WHITEPAPER

Neural Networks-

Comparative Study for Forecasting

VALIANCE

Neural Networks-Comparative Study for Forecasting

Abstract

ANNs (Artificial Neural Networks) have powerful pattern classification and pattern recognition capabilities. Inspired by biological systems, particularly by research into the human brain, ANNs are able to learn from underlying relationships unknown or hard to learn from and generalize from experience. Currently ANNs are being used for a wide variety of tasks in many different fields of business, industry and science etc.

Artificial neural networks, originally developed to mimic basic biological neural systems– the human brain particularly, are composed ofa number of interconnected simpleprocessing elements called neurons or nodes. Each node receives an input signal which is the total "information" from other nodes or external stimuli, processes it locally throughan activation or transfer function and produces a transformed outputsignal to othernodes or external outputs. Although each individual neuron implements its function rather slowly and imperfectly collectively, a network can perform a surprising number of tasks quite efficiently. This information processing characteristic makes ANNs a powerful computational technique with ability to learn from examples and then to generalize to examples never before seen

The traditional approaches to time-series prediction, such as Box-Jenkins or ARIMA method assumes that time series understudy is generated from linear process. Linear models have advantages in that they can be understood and analyzed in great detail, and they are easy to explainand implement. However they may be totally inappropriate if the underlying mechanism is non-linear. As a matter of fact, ANNs are capable of performing nonlinear modeling without a priori knowledge about the relationships between input and output variables. Thus they area more general and flexible modeling tool for forecasting.

Theoretical Approach to Neural Networks

The linear models for regression and classification respectively, are based on linear combinations of fixed nonlinear basis functions φj(x) and take the form

$$y(\mathbf{x}, \mathbf{w}) = f\left(\sum_{j=1}^{M} w_j \phi_j(\mathbf{x})\right)$$

Where f (•) is a nonlinear activation function in the case of classification and is the identity in the case of regression. Our goal is to extend this model by making the basic functions φj(x) depend on parameters and then to allow these parameters to be adjusted, along with the coefficients {wj}, during training. There are, of course, many ways to construct parametric nonlinear basis functions. Neural networks use basic functions that follow the same form so that each basis function is itself a nonlinear function of a linear combination of the inputs, where the coefficients in the linear combination are adaptive parameters.

This leads to the basic neural network model, which can be described, a series of functional transformations. First we construct M linear combinations of the input variables x1, . . . , xD in the form

$$a_j = \sum_{i=1}^{D} w_{ji}^{(1)} x_i + w_{j0}^{(1)}$$

Where j = 1, . . . , M, and the superscript (1) indicates that the corresponding parameters are in the first 'layer' of the network. We shall refer to the parameters w (1) ji as weights and the parameters w(1)j0 as biases, following the nomenclature The quantities aj are known as activations. Each of them is then transformed using a differentiable, nonlinear activation function h(·) to give

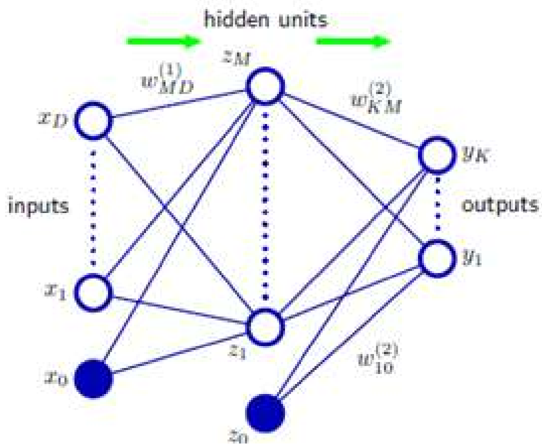
zj = h(aj)

These quantities correspond to the outputs of the basic functions in that, in the context of neural networks, are called hidden units. The nonlinear functions h (·) are generally chosen to be sigmoidal functions such as the logistic sigmoid or the 'tanh'

Following, these values are again linearly combined to give output unit activations

$$a_k = \sum_{j=1}^{M} w_{kj}^{(2)} z_j + w_{k0}^{(2)}$$

Where k = 1, . . . , K, and K is the total number of outputs. This transformation corresponds to the second layer of the network, and again the w(2) k0 are bias parameters. Finally, the output unit activations are transformed using an appropriate activation function to give a set of network outputs yk.



Where k = 1, . . . , K, and K is the total number of outputs. This transformation corresponds to the second layer of the network, and again the w(2) k0 are bias parameters. Finally, the output unit activations are transformed using an appropriate activation function to give a set of network outputs yk.

The Network diagram for the two layers neural network corresponding. The input, hidden, and output variables are represented by nodes and the weight parameters are represented by links between the nodes, in which the bias parameters are denoted by links coming from additional input and hidden variables x0 and z0. Arrows denote the direction of information flow through the network during forward propagation.

Business Context

We had to forecast monthly sales for a retailer on basis of several input parameters. Sales figures hardly showed any trends with monthly sales figure fluctuating as much as 40-50 percent month on month and year on year. Different forecasting techniques were used for our purpose.

Forecasting Techniques Used:

1. Moving average and ARIMA Modeling
2. Linear regression and segment based forecasting
3. Neural networks.

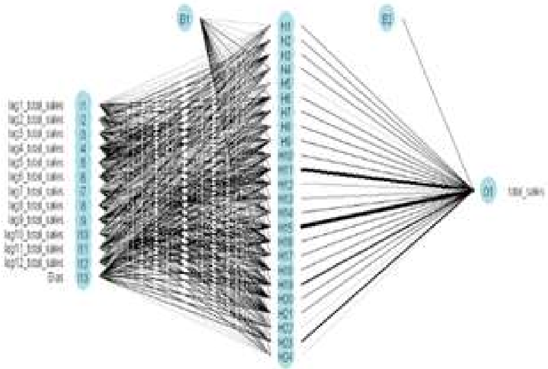
Moving Average and ARIMA Modelling

- The sales series showed an increasing trend but had large noise in it. There was no seasonality pattern in sales. Predicting such a series accurately was a challenge.
- As a first step, the series was smoothened using up to 6 moving average and exponential smoothing with alpha 0.1 to 0.9. The series was forecasted using both moving average and exponential smoothing method, error rate of the series was too high from both the techniques. i.e. 47%
- As a second step, we used ARIMA modeling method for forecasting. The result from ARIMA models were not satisfactory as the average error rate was very high i.e. 32% with a range of 0 to 75%.

Linear Regression and Segment Based Forecasting

- As a third step, we segmented agents into various segments using CART analysis basis various key performance indicators, demographics and other external variables.
- After identification of key segments, we forecasted the sales of each segment but the average error rate was still substantial. i.e. 27% with range of 2% to 65%

Neural Networks



Model Techniques	Average Error Rate
Moving Average and Exp Smoothening	47%
ARIMA	32%
Linear Regression	20%
Neural Networks	10%

Neural Network - Pros

- ANN has the ability to implicitly detect complex nonlinear relationships between dependent and independent variables.
- Prediction accuracy increases to a great extent.

Neural Network - Cons

- Complex to implement in real time.
- Neural network is a black box technique with little explanation of factors that affect forecasts or have a significant relationship with target variable. It's hard to explain why's to a business user.

CONCLUSION

Neural Networks can be highly accurate in making predictions, however, its value lies in making real time predictions in more of an operational role. Examples could be transactional fraud detection, spam filters, web search. Use of neural networks for forecasting strategic parameters like sales can be tricky as there is little to act upon if ANN were to show extreme results.

AUTHORS

- Mr. Trivendra Singh Verma (Business Analyst)
- Vikas Kamra (Co-Founder & CEO)
- Shailendra Kathait (Co-Founder & Chief Data Scientist)