Neural Networks for Retail Sales Forecasting

G. Peter Zhang
Georgia State University, USA

INTRODUCTION

Forecasting of the future demand is central to the planning and operation of retail business at both macro and micro levels. At the organizational level, forecasts of sales are needed as the essential inputs to many decision activities in various functional areas such as marketing, sales, production/purchasing, as well as finance and accounting (Mentzer & Bienstock, 1998). Sales forecasts also provide basis for regional and national distribution and replenishment plans. The importance of accurate sales forecasts to efficient inventory management has long been recognized. In addition, accurate forecasts of retail sales can help improve retail supply chain operation, especially for larger retailers who have a significant market share. For profitable retail operations, accurate demand forecasting is crucial in organizing and planning purchasing, production, transportation, labor force, as well as after sales services.

Barksdale and Hilliard (1975) examined the relationship between retail stocks and sales at the aggregate level and found that successful inventory management depends to a large extent on the accurate forecasting of retail sales. Thall (1992) and Agrawal and Schorling (1996) also pointed out that accurate demand forecasting plays a critical role in profitable retail operations and poor forecasts would result in too many or too few stocks that directly affect revenue and competitive position of the retail business. The importance of accurate demand forecasts in successful supply chain operations and coordination has been recognized by many researchers (Chopra & Meindl, 2001; Lee et al., 1997).

Retail sales often exhibit both seasonal variations and trends. Historically, modeling and forecasting seasonal data is one of the major research efforts and many theoretical and heuristic methods have been developed in the last several decades. Different approaches have been proposed but none of them has reached consensus among researchers and practitioners. Until now, the debate has still not abated in terms of what is the best approach to handle the seasonality.

On the other hand, it is often not clear how to best model the trend pattern in a time series. In the popular Box-Jenkins approach to time series modeling, differencing is used to achieve stationarity in the mean. However, Pierce (1977) and Nelson and Plosser (1982) argue that differencing is not always an appropriate way to handle trend, and linear detrending may be more appropriate. Depending on the nature of the nonstationarity, a time series may be modeled in different ways. For example, a linear or polynomial time trend model can be used if the time series has a deterministic trend. However, if a time series exhibits a stochastic trend, the random walk model and its variations may be more appropriate.

In addition to controversial issues around the ways to model seasonal and trend time series, one of the major limitations of many traditional models is that they are essentially linear methods. In order to use them, users must specify the model form without the necessary genuine knowledge about the complex relationship in the data. This may be the reason for the mixed findings reported in the literature regarding the best way to model and forecast trend and seasonal time series.

One nonlinear model that has recently received extensive attention is the neural network model (NN). The popularity of the neural network model can be attributed to their unique capability to simulate a wide variety of underlying nonlinear behaviors. Indeed, research has provided theoretical underpinning of neural network's universal approximation ability. In addition, few assumptions about the model form are needed in applying the NN technique. Rather, the model is adaptively formed with the real data. This flexible data-driven modeling property has made NNs an attractive tool for many forecasting tasks, as data are often abundant, while the underlying data generating process is hardly known.

In this article, we provide an overview on how to effectively model and forecast consumer retail sales using neural network models. Although there are many studies on general neural network forecasting, few are specifically focused on trending or seasonal time series. In addition, controversial results have been reported in the literature. Therefore it is necessary to have a good summary of what has been done in this area and more importantly to give guidelines that can be useful for forecasting practitioners.

It is important to note that the focus of this article is on time series forecasting methods. For other types of forecasting methods used in retail sales, readers are referred to Green (1986), Smith et al. (1994), and Dominique (1998).

BACKGROUND

Neural networks are computing models for information processing. They are very useful for identifying the functional relationship or pattern in the retails sales and other time series data. The most popularly used neural network model in practice for retail sales is the feedforward multi-layer network. It is composed of several layers of basic processing units called neurons or nodes. For an in-depth coverage of NN models, readers are referred to Smith (1993) and Bishop (1995). A comprehensive review of the NNs for forecasting is given by Zhang et al. (1998).

Before it can be used for forecasting, the NN model must be built first. Neural network model building (training) involves determining the order of the network (the architecture) as well as the parameters (weights) of the model. NN training typically requires that the in-sample data be split into a training set and a validation set. The training set is used to estimate the parameters of some candidate models, among which the one that performs the best on the validation set is selected. The out-of-sample observations can be used to further test the performance of the selected model to simulate the real forecasting situations.

The standard three-layer feedforward NN can be used for time series forecasting in general and retail sales in particular. For one-step-ahead forecasting, only one output node is needed. For multiple-step forecasting, more output nodes should be employed. For time series forecasting, the most important factor in neural networks modeling is the number of input nodes, which corresponds to the number of past observations significantly auto-correlated with the future forecasts. In a seasonal time series such as the retail sales series, it is reasonable to expect that a forecasting model should capture the seasonal autocorrelation that spans at least one or two seasonal periods of, say, 12 or 24 for monthly series.

Therefore in modeling seasonal behavior, it is critical to include in the input nodes the observations separated by multiples of seasonal period. For example, for a quarterly seasonal time series, observations that are four quarters away are usually highly correlated. Although theoretically, the number of seasonal lagged observations that have autocorrelation with the future value can be high, it is fairly small in most practical situations, as empirical studies often suggest that the seasonal autoregressive order be one or at most two (Box & Jenkins, 1976).

There are many other parameters and issues that need to be carefully considered and determined in neural network model building for retail and other time series. These include data preparation, data division and sample size, network architecture in terms of number of hidden and input nodes, training algorithm, model evaluation criteria and so forth. Practical guidelines can be found in many references in the literature including Adya and Collopy (1998), Kaastra and Boyd (1996), and Zhang et al. (1998).

ISSUES AND LITERATURE REVIEW

Modeling seasonal and trend time series has been one of the main research endeavors for decades. In the early 1920s, the decomposition model along with seasonal adjustment was the major research focus due to Person's (1919) work on decomposing a seasonal time series. Different seasonal adjustment methods have been proposed and the most significant and popular one is the Census X-11 method developed by the Bureau of the Census in 1950s and 60s, which has evolved into the current X-12-ARIMA program. Because of the ad hoc nature of the seasonal adjustment methods, several model-based procedures have been developed. Among them, the work by Box and Jenkins (1976) on the seasonal ARIMA model has had a major impact on the practical applications to seasonal time series modeling. This model has performed well in many real-world applications and is still one of the most widely used seasonal forecasting methods. More recently, neural networks have been widely used as a powerful alternative to traditional time series modeling.

In neural network forecasting, little research has been done focusing on seasonal and trend time series modeling and forecasting. In fact, how to effectively model seasonal time series is a challenging task not only for the newly developed neural networks, but also for the traditional models. One popular traditional approach to dealing with seasonal data is to remove the seasonal component first before other components are estimated. Many practitioners in various forecasting applications have satisfactorily adopted this practice of seasonal adjustment. However, several recent studies have raised doubt about its appropriateness in handling seasonality. Seasonal adjustment has been found to lead into undesirable nonlinear properties, severely distorted data, and inferior forecast performance (Ghysels et al., 1996; Plosser, 1979). De Gooijer and Franses (1997) pointed out that "although seasonally adjusted data may sometimes be useful, it is typically recommended to use seasonally unadjusted data". On the other hand, mixed findings have also been reported in the limited neural network literature on seasonal forecasting. For example, Sharda and Patil (1992) found that, after examining 88 seasonal time series, NNs were able to model seasonality directly and pre-deseasonalization is not necessary. Alon et al. (2001) also found that NNs are able to "capture the dynamic nonlinear trend and seasonal patterns, as well as the interactions

3 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:
www.igi-global.com/chapter/neural-networks-retail-sales-forecasting/14567

Related Content

Developing Inter-Organizational-Trust in Business to Business E-Commerce Participation - Case Studies in the Automotive Industry
Pauline Ratnasingam (2002). *Annals of Cases on Information Technology: Volume 4 (pp. 184-194).*
www.irma-international.org/chapter/developing-inter-organizational-trust-business/44506

A Combined Dimensional Kernel Method for Graph Classification
Tiejun Cao (2017). *Journal of Information Technology Research (pp. 22-33).*
www.irma-international.org/article/a-combined-dimensional-kernel-method-for-graph-classification/182710

Agile Methodology in an SaaS Deployment Project: A Pre- and Post-SaaS Deployment Study
Ali Hassan, Soayba Younasand Amiya Bhaumik (2021). *International Journal of Information Technology Project Management (pp. 1-29).*
www.irma-international.org/article/agile-methodology-in-an-saas-deployment-project/273205

The Lonely Comate - The Adoption-Failure of an Intranet-based Consumer and Market Intelligence System
Paul H. J. Hendriksand Wendy H. Jacobs (2003). *Annals of Cases on Information Technology: Volume 5 (pp. 130-150).*
www.irma-international.org/article/lonely-comate-adoption-failure-intranet/44538

The Impact of Information and Communication Technologies on Economic Growth and Electricity Consumption: Evidence from Selected Balkan and Eastern European Countries
Burcu Berke, Gülsüm Akarsuand Gökhan Obay (2017). *Information and Communication Overload in the Digital Age (pp. 176-200).*
www.irma-international.org/chapter/the-impact-of-information-and-communication-technologies-on-economic-growth-and-electricity-consumption/176570