Available online at www.sciencedirect.com
ScienceDirect
Journal of Consumer Psychology 20 (2010) 381–389
Journal of CONSUMER PSYCHOLOGY

Neural network applications in consumer behavior

Richard Briesch*, Priyali Rajagopal

Cox School of Business, Southern Methodist University, USA

Received 25 March 2010; accepted 7 June 2010
Available online 13 July 2010

Abstract

This article introduces the concepts and terminology of artificial neural networks. The approach is demonstrated on data that represent a domain of interest to the consumer psychologist. ANNs are mathematical models that are commonly used in business applications to model relationships between variables. A key strength of ANNs is their flexibility, i.e. their ability to easily accommodate linear and non-linear relationships without any *a priori* functional form specification. They can easily be used to study topics of interest to consumer psychologists such as persuasion, influence, segmentation etc and can offer distinct advantages over traditional statistical techniques such as ANOVA and regression. We demonstrate the application of ANNs in three different areas: regression, non-linear principal component analysis and classification.

Keywords: Methodology; Regression; Structural equations; Discriminant analysis; Nonparametric

Artificial neural networks (ANNs) were originally developed by researchers in computer science as an analogy to the human brain (cf., Hinton, 1992). Neurons receive inputs from other neurons, combine them, and send signals on to other neurons. Some cognitive sciences use the network analog as a fairly literal model of the phenomena being studied, but other disciplines have begun exploring the utility of the network model as a statistical analytical tool.

In particular, consider that some form of the general linear model—analysis of variance or regression—accounts for much of the findings in many social science literatures, even though researchers know that some phenomena are non-linear (e.g. response to fear appeals, the effects of repetition on attitudes). One of the strengths of ANNs is the ability to easily accommodate linear and non-linear relationships; indeed the researcher need not specify functional forms *a priori*. Distributions do not need to be specified, hence there are fewer opportunities for assumptions, such as multivariate normality, to be violated. Naturally, any modeling approach comes with strengths and weaknesses, and Table 1 summarizes key advantages and disadvantages of ANNs relative to the more frequently used general linear models (e.g. regression). An early disadvantage that had been heavily cited was the requirement of large sample sizes, yet increasingly marketers and consumer psychologists are working with extremely large datasets (e.g., companies' CRM databases, online records and the like).

Broadly speaking, ANNs can be classified along two dimensions: 1) the type of learning algorithms used—where supervised and unsupervised are the two extreme pure forms, and 2) whether the networks are recursive or non-recursive. In this literature, the term "learning" comes from the analogy to the human brain, and refers to the iterative methods used to estimate the parameters of the model.

To elaborate on the first dimension, in "supervised learning," the dependent variable(s) is (are) observed, the learning algorithms define an error function to be minimized (e.g., squared distance between the observed and predicted variables, like in regression) and then a gradient-based optimization technique is used to estimate the parameters. In "unsupervised learning," the dependent variable is not observed (much like the "truth" not being known when conducting a cluster analysis), but the algorithms estimate the parameters to implicitly or explicitly maximize similar objective functions (Cheng & Titterington, 1994).

To understand the difference between recursive and non-recursive networks, consider Fig. 1. It depicts three types of layers (or nodes) within a network: 1) input, 2) hidden and 3) output. There

* Corresponding author.
E-mail addresses: rbriesch@mail.cox.smu.edu (R. Briesch), priyalir@mail.cox.smu.edu (P. Rajagopal).

1057-7408/$ - see front matter
doi:10.1016/j.jcps.2010.06.001

Table 1
Comparison of ANNs with general linear models.

	ANNs	Generalized linear models (ANOVA, Linear Regression, Discriminant analysis)
Advantages	Can easily handle non-linear relationships between dependent and independent variables. There is no need to know the exact form of the non-linearity in advance.	Some models can incorporate non-linear relationships (e.g. use of interaction terms), but requires significant effort. Further, the exact form of the non-linearity needs to be known in advance.
	Can handle interactions between several independent variables easily.	Limited number of interactions between dependent variables can be handled.
	Robust to model misspecifications. Since the network "learns" the exact relationship between the independent and dependent variables through multiple iterations, the mathematical form of the model is not pre-specified.	Sensitive to model misspecifications.
	No significant effect of multicollinearity in independent variables.	Models can be significantly compromised due to multicollinearity of the independent variables.
	More robust to missing data and inaccurate data.	More sensitive to missing data and inaccurate data.
Disadvantages	Lack of theoretical background. ANNs are essentially data fitting and prediction tools; hence they are similar to "black boxes" in that no theoretical framework is required for their working and the results from these models may not be theoretically based.	Theoretical frameworks can be used to specify the model relationships through coefficient restrictions and/or the shape of the response function.
	Possibility of model over-fitting since there are no prior theoretical relationships to function as constraints on model flexibility.	Relatively lower possibility of model over-fitting since the model is constrained by prior theoretical relationships.
	Difficult to interpret the results of a 'black box.' Graphing and 'elasticities' are generally required.	Relationship between independent variable and dependent measure summarized in single coefficient.
	Require many decisions to make (e.g., number of hidden nodes, structure of the network, type of learning algorithm) which can render model specification difficult.	Fewer decisions and hence relatively straightforward to specify models.

can be zero or more hidden layers within a network. The input layer is analogous to independent variables, as this is where they enter the network. In Fig. 1, there are two inputs (x_1 and x_2), so there are two input nodes. The output layer is analogous to the prediction of the network model, whether it is forecasting, cluster membership, classification, probability of occurrence, etc. In Fig. 1, there is a single output variable, y', so there is one output node. Depending upon the application, y' could be the forecasted value of the observed dependent variable y. The arrows in Fig. 1 between the input layer and hidden layer, and between the hidden layer and output layer represent coefficients to be estimated.

The nodes which are not input or output nodes are the hidden nodes. Hidden nodes are similar to the latent or unobservable variables that are often related to consumer attitudes or behavior

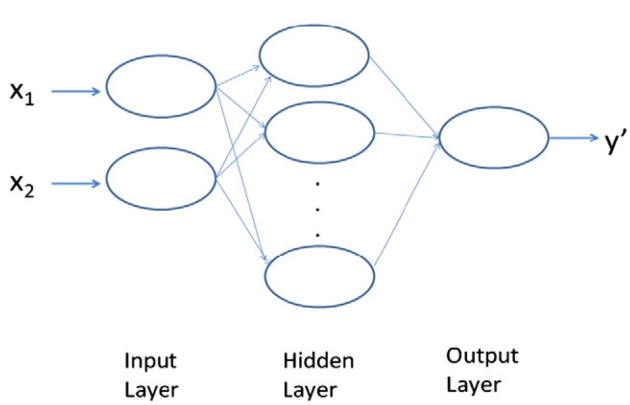

Fig. 1. Depiction of artificial neural network.

(e.g. personality differences). They can capture complex non-linearities and interactions. There is only one layer in Fig. 1 because no hidden nodes are connected to other hidden nodes. The network depicted in Fig. 1 is non-recursive because there are no loops among the hidden nodes that would result in the output of a node also being an input to the same node. Non-recursive networks are also called "Feed-Forward" ANNs. From a mathematical standpoint, Feed-Forward ANNs can be written as $Y = f(X) + \varepsilon$, where ε is the error term, Y is the observed dependent variable and $f(X)$ is some function of the independent variables, X.

While there are applications where recurrent ANNs and unsupervised learning can be used in marketing research, this article focuses on feed-forward ANNs because of their close relationship to regression and applicability to behavioral research. The remainder of this article is organized as follows. The next section describes the relationship between feed-forward ANNs and statistics and includes a discussion of some of the limitations of ANNs from a statistical perspective. The following section provides some examples of neural networks that address specific applications. The article closes with conclusions and discussion.

Relationship to statistics/econometrics

Relatively early in the development of ANNs, several researchers noticed the relationship of ANNs to a variety of statistical techniques (cf., Cheng & Titterington, 1994; White, 1989). Hornick, Stinchcombe, and White (1989) showed that feed-forward neural networks are "universal approximators."