Forecasting of Day Ahead Electricity Prices using Principal Component Analysis and a Neural Network

UCD Michael Smurfit Graduate Business School

MSc Business Analytics Dissertation

Eoin Kearney (eoin.kearney.1@ucdconnect.ie)

UCD Michael Smurfit Graduate Business School, Supervisor: Dr Paula Carroll

Introduction

In April 2011 both the commercial and domestic Irish electricity markets were deregulated in an effort to improve competition and reduce costs to end users. In a competitive electricity market it is important that wholesale electricity prices are forecast as accurately as possible to ensure the cheapest possible prices can be passed on to end users while maintaining margins for generators and suppliers.

Research Question

Can dimensionality reduction through Principal Component Analysis improve the performance of a Multilayer Perceptron Neural Network in the prediction of day ahead electricity prices?

Methodology

Principal Component Analysis (PCA)

PCA orthogonally transforms data into linearly uncorrelated variables (components). Components are then sorted so that each contains the highest possible variance under the constraint of remaining uncorrelated with the other components. This allows for dimensionality reduction by removing components that account for only small amounts of variance within the data. PCA was used at the pre-processing stage to produce two dimensionally reduced input datasets. The first reduced the number of input parameter from 8 to 6 (PCA6) whilst maintaining >99% of the total variance of the data and a second dataset with parameters reduced to 5 (PCA5) whilst maintaining >95% of the total variance of the data. A third input dataset (NN input) consisting of original 8 input parameters without PCA analysis was also examined for comparison.

Multilayer Perceptron Neural Network (MLPNN)

A MLPNN approach was taken to predict electricity prices for each of the three input datasets by using the MONMLP package in the R software environment. A MLPNN contains one or more hidden layers (only 1 hidden layer was examined in this research), one or more weights, an activation function between input and hidden layers plus a bias. Activation functions are required to deal with nonlinearity. A logistic function between input and hidden layer and a linear function between hidden and output were used. Each of the datasets were trained with 70% of the same randomly selected data with the remaining 30% of the data used for testing. Forty eight models in total were trained and tested with the number of hidden weights varying from 1 to 16 for each of the datasets.

Data

Historic electricity prices and demand were sourced from Single Electricity Market Operator (SEMO) and Eirgrid while commodity prices were obtained from ICE. The obtained data covered the time period covering January 2012 through to February 2014. The following 8 input parameters were chosen from the available data: change in gas prices, change in coal prices, change in carbon prices, prior day electricity price (SMP), total demand, renewable demand, gas demand and other demand.

The following Data Processing activities were performed:

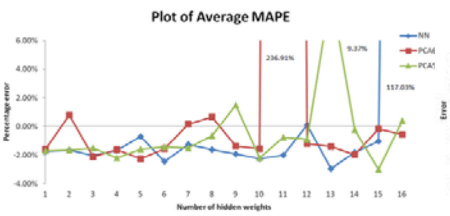
- Data cleaning replacing missing values etc
- Data aggregation and extraction
- Data normalisation by scaling in range 0-1
- Data pre-processing using Principal Component Analysis

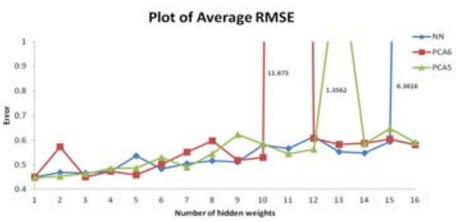
Results

Errors

The forty eight models (sixteen for each dataset) with 1 to 16 hidden weights were compared by errors using Mean Absolute Percentage Error (MAPE) and Root Mean Squared Error (RMSE) per formulae below:

$$\text{MAPE} = \frac{\frac{\sum |\text{Actual} - \text{Forecast}|}{Actual} X\ 100}{n} \qquad \text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(X_{obs} - X_{mo\,del})^2}{n}}$$





From the above graphs the PCA5 models have less variation in errors than the PCA6 and NN input models. In the case of all three model types there were extreme errors at a certain number of hidden weights; 10 for PCA6 with a MAPE of 236.9% and a RMSE of 11.67, 13 for PCA5 with a MAPE of 9.73% and a RMSE of 1.3562 and 16 for NN with a MAPE of 117% and a RMSE of 6.36. A full comparison of models was performed using Akaike's Information Criterion, for each of the three model types the 1 hidden weight model was found to have the highest likelihood of being the best model. Comparing the 1 hidden weight models, the PCA5 model with one hidden weight was deemed to have a 90.1% likelihood of being the best model for the data tested in this research, followed by 8.8% for PCA6 and 0.35% for NN.

Runtime

The average runtimes for training of the PCA5 and PCA6 models were reduced by 39.9% and 25.2% in comparison to the NN input models.

Conclusions

Overall from the models and data examined in this research, PCA appears to be an effective method for reducing the number of input parameters and in this case has resulted in a reduction in MAPE and RMSE errors. In terms of the number of hidden weights one appears to give the best models out of 1 to 16 hidden weights tested. However the models were examined using Akaike's Information Criterion which penalises models based on the number of parameters. PCA pre-processing with input reduction has also resulted in a significant reduction in training time.