Chapter 12
Artificial Neural Networks and Discrete Choice Models:
Sales Forecast in Supermarket Products

Paulo Botelho Pires
https://orcid.org/0000-0003-3786-6783
CEOS, Polytechnic of Porto, Portugal

José Duarte Santos
https://orcid.org/0000-0001-5815-4983
CEOS, ISCAP, Polytechnic of Porto, Portugal

ABSTRACT

The performance of artificial neural networks was compared with the performance of discrete choice models in predicting the purchase of products with weak involvement. A comprehensive literature review on the main paradigms of artificial neural networks was carried out, namely variants of the back-propagation algorithm, radial basis function, and genetic computing. Within the class of discrete choice models, the authors restricted the comparison to the multinomial logit model and the mixed logit. The performance of the models was measured in a database of grocery purchases in supermarkets. Artificial neural networks outperformed discrete choice models in predicting sales in supermarkets, and both types of models demonstrated strong predictive power. As a result, both can be reliably used in marketing to estimate individual or collective probabilities of supermarket product purchases.

INTRODUCTION

Today the availability and access to information about consumer behavior is an irreversible reality. In this context, a set of models is investigated that allow the information from consumer behavior to be analyzed and, subsequently, to gauge future behavior. The research has been restricted to two disciplinary areas: discrete choice models and artificial neural networks. This choice was mainly due to the capac-

DOI: 10.4018/978-1-6684-8574-3.ch012

ity evidenced by the models and the consequent repercussion. The secondary motivation, but no less important, was based on the evaluation and comparison of the predictive ability of this set of models. Coming from different scientific areas and based on their own concepts and deductions, these models have been applied to marketing problems with undeniable success.

Given the proliferation of paradigms, the scope of the study had to be limited to a smaller number. Thus, in discrete choice models, the study focused on the Multinomial Logit model, considered by many to be the base and simplest model, and on the Mixed Logit model, the representative of a class of flexible and more powerful models. In artificial neural networks, the study was restricted to unidirectional architectures, summarizing the variants of the backpropagation algorithm, the Radial Basis Function, and the integration of genetic algorithms and artificial neural networks.

BACKGROUND

The literature review is logically divided into two parts and five sections. The first part contains a review of the main paradigms of artificial neural networks. A review of discrete choice models is included in the second part, while the literature review about the comparison of artificial neural networks and discrete choice models is placed in the third part.

Artificial Neural Networks

The Backpropagation Algorithm

The Backpropagation (BP) algorithm is the best-known and most widely used of all learning algorithms for Artificial Neural Networks (ANN). First formulated for learning generic networks by Paul Werbos in 1974, it was later adapted by three groups of researchers as a learning algorithm for multilayer ANNs. The BP algorithm is a supervised learning algorithm and can be applied following an incremental, batch, or intermediate learning method (Heskes & Wiegerinck, 1996; Möller, 1990; Torresen, 1997). The BP algorithm is the application of the gradient or steepest descent method to an ANN with unidirectional architecture, with the mean squared error generally chosen as the error function.

The BP algorithm in its most elementary form, while having been considered an important advance in the evolution of ANNs, has a very slow learning phase, making its practical applicability very limited (Hagan et al., 1996). An ANN with a large sample size causes a very long learning phase, with thousands of iterations and epochs (Reed & Marks, 1999). In the BP algorithm, the error function generally used is the mean of the sum of the squared error (batch learning) or the squared error (incremental learning). These functions, when applied to ANN with multiple layers, generate very complex surfaces, characterized by the existence of several minima. In certain problems, this characteristic results in the convergence of the algorithm to a local minimum (Hagan et al., 1996).

Backpropagation Momentum Algorithm

The BPMO (BP Momentum) algorithm requires an additional coefficient γ, called momentum, and the changes to the values of the connections, $\{k+1\}$, depend on several previous iterations or epochs (depending on the learning method) (Fausett, 1994). In its simplest form (first order), the BPMO algorithm

Related Content

Sustainable Approach towards Thai SMEs: The Effects of Country Philosophy of Sufficiency Economy and Government Support
Siriluck Thongpoon, Noor Hazlina Ahmadand Imran Mahmud (2018). *Operations and Service Management: Concepts, Methodologies, Tools, and Applications (pp. 1057-1078).*
www.irma-international.org/chapter/sustainable-approach-towards-thai-smes/192520

Reducing the Negative Effects of Psychological Contract Breach during Management-Imposed Change: A Trickle-Down Model of Management Practices
Melanie De Ruiter, Robert J. Blommeand René Schalk (2016). *Organizational Change Management Strategies in Modern Business (pp. 122-142).*
www.irma-international.org/chapter/reducing-the-negative-effects-of-psychological-contract-breach-during-management-imposed-change/140326

The Impact of Enhancing Social Media Marketing Knowledge on Customer Attraction and Engagement for University Organizational Growth and Development from the TRACK Theory: The Context of Mpumalanga University of South Africa
Glenton Khulani Samboand Austin Musundire (2020). *International Journal of Applied Management Theory and Research (pp. 19-40).*
www.irma-international.org/article/the-impact-of-enhancing-social-media-marketing-knowledge-on-customer-attraction-and-engagement--for-university-organizational-growth-and-development-from-the-track-theory/260736

Recruiting for Resilience: C-Suite Leaders in the Life Sciences Share Lessons Learned
Helen Mary Meldrum (2022). *International Journal of Applied Management Theory and Research (pp. 1-18).*
www.irma-international.org/article/recruiting-for-resilience/288506

The Tacit Knowledge-Centric Innovation: Toward the Key Role of Customer Experiential Knowledge
Dhouha Jaziri Bouaginaand Abdelfattah Triki (2015). *Handbook of Research on Effective Project Management through the Integration of Knowledge and Innovation (pp. 75-100).*
www.irma-international.org/chapter/the-tacit-knowledge-centric-innovation/124713