## Market modelling · 3
# Neural Networks

Chuck Chakrapani

As we noted earlier, there are many kinds of models. In this series we will look at a few of these models.

A model that is currently in vogue is known as the Neural Network. The Neural Network model is not new. It is more than 50 years old, although it is currently enjoying wide popularity. The Neural Network model is particularly suited to large databases, although its applicability can be much wider than that.

To understand Neural Networks, we may want to start with a review of Multiple Regression Analysis. In multiple regression analysis, we aim to predict the effect of a number of variables (such as product ratings) on a single variable (such as propensity to buy). The resultant equation is of the form:

Propensity to buy =
b1 * Price
b2 * Durability
b3 * Appearance

and so on.

In statistical notation it is written as:
y = b1x1 + b2x2+b3x3+...bnxn

The 'b' coefficients are the derived weights assigned to each attribute. Such weights are generated by the technique using the data which is being analyzed. It is helpful to remember the structure of the regression equation when discussing Neural Networks.

**The context**
Let us consider the example of a financial organization that has mailed out a promotional mailing for an investment type product to 10,000 of its customers. If 1,500 of these people respond to the promotional offer, they can be so identified along with the other relevant details such as the size of purchase.

The sample of 10,000 at this stage is divided into two parts: one part is used to develop the model and the other to validate the model (the holdout sample).

**Input variables**
We may want to start with the attributes that are available for each customer on the data base. These could be attributes such as:

- How long the person has been a customer
- Size of the account
- Area of dwelling
- Gender
- Income
- Home ownership

**How it works**
A starting weight is assigned to each of these attributes. These attributes are similar to the 'b' weights in a regression equation. The only difference here is that the weights are assigned arbitrarily and not on the basis of any statistical calculations. In other words we calculate 'y' exactly as we did in Multiple Regression:

I = w1x1 + w2x2+w3x3+...wnxn

As you will note, here we use the notation I (Internal Activation Quantity) instead of 'y' and 'w' rather than 'b' to indicate the coefficients. As noted, while 'b' coefficients are estimated from the data, 'w' coefficients are random initial weights that bear no relationship to the data.

The Internal Activation Quantity is subjected to a nonlinear transfer function. This function converts Internal Activation into a number that falls between -1.0 to +1.0.

As with a multiple regression dummy variable, customers who responded are coded as 1 and those who did not as 0. Thus each customer can assume only one of the two values-a 0 or a 1.

The Internal Activation quotient however is free to generate any value between -1.0 and +1.0. The equation is applied to each customer separately. Suppose customer A is a buyer. He is coded as 1. If the generated I is 0.9, then the deviation or 'error' is -0.1. Suppose B is a non-buyer. She is coded as 0. If the generated I is 0.8, then the deviation or 'error' is +0.8. Because each individual's characteristics (x's) are different, the same equation will produce different errors for different customers. Again this is akin to applying the regression equation to each individual. For some the equation will fit well and for others less well.

The purpose of the neural network is to iteratively work on the coefficients such that the overall error is minimized.

So, in its basic structure, Neural Networks resemble multiple regression equations. Neural Networks iteratively refine the coefficients while multiple regression derives the coefficients based on some criterion such as the method of least squares.

**The hidden layer**
If that was all there was to Neural Networks, there wouldn't be much different from the multiple regression analysis.

Even the iterative processes is not a novelty to marketing researchers. Other techniques such as conjoint analysis use iterative processes to reproduce the rankings given by consumers.

There is obviously more to Neural Networks. Consider that what we have discussed so far is emanating from one node. If there are many such nodes, these nodes themselves become the input for the next level of Neural Network Analysis.

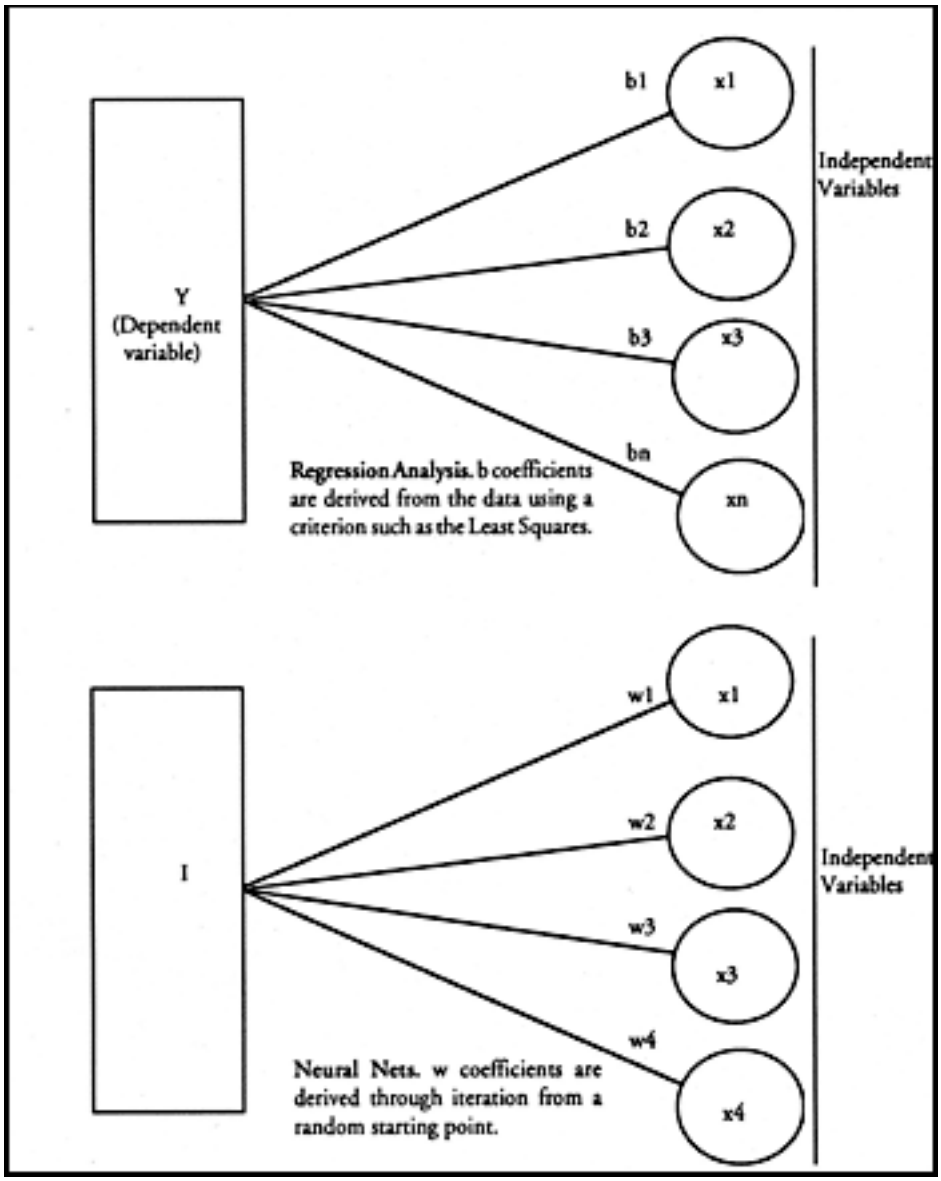
In fact, Neural Networks can be multilayered with each layer affecting the layer above it.

As an example, suppose *Networth* and *Socioeconomic* status influence purchase decisions. These two 'nodes' are not directly measured but are derived from other input variables (see diagram on last page). In this instance the nodes are estimated by the input variables and buying behaviour in turn is estimated by the nodes. There can be several such hidden layers.

**How many iterations?**
The technique attempts to get the I of buyers as close to 1 as possible and the I of non-buyers as close to 0 as possible.

But perfect convergence to these values is seldom possible. In theory, such iterations can go on forever, especially when there is no convergence. However, many programs that handle Neural Network problems are given an upper limit of iterations (which could be as high as 200,000) and reasonable convergence can be obtained in less than an hour on a high speed desktop computers.

**Multiple Regression and Neural Nets: Structural Similarities**




**Validating the model**

Even when there is reasonable convergence, we cannot be sure that the model is correct. This is because the model is not validated on any other set of data.

This is the reason we set aside a holdout sample. We first develop the model using half the sample. Now we take the results of that model and apply them to the holdout sample.

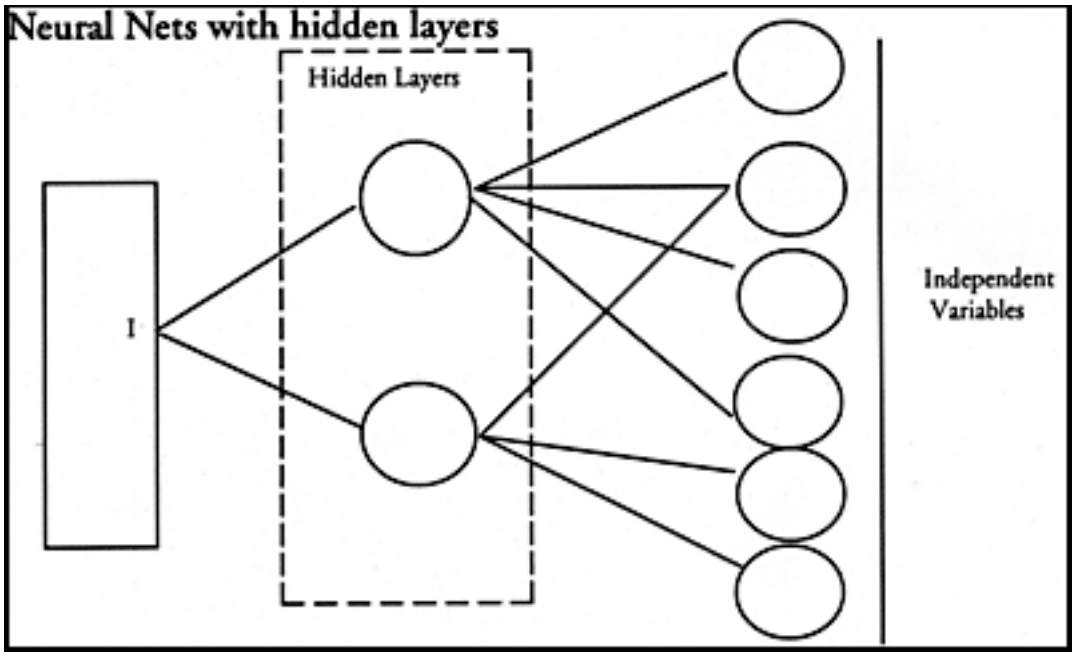
A score is arrived at for each individual which ranges from .000 to .999 which in fact is the probability of a 'correct prediction'. (In practice, these numbers are multiplied by 1,000 and the probability of .990 will simply be referred to as 990.)

Now these predictions can be checked against the actual results we already have at the aggregate level.

What if the model converges well in the initial sample but does poorly when applied to the holdout sample?

This really means that the model does not work well and cannot be used for further work.

If the model works well on the holdout sample, it would increase our confidence in the model and we may want to apply the model to other customers to assess their propensity to buy.




**Summary of the model**
The steps involved in the Neural Network model can be summarized as follows:

1. Identify the attribute of interest. eg. Buyers and Non-buyers.

2. Identify the possible variables that could contribute to the criterion variable identified in Step (1).

3. Divide the sample into the Model sample and the Holdout sample.

4. Develop the conceptual links as to what attributes (2 above) contribute to what other attributes which in turn contribute to the criterion variable.

5. Develop the model using a suitable computer program.

6. Apply the model to the Holdout sample.

7. If it works with the Holdout sample, extend the model to other data. If it fails re-specify the model and start again.

**Uses of the model**
The uses of these models are fairly obvious. For instance, if we know the probability of a person responding to an offer, we can restrict our promotional activities to people who are likely to respond to our offer rather than extending it to every customer.

But that is only the most obvious of uses. With a bit of creativity, Neural Nets, like many other models, can be applied to a wide variety of problems.

A case in point is the application of Neural Nets to analyze qualitative data. Canadian researchers Karl Moore, Robert Burbach and Roger Heeler report an application inwhich they input a text file through a neural network process to cluster key words. The objective is to arrive at sorting and fine tuning of concepts and understand the intricate interrelationships among them. (I realize that this description is rather sketchy. Interested readers are referred to *"Using Neural Nets to Analyze Qualitative Data"* in the Winter 1995 issue of the *Marketing Research* magazine published by the AMA.)

*Dr Chuck Chakrapani is a Toronto-based consultant, author and speaker who works internationally.*

*permission from the author.*