The Consumer Prototype

Explaining the Underlying Psychological Factors of Consumer Behaviour with Artificial Neural Networks

Max N. Greene

Cardiff Business School, Cardiff University, Aberconway Building, Colum Drive, Cardiff, U.K.

1 STAGE OF THE RESEARCH

The interdisciplinary nature of the research project offers both distinctive advantages and challenges. Coming from the fields of strategic marketing and psychology, the field of Artificial Intelligence (AI) has not received a comprehensive consideration in the past. Therefore, the initial stages involved familiarization with the foundational work of Simon, Newell, Kurzweil, along with the philosophical discussions of Dennett, Minsky, Bechtel, and others – and is an ongoing process.

At the same time, negotiations regarding the secondary data from Office for National Statistics took place, completed successfully earlier this year. As a result, dataset was acquired that contains particularly useful for our purposes very extensive transactional consumer purchasing data (author is very grateful to ONS for the opportunity). Preliminary investigations are being carried out presently. Initial modelling stages are due to follow.

2 OUTLINE OF OBJECTIVES

The prospect of examining the underlying psychological (or psychosocial) factors that may explain behaviour – consumer behaviour in this instance – is intriguing, and could offer a substantial benefit for such fields like marketing, psychology, and others. Research of this type poses certain difficulties however, primarily concerned with the fact that consumers are human persons. The study of human behaviour is difficult not only due to the problems concerning the generalizability of the matter, but also due to the high cost of some methods that offer high level of confidence such as eye movement tracking and FMRI.

This research is primarily concerned with developing an artificial consumer prototype employing artificial neural networks (commonly referred to as neural networks, NNs) to subsequently examine varying underlying network structures and inherent interconnectivity in attempt to provide a descriptive and consequently a prescriptive account of human consumer decision-making ability.

3 RESEARCH PROBLEM

The research problem is twofold. On the one hand, the philosophical implications need be addressed that primarily revolve around the adequacy of employing an artificial agent to study the underlying phenomena of purchasing decisions versus studying actual human consumer behaviour. The transferability and extrapolation of insight acquired with the use of artificial agent towards the human consumer requires attention as well.

On the other hand, the actual development of the functional network based on the consumer purchasing data required to develop an artificial consumer prototype for the subsequent examinations is a formidable task in its own right. NNs would seem to have a special appeal for such a task, as the models learn the patterns in the data over numerous iterations and settle into a stable state as a result, once the predefined learned parameters have been attained and the network is no longer able to improve. At that time, after the network is optimized, variable contribution analysis will be carried out. For comparative purposes, a number of networks will be developed with varying degree of complexity, architecture and training algorithms.

The interpretation of the observed changes that occurs in different network architectures will take us back to the first part of the research problem, as it would involve a comprehensive philosophical discussion of the implications the results may entail.

4 STATE OF THE ART

In the high-level task of pattern recognition while

N. Greene M. (2013).
The Consumer Prototype - Explaining the Underlying Psychological Factors of Consumer Behaviour with Artificial Neural Networks.
In *Doctoral Consortium*, pages 38-43

examining complex behaviour phenomena, linear models could only be useful in explaining linear relations. For the purposes of the present discussion however this would be insufficient, as consumer behaviour and the process of decision-making in a modern market and socio-economic environment is without a doubt a very intricate and multifarious phenomenon composed of a large number of interrelated developments, where simple changes in one part of the system are able to produce complex effects throughout. It has been indeed a common practice to attempt to decompose the larger phenomena and isolate the process into individual elements for the subsequent analysis controlling for all other variables. The learning thus obtained could then be propagated to the higher level of the process. This method however is very inefficient and poses a serious scalability problem – that is of course in addition to the limitation concerning the ability of researcher to identify the individual parts of the process correctly (the task some believe to be impossible). A better method would be to examine the relations between all components concurrently.

NNs are able to examine all variables and account for nonlinear relations within the data once the hidden layers are introduced into the model structure. This offers high predictive capacity, but not only that. The weights could be examined for explanatory purposes and are able to provide an insight into the intrinsic nature of the process. Consumer decision making is an intricate continuous behaviour exhibited by persons that NNs seem to be particularly suited for as a method of analysis for a number of reasons. First, NNs model architecture resembles physiological inner workings and structure of a human brain – making it a particularly good fit to study human processes. Second, connectionism (the theoretical framework of NNs) is a set of approaches in the fields of AI and cognitive psychology that is particularly suited for modelling behaviour as the emergent processes of interconnected networks of simple units from the conceptual point. The hidden layers and nodes that are developed in the process of training a NNs model (NNs models repeatedly intake data and adjust the weights in the process up to a point of equilibrium where the model cannot improve anymore – method commonly referred to as training as it indeed resembles the process of training in the traditional sense) are not like input and output variables that come from the data, but could rather represent underlying abstract concepts and latent variables identified in the process of training that play a major role in explaining the relation between the input and the output layers.

This research project will address the idea of interpreting the number of hidden layers, nodes and weight values in NNs models in attempt to provide an explanatory account of consumer behaviour.

4.1 Artificial Neural Networks

There are a number of qualitative differences that set NNs apart for other AI approaches, namely learning and representation. Other distinguishing features worth mentioning are inherent parallelism, nonlinearity, and ability to exhibit exceptional performance with noisy data (Gallant, 1993).

Machine learning broadly refers to ability of a model to improve its performance based upon input information. It is generally considered that research on machine learning presents the highest potential to eventually develop models able to perform complicated AI tasks, as algorithms that learn from training and experience are superior to those based on a subset of contingency rules developed by human scientists. Machine learning may be divided into supervised learning and unsupervised learning.

4.1.1 Supervised Learning

Supervised learning is a group of learning algorithms that analyze training data (i.e. labelled data: pairs of input and output values) to produce an inferred or a regression function able to predict the correct output for any input. It is required for the learning algorithm to make certain generalizations from the training data that could be used to analyze previously unseen data – a process that is analogous to concept learning in human and animal psychology. Feed forward networks are the most common representative, and will be used for the purposes of present research.

4.1.2 Unsupervised Learning

Unsupervised learning refers to the machine learning problem aimed to determine the underlying structure of unlabelled data. In unlabelled data, there is no error signal to evaluate possible solution, and therefore the algorithm relies on techniques such as clustering that examine the core features of the data – self-organizing map is one such algorithm often used in NNs models, and will be used for the purposes of this research project.

4.2 Interpreting NNs Models Output Parameters

A number of ways may provide an insight into what happens inside the NNs model and help interpret the result. Some of the most common methods assess how the number of hidden layers and nodes affects the predictive and explanatory capacity of the model. A number of algorithms have been devised to make use of the weight values from NNs model output. Model architecture pruning techniques have also been shown to have a positive outcome in developing models with improved out of sample testing faculties. In the following sections, these methods are briefly discussed.

4.2.1 Number of Hidden Layers and Nodes

Model size matters. It has been shown that large models used to analyze extensive datasets show better predictive capacity.

Once the models are developed it is imperative to have a look into the optimal model structure however. It is indeed true that the larger models would offer higher predictive capacity and increase in the model fit, but at the same time, larger models need be penalized according to the Occam's razor principle. For example, one method to evaluate the model performance and select the optimal structure is described by Huang et al., (2004). Their method eliminates the independent variables that do not carry sufficient predictive and explanatory capacity and therefore do not need to be considered in the model. Thus, the model structure is simplified resulting in higher AIC and BIC values as both methods penalize model size while maximizing the model performance at the same time.

4.2.2 Model Pruning

Model architecture plays an important role in model adaptive performance.

While exploring environmental conditions that may have an effect on fish population, Olden and Jackson (2001) compared traditional statistical approaches with NNs models. In the NNs model structure, the connection weights between neurons are the associative links that signify the relation between the input and output variables and therefore are the key to solving the problem. Connection weights signify the influence each input variable is able to exert on the output, and dictate the direction of the influence. Input variables with large connective weights carry higher signal transfer capacity and therefore exert higher influence on the output variable. Excitatory effect (incoming signal increased with positive output effect) is represented by the positive connection weight and inhibitory effect (incoming signal reduced with negative output) is represented by the negative connection weight.

Even if it is possible to assess the overall contribution of input variables employing these approaches, the interpretation of interactive relations within the data presents an increasingly difficult undertaking, as the interactions between the variables in the network require immediate examination. Even a small network would contain a increasingly large number of connections, making the interpretation increasingly difficult. One way to manage this is through pruning connections with small weights that do not exert significant influence over the network structure and output (Bishop, 1995). Deciding which weights to remove or keep however is a task that requires substantial effort. Following the NNs approach, Olden and Jackson (2001) were able to develop and describe a randomization test to address this task. As a result, Olden and Jackson (2001) were able to provide a predictive and explanatory insight into nonlinear complex relations of ecological data (a task that poses a serious problem for traditional statistical approaches as species often exhibit nonlinear response to environmental conditions). In the course of detailed evaluation of NNs and traditional models it was shown that partitioning the predictive performance of the model into measures such as sensitivity (ability to predict the presence) and specificity (ability to predict the absence) allows for a more efficient way to assess the model strengths, weaknesses, and applicability. It is also shown that NNs are a useful approach for examining the interactive effects and factors. Both empirical and simulated datasets were used for comparative purposes, and show superior predictive performance of NNs models over traditional regression approaches (Olden and Jackson, 2001).

Building upon their work, approach that Olden and Jackson (2002) propose in their following publication provides the facility to eliminate irrelevant connections between neurons whose weights do not significantly influence the network output (i.e. predicted response variable), thus facilitating the interpretation of individual and interacting contributions of the input variables in the network. The approach is able to identify variables that provide a significant contribution to network predictive capacity, which effectively constitutes a

NNs variable selection method.

4.2.3 Interpreting Model Weights

Relatively few studies are carried out with the aim of developing methods for variable contribution analysis in NNs models – perhaps at least in part due to seeming complexity of the task.

Variable contribution analysis methods have been examined and compared by Gevrey, Dimopoulos and Lek (2003). One of the seven methods they surveyed included a computation that used connection weights to provide explanatory dimension to a NNs model using ecological data. First proposed by Garson (1991) and later further investigated by Goh (1995), the procedure is set to determine the relative importance of the inputs by partitioning the connection weights. Essentially, hidden-output connection weight of hidden neurons is partitioned into components associated with the input neurons (please see more in Appendix A of the (Gevrey et al., 2003)). Authors concluded that method that uses connection weights was able to provide a good classification of input parameters even though it was found to lack stability.

One of the concerns conveyed regarding the otherwise extensive investigation of different methods was that the dataset originally employed in 2003 study (Gevrey et al., 2003) was empirical, and therefore did not allow to ascertain the factual precision and accuracy of each method as the true relations between the variables are not known (Olden et al., 2004). Instead, the artificial dataset was created using the Monte Carlo simulation and employed to assess true accuracy of each method using the dataset with defined and therefore knows relations. Results showed that weights method that uses input-hidden and hidden-output connection weights showed consistently best results out of all methods assessed, contrary to Gevrey et al., (2003) findings. Additionally, the weights method was able to accurately identify the predictive importance ranking, whereas other methods were only able to identify the first few if any at all (Olden et al., 2004).

Olden and Jackson (2002) also used ecological data to demonstrate the predictive and explanatory power of NNs. A number of methods surveyed, including Neural Interpretation Diagram, Garson's algorithm and sensitivity analysis, aid in understanding the mechanics of NNs and improve the explanatory power of the models. Interpretation of statistical models is imperative for acquiring knowledge about the causal relationships behind the phenomena studied. They also propose a randomization approach for statistical evaluation of the importance of connection weights and the contribution of input variables in the neural network (already discussed in details in the sections above).

Nord and Jacobsson (1998) have also addressed the issue of explaining and interpreting NNs structure and developed algorithms for variable contribution analysis. The study compared the proposed novel algorithmic approach for NNs model interpretation with the analogous variable contribution method of partial least squares regression. Sensitivity analysis is also performed through setting each input to zero in a sequential manner. Linear regression coefficients for each of the input variables have also been generated for the purposes of examining the variable contribution direction. The results of the two approaches are then reviewed and compared to the results of the partial least squares regression. What the study is able to reveal is that in the linear dataset both the partial least squares regression and NNs models show similar performance in the variable contribution task, whereas with the nonlinear data the differences become obvious (Nord and Jacobsson, 1998).

Andersson et al., (2000) present two methods to study variable contribution in NNs models: (1) a variable sensitivity analysis and (2) method of systematic variation of variables. Variable sensitivity analysis is based on setting the connection weights between the input and hidden layer to a zero in a sequential manner, whereas the systematic variation of variables method is based on keeping the other variables constant or manipulated simultaneously. In the course of their study, it is shown that there is a high similarity between the method proposed by the authors for the variable contribution analysis in NNs models and the nature of the processes used to develop the synthetic datasets used. Thus, it is shown that the NNs models are suitable not only for the function approximation in nonlinear datasets, but are also able to accurately reflect the characteristic qualities of the input data. The transparency of highly interconnected NNs models could be demonstrated in response to the 'black box' argument. Presented method is then able to generate information about the variables that could be useful in examination and interpretation of variable contribution and relations.

The discussed earlier method of Nord and Jacobsson (1998) is based on the saliency estimation principles (such as Optimal Brain Surgeon, Optimal Brain Damage, etc.) as it estimates the consequence of weight deletion on prediction error. The difference with the method proposed by Andersson, Aberg and Jacobsson (2000) is in the way estimation is carried out (theoretical calculation in saliency estimation methods as opposed to experimentally derived values presented (Andersson et al., 2000), and builds upon the findings of Nord and Jacobsson (1998). In the course of analysis, a systematic variable contribution analysis is carried out on a highly interconnected network structure, including the signal separation exercise, employing a number of synthetic and empirical dataset to provide additional information on the methods considered, including the ability to show graphically the variable interdependencies. Other research is based on the principle of systematic variable variation and not the connection weights. Information obtained in such a way could constitute an analytical basis for a comprehensive variable contribution analysis and variable selection procedure survey (Nord and Jacobsson, 1998).

5 METHODOLOGY

Research project would include two phases. First, smaller data subset is used to develop and optimize the procedure, carry out all the preliminary analyses and produce the programming code:

- Regression models developed for exploratory and descriptive purposes to examine the data and carry out simple linear modelling;
- NNs as a primary method of analysis to develop complex nonlinear models;
- NNs learning algorithms are examined and selected for consecutive modelling;
- Various network architectures are studied and optimized employing pruning methods.

In the second stage, the procedure established in the first stage will be followed using a full data, effectively scaling up the analyses to the full power (if necessary, intermediate transitional stages may be incorporated to gradually scale up the procedure):

- Full scale network architectures optimized;
- Variable contribution analysis carried out;
- Network structure examined and interpreted in the context of consumer behaviour.

6 EXPECTED OUTCOME

The expected outcome of this research should produce an artificial consumer prototype based on the actual human consumer purchasing data, with attempt to identify complex latent underlying factors that may influence the artificial behaviour. These findings will then be extrapolated to human consumers, aimed to provide the insight into the underlying psychological factors of human consumer behaviour.

The philosophical contemplations presented here should promote the advancement of interdisciplinary research, facilitating cooperation between fields such as psychology, strategic marketing and artificial intelligence, and provide significant benefit in acceptance and advance of computational methods to study consumer behaviour. This should serve as a catalyst for a broader dialogue between the marketing professionals in the industry that express demand in highly accurate forecasting and business intelligence tools, and the researchers in the field of consumer behaviour.

DISCLAMER

Data supplied by TNS UK Limited. The use of TNS UK Ltd data in this work does not imply the endorsement of TNS UK Ltd. in relation to the interpretation or analysis of the data. All errors and omissions remain the responsibility of the authors.

REFERENCES

Andersson, F., Aberg, M., & Jacobsson, S., 2000. Algorithmic approaches for studies of variable influence, contribution and selection in neural networks. In *Chemometrics and intelligent laboratory systems*.

Bishop, C., 1995. Neural networks for pattern recognition: *Oxford university press*.

Gallant, S. I., 1993. Neural network learning and expert systems: *The MIT Press*.

Garson, D., 1991. Interpreting neural-network connection weights. In *AI expert*.

Gevrey, M., Dimopoulos, I., & Lek, S., 2003. Review and comparison of methods to study the contribution of variables in artificial neural network models. In *Ecological Modelling*.

Goh, A., 1995. Back-propagation neural networks for modeling complex systems. In *Artificial Intelligence in Engineering*.

Huang, Z., Chen, H., Hsu, C.-J., Chen, W.-H., & Wu, S., 2004. Credit rating analysis with support vector machines and neural networks: a market comparative study. In *Decision support systems*.

Nord, L. I., & Jacobsson, S. P., 1998. A novel method for examination of the variable contribution to computational neural network models. In *Chemometrics and intelligent laboratory systems*.

Olden, J. D., & Jackson, D. A., 2001. Fish–habitat relationships in lakes: gaining predictive and explanatory insight by using artificial neural networks. In *Transactions of the American Fisheries Society*.

Olden, J. D., & Jackson, D. A., 2002. A comparison of statistical approaches for modelling fish species distributions. In *Freshwater Biology*.

Olden, J. D., Joy, M. K., & Death, R. G., 2004. An accurate comparison of methods for quantifying variable importance in artificial neural networks using simulated data. In *Ecological Modelling*.