# Artificial Neural Networks

8

Learning Objectives

* Understand what are ANNs and why they are useful
* Know the design principles of ANN
* Learn about representation of the elements of an ANN
* Know the many architectures of ANNs
* Understand how ANNs are developed and trained
* Appreciate the many advantages and disadvantages of ANNs

### INTRODUCTION

Artificial Neural Networks (ANNs) are inspired by the information processing model of the brain. The human brain consists of billions of neurons that link with one another in an intricate pattern. Every neuron receives information from many other neurons, processes it, gets excited or not, and passes its state information to other neurons.

Just like the brain is a multipurpose system, so also the ANNs are very versatile systems. They can be used for many kinds of pattern recognition and prediction. They are also used for classification, regression, clustering, association, and optimization activities. They are used in finance, marketing, manufacturing, operations, information systems applications, and so on.

ANNs are composed of a large number of highly interconnected processing elements (neurons) working in multilayered structures that receive inputs, process the inputs, and produce an output. An ANN is designed for a specific application, such as pattern recognition or data classification, and trained through a learning process. Just like in biological systems, ANNs make adjustments to the synaptic connections with each learning instance.

ANNs are like a black box trained into solving a particular type of problem, and they can develop high predictive powers. Their intermediate synaptic parameter values evolve as the system obtains feedback on its predictions, and thus an ANN learns from more training data (Figure 8.1).

Input

Output

Artificial Neural Network (ANN)

FIGURE 8.1 General ANN Model

#### Caselet: IBM Watson – Analytics in Medicine

*The amount of medical information available is doubling every five years and much of this data is unstructured. Physicians simply don’t have time to read every journal that can help them keep up to date with the latest advances. Mistakes in diagnosis are likely to happen and clients have become more aware of the evidence. Analytics will transform the field of medicine into evidence-based medicine. How can healthcare providers address these problems?*

*IBM’s Watson cognitive computing system can analyze large amounts of unstructured text and develop hypotheses based on that analysis. Physicians can use Watson to assist in diagnosing and treating patients. First, the physician might describe symptoms and other related factors to the system. Watson can then identify the key pieces of information and mine the patient’s data to find relevant facts about family history, current medications and other existing conditions. It combines this information with current findings from tests, and then forms and tests a hypothesis by examining a variety of data sources—treatment guidelines, electronic medical record data and doctors’ and nurses’ notes, as well as peer-reviewed research and clinical studies. From here, Watson can provide potential treatment options and its confidence rating for each suggestion.*

*Watson has been deployed at many leading healthcare institutions to improve the quality and efficiency of healthcare decisions, to help clinicians uncover insights from its patient information in electronic medical records (EMR), among other benefits.*

1. *How would IBM Watson change medical practices in the future?*
2. *In what other industries and functions could this technology be applied?*

### BUSINESS APPLICATIONS OF ANN

Neural networks are used most often when the objective function is complex, and where there exists plenty of data and the model is expected to improve over a period of time. A few sample applications are as follows

* They are used in stock price prediction where the rules of the game are extremely complicated, and a lot of data needs to be processed very quickly.
* They are used for character recognition, as in recognizing hand-written text, or damaged or mangled text. They are used in recognizing fingerprints.

These are complicated patterns and are unique for each person. Layers of neurons can progressively clarify the pattern leading to a remarkably ac- curate result.

* They are also used in traditional classification problems, like approving a financial loan application.

### DESIGN PRINCIPLES OF AN ARTIFICIAL NEURAL NETWORK

1. A neuron is the basic processing unit of the network. The neuron (or processing element) receives inputs from its preceding neurons (or PEs), does some nonlinear weighted computation on the basis of those inputs, transforms the result into its output value, and then passes on the output to the next neuron in the network (Figure 8.2). x’s are the inputs, w’s are the weights for each input, and y is the output.

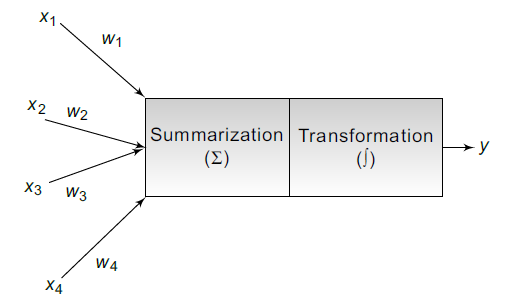
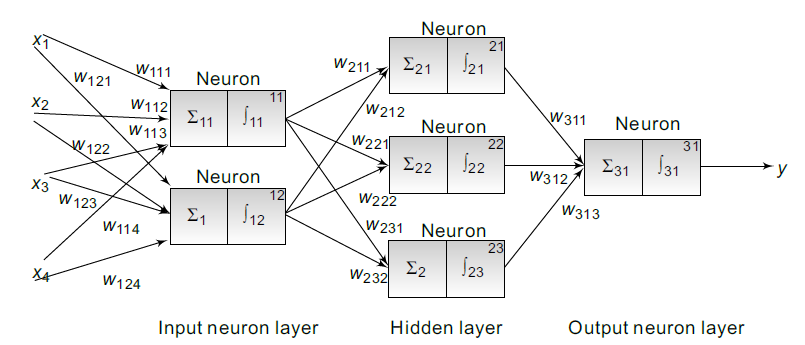


FIGURE 8.2 Model for a Single Artificial Neuron

1. A neural network is a multilayered model. There is at least one input neuron, one output neuron, and at least one processing neuron. An ANN with just this basic structure would be a simple, single-stage computational unit. A simple task may be processed by just that one neuron and the result may be communicated soon. ANNs however, may have multiple layers of processing elements in sequence. There could be many neurons involved in a sequence depending upon the complexity of the predictive action. The layers of PEs could work in sequence or in parallel (Figure 8.3).
2. The processing logic of each neuron may assign different weights to the various incoming input streams. The processing logic may also use nonlinear transformation, such as a sigmoid function, from the processed values to the output value. This processing logic and the intermediate weight and processing functions are just what works for the system as a whole, in its

FIGURE 8.3 Model for a Multilayer ANN

objective of solving a problem collectively. Thus, neural networks are considered to be an opaque and a black-box system.

4. The neural network can be trained by making similar decisions over and over again with many training cases. It will continue to learn by adjusting its internal computation and communication based on feedback about its previous decisions. Thus, the neural networks become better at making a decision as they handle more and more decisions.

Depending upon the nature of the problem and the availability of good training data, at some point, the neural network will learn enough and begin to match the predictive accuracy of a human expert. In many practical situations, the predictions of ANN, trained over a long period of time with a large number of training data, have begun to decisively become more accurate than human experts. At that point, ANN can begin to be seriously considered for deployment in real situations in real time.

### REPRESENTATION OF A NEURAL NETWORK

A neural network is a series of neurons that receive inputs from other neurons. They do a weighted summation function of all the inputs, using different weights (or importance) for each input. The weighted sum is then transformed into an output value using a transfer function.

Learning in ANN occurs when the various processing elements in the neural network adjust the underlying relationship (weights, transfer function, etc.)

between input and outputs, in response to the feedback on their predictions. If the prediction made was correct, then the weights would remain the same, but if the prediction was incorrect, then the parameter values would change.

The Transformation (Transfer) Function is any function suitable for the task at hand. The transfer function for ANNs is usually a nonlinear sigmoid function. Thus, if the normalized computed value is less than some value (say, 0.5) then the output value will be zero. If the computed value is at the cut-off threshold, then the output value will be 1. It could be a nonlinear hyperbolic function in which the output is either –1 or 1. Many other functions could be designed for any or all of the processing elements.

Thus, in a neural network, every processing element can potentially have a different number of input values, a different set of weights for those inputs, and a different transformation function. Those values support and compensate for one another until the neural network as a whole learns to provide the correct output, as desired by the user.

### ARCHITECTING A NEURAL NETWORK

There are many ways to architect the functioning of an ANN using fairly simple and open rules with a tremendous amount of flexibility at each stage. The most popular architecture is a feedforward, multilayered perceptron with back-propagation learning algorithm. That means there are multiple layers of PEs in the system and the output of neurons are fed forward to the PEs in the next layers; and the feedback on the prediction is fed back into the neural network for learning to occur. This is essentially what was described in the earlier paragraphs. ANN architectures for different applications are shown in Table 8.1.

Table 8.1 ANN Architecture for Different Applications

|  |  |
| --- | --- |
| Classification | Feedforward networks (MLP), radial basis function and proba- bilistic |
| Regression | Feedforward networks (MLP), radial basis function |
| Clustering | Adaptive resonance theory (ART), Self-organizing maps (SOMs) |
| Association Rule Mining | Hopfield networks |

### DEVELOPING AN ANN

It takes resources, training data, skill and time to develop a neural network. Most data mining platforms offer at least the MultiLayerPerceptron (MLP) algorithm to implement a neural network. Other neural network architectures include probabilistic networks and self-organizing feature maps.

The steps required to build an ANN are as follows

1. Gather data and divide into training data and test data. The training data needs to be further divided into training data and validation data.
2. Select the network architecture, such as Feedforward network.
3. Select the algorithm, such as Multi Layer Perception.
4. Set network parameters.
5. Train the ANN with training data.
6. Validate the model with validation data.
7. Freeze the weights and other parameters.
8. Test the trained network with test data.
9. Deploy the ANN when it achieves good predictive accuracy.

Training an ANN requires the training data be split into three parts (Table 8.2)

Table 8.2 ANN Training Datasets

|  |  |
| --- | --- |
| Training Set | This dataset is used to adjust the weights on the neural network (~ 60%). |
| Validation Set | This dataset is used to minimize overfitting and verifying accuracy (~ 20%). |
| Testing Set | This dataset is used only for testing the final solution in order to confirm the actual predictive power of the network (~ 20%). |
| *k*-fold Cross Validation | This approach means that the data is divided into *k* equal pieces, and the learning process is repeated *k* times with each piece be- coming the training set. This process leads to less bias and more accuracy, but is more time consuming. |

### ADVANTAGES AND DISADVANTAGES OF USING ANNs

There are many benefits of using ANN. Some are given below

* ANNs impose very little restrictions on their use. ANN can deal with (identify/model) highly nonlinear relationships on their own, without much work

from the user or analyst. They help find practical data-driven solutions where algorithmic solutions are nonexistent or are too complicated.

* There is no need to program neural networks as they learn from examples. They get better with use, without much programming effort.
* They can handle a variety of problem types, including classification, clustering, associations, etc.
* ANNs are tolerant of data quality issues and they do not restrict the data to follow strict normality and/or independence assumptions.
* They can handle both numerical and categorical variables.
* ANNs can be much faster than other techniques.
* Most importantly, they usually provide better results (prediction and/or clustering) compared to statistical counterparts, once they have been trained enough.

The key disadvantages arise from the fact that they are not easy to interpret or explain or compute.

* They are deemed to be black-box solutions, lacking explainability. Thus, they are difficult to communicate about, except through the strength of their results.
* Optimal design of ANN is still an art. It requires expertise and extensive experimentation.
* It could be difficult to handle a large number of variables (especially the rich nominal attributes).
* It takes large datasets to train an ANN.

## Conclusion

Artificial neural networks are complex systems that mirror the functioning of the human brain. They are versatile enough to solve many data mining tasks with high accuracy. However, they are like black boxes and they provide little guidance on the intuitive logic behind their predictions.

## Questions

1. What is a neural network? How does it work?
2. Compare a neural network with a decision tree.
3. What makes a neural network versatile enough for supervised as well as nonsupervised learning tasks?

4. Examine the steps in developing a neural network for predicting stock prices. What kind of objective function and what kind of data would be required for a good stock price predictor system using ANN?

## True/False

1. ANN is a machine learning technique.
2. ANNs are like a black box trained into solving a particular type of problem.
3. ANNs can be used for Classification, Clustering, Association Rules, and Optimization activities.
4. The processing logic of each neuron may assign different weights to the various incoming input streams.
5. A neural network is a multilayered model.
6. An ANN should have at least one input neuron, one output neuron, and at least three processing neurons.
7. ANN can deal with (identify/model) nonlinear relationships but with much work from the user or analyst.
8. Learning in ANN occurs when the various processing elements in the neural network adjust the underlying relationship (weights, transfer function, etc.) between inputs and outputs.
9. The most popular ANN architecture is self-organizing maps based learning algorithm.
10. IBM Watson is a form of a neural network.