Development and Application of Back-Propagation-Based Artificial Neural Network Models in Solving Engineering

Saleh Mohammed Al-Alawi

Department of Electrical Engineering, College of Engineering, Sultan Qaboos University, P.O. Box 33, Al Khod 123, Muscat, Sultanate of Oman.

ABSTRACT: Artificial Neural Networks (ANNs) are computer software programs that mimic the human brain’s ability to classify patterns or to make forecasts or decisions based on past experience. The development of this research area can be attributed to two factors, sufficient computer power to begin practical ANN-based research in the late 1970s and the development of back-propagation in 1986 that enabled ANN models to solve everyday business, scientific, and industrial problems. Since then, significant applications have been implemented in several fields of study, and many useful intelligent applications and systems have been developed. The objective of this paper is to generate awareness and to encourage applications development using artificial intelligence-based systems. Therefore, this paper provides basic ANN concepts, outlines steps used for ANN model development, and lists examples of engineering applications based on the use of the back-propagation paradigm conducted in Oman. The paper is intended to provide guidelines and necessary references and resources for novice individuals interested in conducting research in engineering or other fields of study using back-propagation artificial neural networks.

KEYWORDS: Artificial Neural Network Applications, Engineering Problems, Back-Propagation, Forecasting, Classification, Oman.

1. Introduction

Artificial Neural Networks (ANNs) is a research area that has evolved from Artificial Intelligence (AI) research. Artificial Intelligence, on the other hand, is a branch of computer science. This branch is concerned with designing computer systems that exhibit characteristics associated with intelligent human behavior. Artificial Intelligence research is based on many interrelated sciences and technologies such as engineering, management science, computer science, psychology, philosophy, and linguistics, and covers a wide range of applications. In addition, ANNs and AI provide the scientific foundation for many other growing commercial technologies such as machine learning, expert systems, natural language processing, computer vision and robotics, speech recognition systems, automatic programming, and computer-aided instructions.
ANNs are computer programs that are trained in order to recognize both linear and nonlinear relationships among the input and the output variables in a given data set. In general, ANN applications in engineering have received wide acceptance. The popularity and acceptance of this technique stems from ANNs features that are particularly attractive for data analysis. These features include handling of fragmented and noisy data; speed inherent to parallel distributed architectures, generalization capability over new data, ability to effectively incorporate a large number of input parameters, and its capability of modeling nonlinear systems. Due to these distinctive features, Artificial Neural Networks are used to add intelligent capabilities to computer systems. ANN models allow computer systems to process and recognize different voices, read in text, recognize and classify objects, sense the environment and control robotic movements, predict future trends, and even decide whether to grant a bank loan to a specific customer or not.

2. Artificial Neural Network Concepts

2.1 Back-Propagation Paradigm

One of the most common and frequently used ANN paradigms is the Back-propagation paradigm (Simpson, 1990). This supervised learning method was developed by Rumelhart based on the generalization of the least mean square error (LMS) algorithm. The Back-propagation algorithm uses the gradient descent search technique to minimize a cost function equal to the mean square difference between the desired and the actual net output. The network is trained by selecting small random weights and internal threshold, and then presenting all training data repeatedly by using the supervised training technique. The weights are changed until the network reaches the desired error level or the cost function is reduced to an acceptable value.

2.2 ANN Architecture

The major building block for any ANN architecture is the processing element or neuron. These neurons are located in one of three types of layers: the input layer, the hidden layer, or the output layer. The input neurons receive data from the outside environment, the hidden neurons receive signals from all of the neurons in the preceding layer, and the output neurons send information back to the external environment. These neurons are connected together by a line of communication called connection. Stanley (1990) indicated that the way in which the neurons are connected to each other in a network typology has a great effect on the operation and performance of the network. ANN models come in a variety of typologies or paradigms. Simpson (1990) provides a coherent description of 27 different popular ANN paradigms and presents comparative analyses, applications, and implementations of these paradigms.

2.3 ANN Operations

In the back-propagation (BP) architecture, shown in Figure 1, each element or neuron receives input from the real-world environment or from other processing elements, processes this input, and produces a specific output. Generally, many of these processing elements perform their operations at the same time. This parallelism is a unique feature of the ANN that distinguishes it from the serial processing that is usually performed by conventional computer systems. Each neuron has a straightforward assignment. Input coming to the neuron is associated with a weight indicating its strength. In the neuron, the values of the input are multiplied by the corresponding weights and all products are added to obtain a net value (net). After summation, the net input of the neurons is combined with the previous state of the neurons to produce a new activation value. Whether the neurons fire or not will depend on the magnitude of this value. The activation is then passed through an output or a transfer function (f) that generates the actual neuron output. The transfer function modifies the value of the output signal. This function can be either a simple threshold function that only produces output if the combined input is greater than the threshold value, or it can be a continuous function that changes the output based on the strength of the combined input.
Typical transfer functions employed in building ANN applications include a linear threshold transfer function, step function, sigmoid function, and others.

\[ Net_i = \sum_{j=1}^{n} W_{ij} O_j \]

**Figure 1.** A typical back-propagation architecture.

### 2.4 ANN Training

The first and the most critical step in developing an effective ANN model is input and output definition and data preparation. This includes identifying variables of interest, gathering the relevant data and inspecting them for possible errors, missing values, and outliers. Data accuracy is vital for the development of an efficient model that can provide accurate prediction. If incorrect or erroneous data are fed to the model, this will result in incorrect prediction. As the saying goes, "garbage in, garbage out".

Once the ANN model architecture is defined, data are collected and fed to the model. The network is then trained to recognize the relationships between the input and output parameters. The BP algorithm uses the supervised training technique. In this technique, the interlayer connection weights and the processing elements’ thresholds are first initialized to small random values. The network is then presented with a set of training patterns, each consisting of an example of the problem to be solved (the input) and the desired solution to this problem (the output). These training patterns are presented repeatedly to the ANN model, and the error between actual and predicted results is calculated. Weights are then adjusted by small amounts that are dictated by the General Delta Rule (Rumelhart et al., 1988). This adjustment is performed after each iteration whenever the network’s computed output is different from the desired output. This process continues until weights converge to the desired error level or the output reaches an acceptable level. Simpson (1990) describes the system of equations that provides a generalized description of how the learning process is performed by the BP algorithm.

### 2.5 ANN Testing and Validation

The ANN model can sometimes learn different features other than the relationships in the data. It also can memorize the data or part of this data without learning the relationships between variables or trends in the data. Hence, to insure network accuracy and the generalization capability, the network must be tested on a continuous basis and should be monitored during the training and testing operations. The testing operation involves passing a separate testing set to the trained ANN model and recording the results. These results are compared to actual results. The trained model is assumed to be successful if the model gives good results for that test set. To insure that ANN models provide correct prediction or classifications, the prediction results produced by ANN
3. Developing an Artificial Neural Network Model

**Step #1:** ANN model development starts by first conducting a feasibility study and validating the proposed application. Bailey and Thompson (1990) pointed out some common characteristics of a successful neural network application. They suggested that the application must be data-intensive and dependent upon multiple interacting parameters. The problem area should be rich in historical data or examples. The data set available may be incomplete, contain errors, and describe specific examples. The discriminator or function to determine solutions is unknown or expensive to discover, and the problem should require qualitative or complex quantitative reasoning. Once the application is judged to be feasible and valid, resource constraints (time, equipment, money) should be evaluated. Data, sources, and solution requirements should also be identified and appropriate data should be secured.

**Step #2:** The next step in the ANN development process is data preparation and training. The ability of the ANN to effectively learn the training set and provide accurate results is dependent upon the data preparation activity. Data preparation for modeling can be broadly classified into three distinct areas: data specification, in which variables of interest are identified and collected; data inspection, in which data is examined and analyzed; and data pre-processing, in which some data may be restructured or transformed to make it more useful.

Data specification involves two primary activities: variable selections and determining data sources. For example, there are many social, economical, and weather variables that could possibly affect the demand forecast. A wish list of these variables for model building should be generated by the planner through scanning available literature, consulting experts in the area in question, and by conducting brainstorming sessions with colleagues. Variables in this list should then be examined to assess whether historical data for such variables are available or not. For variables with readily available historical data, data sources should be identified and the data should be collected. Once data for a set of candidate variables are collected, data analysis should then be used to weed out the potential input variables from the wish list generated so that only the most relevant variables are used to develop the forecasting model. To do this, several statistical methods are available for determining linear significance of variables. Some of the more popular statistical techniques used are the coefficient of correlation (R), the coefficient of determination $R^2$, and the ordinary least squares (OLS) regression analysis. A detailed discussion of these statistical techniques is beyond the scope of this article, but treatments of these techniques can be found in Bunn and Farmer (1985), Mendenhall and Beaver (1994), and Burden and Fairies (1985).

As many people discover when they try to model real-world problems and processes, clean data is a luxury that is all too rare. Data collected from the different sources are generally noisy, contain many gaps and outliers, and are poorly distributed. These issues, if not properly addressed prior to the model's development, could lead to inaccurate and unreliable prediction. Collected data, hence, should be inspected well and analysed carefully. The first step in data inspection is to examine individual variables for erroneous values and to remove these values from the data set after careful analysis and only if these values prove to be erroneous. Each variable should also be inspected for outliers as well as missing data.

Once the most significant input variables are selected and carefully inspected, the forecaster should then examine the distribution of each of these variables. The shape of the distribution will indicate to the planner whether a particular variable needs data pre-processing. Data pre-processing may involve any mathematical operations. Common techniques include calculating sums, differences, differentials, inverses, powers, roots, averages, etc. Anderson (1990) and Lawrence (1991) provide a detailed description on how to perform this important task.

**Step #3:** The third step in the development process is the selection of an appropriate neural network paradigm. ANN models come in a variety of typologies or paradigms. Simpson (1990) provides a coherent description of 27 different popular ANN paradigms and presents comparative...
analyses, applications, and implementations of these paradigms. The selection among these paradigms should be based on the application requirements and the available neural network software containing the specific paradigm. Some of the factors that should be considered in the selection process include neural network size, required output type, method of training, and the time available for model development and testing.

**Step #4:** Having selected the appropriate paradigm, the fourth step is to determine the network's architecture design and to select its parameters. This process involves the selection of the number of input nodes, hidden nodes, and output nodes. In addition, it also involves the selection of the network parameters such as the transfer function, learning algorithm, learning rate, momentum, and learning threshold.

**Step #5:** The next step in this process is training the model. Training involves presentation of the training set to the network and periodical monitoring of the network's performance. This is accomplished automatically by the appropriate Back-propagation simulation software that was selected in step #3. Based on the user choice, training cases can be presented to the network either sequentially or by following a random process. During the training process, one or several of the network parameters are changed to improve the network's performance. This process continues until weights converge to the desired error level or the output reaches an acceptable level.

**Step #6:** After training is complete, testing and validation is the final step in the development process. It is important to test the resulting ANN model against both the training set and the test set. The test set should contain examples of input vectors that the network did not encounter previously. This test is a benchmark that determines how well and how accurate the trained network is performing. Model validation, on the other hand, deals with comparing the results of the developed model to results obtained from common or classical models or techniques being used by the industry. The training, testing and validation process is explained in the NeuroShell™ simulation package (1991).

4. **An Example of an ANN Application**

The following is an illustrative example of an ANN application to forecast electrical demand for a Muscat power system. Detailed steps of this example can be found in Islam et al. (1995).

For any utility, medium-term load and energy forecasting is useful in planning fuel procurement, reserve margin, scheduling unit maintenance, diversity interchange, and system expansion planning. This type of forecast is normally prepared for the range of one to five years.

5. **Problem Definition and its Importance**

According to the historical monthly peak load and energy data collected from 1986 to 1992, the system's load and energy consumption appeared to be more or less cyclic, keeping in harmony with temperature, which varies from an extreme maximum of 48°C in summer to an extreme minimum of 10°C in winter. Load demand, therefore, varies considerably from hour to hour. It is apparent that the electrical load and energy consumption pattern of this power system depends heavily on weather. On the other hand, the growth in load and energy demand depends largely on the number of consumers connected to the system. Variables such as temperature, humidity, wind speed, number of connections, and other variables can be used to develop load and energy models for the Muscat Power System. In fact, the number of such variables is large and, depending on the type and nature of the forecast, should be carefully selected. The selection criteria could be based on human intuition, knowledge and experience and should be validated using statistical techniques to determine their contribution and correlation to the load or energy.

6. **Data Requirements and Data Processing**

In medium-term load forecasting, generally, there are two forecasts that are prepared. These are the load forecast and the energy forecast. To develop these forecasts, the following variables
were identified using human intuition, brainstorming sessions, and consultation with experts in the area: Absolute Maximum Temperature ($T_{\text{max}}$), Average Maximum Temperature ($T_{\text{avmax}}$), Average Maximum Relative Humidity ($\text{RH}_{\text{avmax}}$), Average Relative Humidity ($\text{RH}_{\text{av}}$), Wind Speed ($W$), Duration of Bright Sunshine ($S$), Global Radiation ($R$), Precipitation ($\text{PR}$), Vapor Pressure ($\text{VP}$), Degree Days ($\text{DD}$), Comfort Index ($\text{CI}$), and Number of Connections ($\text{CON}$). Data for all of the above variables, with the exception of CI and CON, were collected from the historical records of the Ministry of Housing Electricity and Water and from the records provided by an automated weather station, while CI and CON are processed variables (Bunn and Farmer, 1985). The collected data were then examined to remove errors and outliers and to replace missing values. In addition, a correlation analysis was performed to select the appropriate variables suitable for the load and energy models. As a result of this test, the variables PR and VP were eliminated from the energy model because of their low correlation and contribution. It was also interesting to find that although some variables like $W$ and $\text{RH}_{\text{max}}$ had strong correlation to monthly electrical energy consumption, they had very little correlation to monthly peak load.

7. Developing the Model

Using the historical data for the selected variables, monthly peak load and energy consumption forecasting models were developed using an artificial neural network (ANN) simulation package, NeuroShell™ (1991). For the Energy Model, monthly data from 1986 to 1990 were used in model development. For the Load Model, the input variables selected were $T_{\text{pkld}}$, $\text{RH}_{\text{pkld}}$, $T_{\text{max}}$, and $\text{CON}$. The monthly historical data used for developing the model also covered the same periods as in the energy model. Similarly, in validating the ANN models' results, Socio-Economic models (Barakat and Al Rashad, 1993) were also generated using the same variables and historical data. These models were particularly selected for comparison since they have demonstrated giving better accuracy than the Box and Jenkins models, and they are more suited to high growth systems such as the Muscat Power System.

8. Validation of Results

To test and validate the forecasts generated by Socio-Economic (SE) models and ANN models, monthly historical data for 1991 and 1992 were used to test these models' prediction capabilities. The resulting forecasts were then compared to the actual results, and statistical numerical measures were then calculated. For the SE models, the mean absolute percentage error (MAPE) for the energy model was approximately 10.969 while the load model was 10.786. The testing set $R^2$ was 0.946 and 0.719 for the two models, respectively. In comparison, the ANN-based energy model's MAPE was 1.787 and the load model was 1.870. The testing set $R^2$ was 0.996 and 0.989, respectively. Table 1 shows the comparison of these results. The monthly actual and forecasted energy consumption and peak load for 1991 and 1992 are shown in Figures 2 and 3. From this Figure, we can see that the SE models as well as the other models did not provide highly accurate results as the ANN models did.

<table>
<thead>
<tr>
<th>Technique</th>
<th>ME</th>
<th>MAD</th>
<th>MAPE</th>
<th>$R^2$</th>
<th>Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy Model (Socio-Economic)</td>
<td>18638.17</td>
<td>25783.50</td>
<td>10.969</td>
<td>0.946</td>
<td>89.03</td>
</tr>
<tr>
<td>Energy Model (ANN)</td>
<td>489.67</td>
<td>4839.83</td>
<td>1.787</td>
<td>0.996</td>
<td>98.21</td>
</tr>
<tr>
<td>Load Model (Socio-Economic)</td>
<td>21.944</td>
<td>51.53</td>
<td>10.786</td>
<td>0.719</td>
<td>89.21</td>
</tr>
<tr>
<td>Load Model (ANN)</td>
<td>-1.846</td>
<td>1.11</td>
<td>1.870</td>
<td>0.989</td>
<td>98.13</td>
</tr>
</tbody>
</table>
Where ME=Mean Error, MAD=Mean Absolute Deviation, MAPE=Mean Absolute Percentage Error.

Figure 2. Actual and forecasted results for the load model.

Figure 3. Actual and forecasted results for the energy consumption model.

9. Engineering Applications

A clear guideline on how the steps are implemented to develop ANN models for different engineering applications can also be found in the following ANN application papers in different engineering fields. The author and his colleagues wrote the ANN application papers.

10. Electrical Engineering
1. Short-term Load Forecasting Using Artificial Neural Networks (Al-Alawi and Islam, 1995).
4. Tuning of SVC Damping Controllers Over Wide Range of Load Models Using Artificial Neural Network (Ellithy and Al-Alawi, 2000).
5. Tuning Power System Stabilizers over a Wide Range of Load Models Using Artificial Neural Networks (Ellithy et al, 1997).
6. ANN-Based Load Identification and Control of AC Voltage Regulator (Gastli et al 2000).
7. Statistical Signal Characterization-Artificial Neural Network Based Hybrid System for Electrocardiogram Interpretation (Al-Alawi et al 1998).

11. Mechanical Engineering

1. Experimental Investigation and Failure Analysis of Fastened GRP Under Bending Using the Finite Element Method and Artificial Neural Network Modeling (Seibi and Al-Alawi, 1999).
3. Prediction of Fracture Toughness Using Artificial Neural Networks (Seibi and Al-Alawi, 1997).
5. Design of Fiberglass/Copper Moulds Using Finite Element Analysis (Seibi and Al-Alawi, 1999).

12. Petroleum and Mineral Resources Engineering

2. Establishing PVT Correlation for Omani Oils (Boukadi et al 1998).
6. Analysis and Prediction of Oil Recovery Efficiency in Limestone Cores Using Artificial Neural Networks (Boukadi and Al-Alawi, 1997).
8. Preliminary Studies on Using Artificial Neural Networks to Predict Sedimentary Facies of the Petro-Carboniferous Glacigenic Al Khlata Formation in Oman (Schuniker et al 1999).
10. The Application of Artificial Neural Networks to Reservoir Engineering (Kalam et al 1995).

13. Civil Engineering


14. Other Applications

1. Forecasting Fish Exports in the Sultanate of Oman Using Artificial Neural Networks (Luqman and Al-Alawi 2000).

15. Conclusion

In the past ten years, the international community has given considerable attention to developing more accurate systems and models based on Artificial Intelligence techniques such as artificial neural networks, expert systems and fuzzy logic. These techniques have successfully been applied in a variety of fields reporting higher accuracy compared to other classical models and methods.

This paper provides basic ANN concepts, outlines steps used for ANN model development, and lists examples of ANN-based engineering applications conducted in Oman. The paper is intended to provide guidelines and necessary references and resources for individuals interested in conducting research in engineering or other fields of study using back-propagation artificial neural networks. It is recommended; therefore, to explore, learn and use such advanced techniques in order to solve some of the engineering problems and to survive current economic conditions.

References

SALEH MOHAMMED AL-ALAWI


Appendix A: Artificial Neural Networks Resources

Books, journals and other resources on ANNs included in this guide are obtained from the following WebPages:


Please see the above WebPages for more detailed information and full review.

Some of the best popular introduction to ANNs:


Some of the best introductory books for business executives:

Some of the best books on using and programming ANNs:


Some of the best intermediate textbooks on ANNs:


Some of the best books on neurofuzzy systems:


Some of the best comparison of ANNs with other classification methods:


Other notable books:

Feedforward networks:


Time-series forecasting:


Fuzzy logic and neurofuzzy systems:


Optimization:


General Books for the Beginner:

Not-quite-so-introductory literature:


Subject: Journals and magazines about Neural Networks:


Title: Neural Computation, Publish: MIT Press Address: MIT Press Journals, 55 Hayward Street Cambridge, MA 02142-9949, USA, Phone: (617) 253-2889 Freq.: Quarterly (vol. 1 in 1989) Cost/Yr: Individual $45, Institution $90, Students $35; Add $9 Outside USAISSN #: 0899-7667URL: [http://mitpress.mit.edu/journals-legacy.tcl](http://mitpress.mit.edu/journals-legacy.tcl)

Title: NEURAL COMPUTING SURVEYS Publish: Lawrence Erlbaum Associates Address: 10 Industrial Avenue, Mahwah, NJ 07430-2262, USA. Freq.: Yearly Cost/Yr: Free on-line, ISSN #: 1093-7609URL: [http://www.icsi.berkeley.edu/~jagota/NCS/](http://www.icsi.berkeley.edu/~jagota/NCS/)

Title: IEEE Transactions on Neural Networks, Publish: Institute of Electrical and Electronics Engineers (IEEE) Address: IEEE Service Center, 445 Hoes Lane, P.O. Box 1331, Piscataway, NJ, 08855-1331 USA. Tel: (201) 981-0060 Cost/Yr: $10 for Members belonging to participating IEEE societies, Freq.: Quarterly (vol. 1 in March 1990) URL: [http://www.ieee.org/nnc/pubs/transactions.html](http://www.ieee.org/nnc/pubs/transactions.html)


Title: **International Journal of Neural Networks**, Publish: Learned Information Freq.: Quarterly (vol. 1 in 1989), Cost/Yr: 90 pounds, ISSN #: 0954-9889

**Subject: Conferences and Workshops on Neural Networks:**
The journal "Neural Networks" has a list of conferences, workshops and meetings in each issue. It is also available from [http://www.ph.kcl.ac.uk/neuronet/bakker.html](http://www.ph.kcl.ac.uk/neuronet/bakker.html).


Conferences, workshops, and other events concerned with neural networks, inductive learning, genetic algorithms, data mining, agents, applications of AI, pattern recognition, vision, and related fields are listed at Georg Thimm's web page [http://www.drc.ntu.edu.sg/users/mgeorg/enter.epl](http://www.drc.ntu.edu.sg/users/mgeorg/enter.epl)

**Subject: Freeware, Shareware, and Commercial Software for Neural Networks:**
- Several source codes for ANN in C/C++ and Java
- A review of 44 freeware and shareware packages for ANN simulation
- A review of 40 commercial software packages for ANN simulation

Please see the above WebPages for detailed information or for full review: