

Business Intelligence Competition, BI CUP 2006

Entrant:



Philip Brierley
philb@philbrierley.com
www.tiberius.biz

Problem description:

A public transportation company is expecting increasing demand for its services and is planning to acquire new buses and to extend its terminals. These investments require a reliable forecast of future demand which should be based on historic demand stored in the company's data warehouse. For each 15-minute interval between 6:30 hours and 22 hours the number of passengers arriving at the terminal has been recorded and stored. As a forecasting consultant you have been asked to forecast the number of passengers arriving at the terminal.

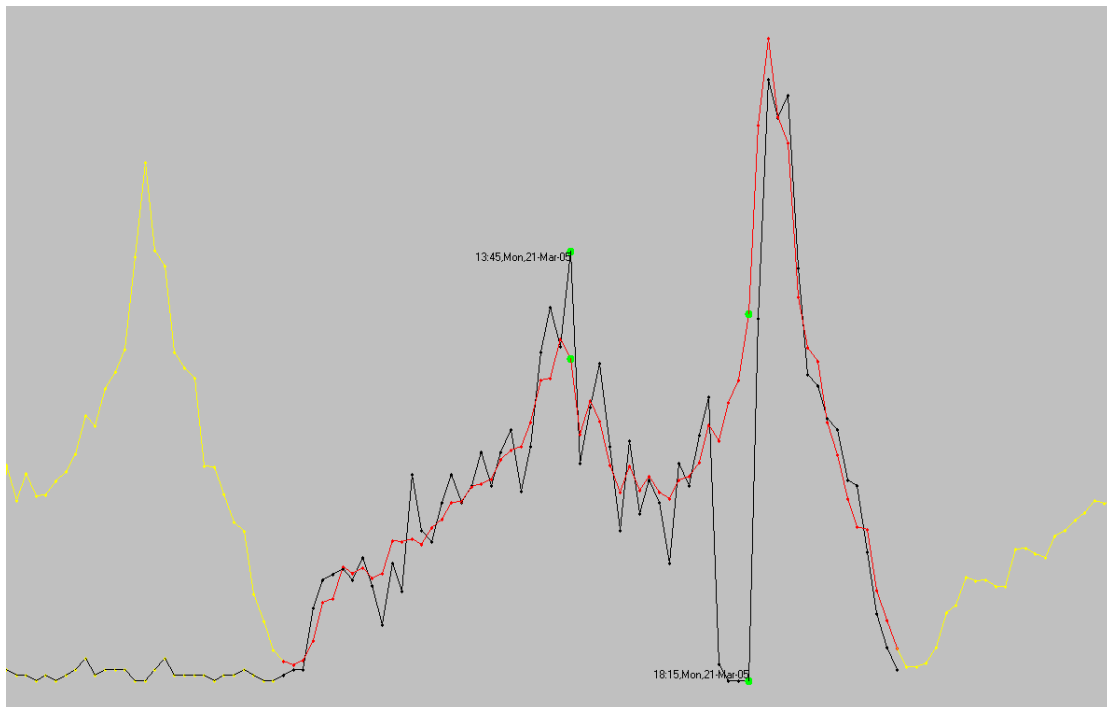
Methodology:

A multiplayer perceptron (Tiberius – see www.tiberius.biz) neural network was the modelling tool used.

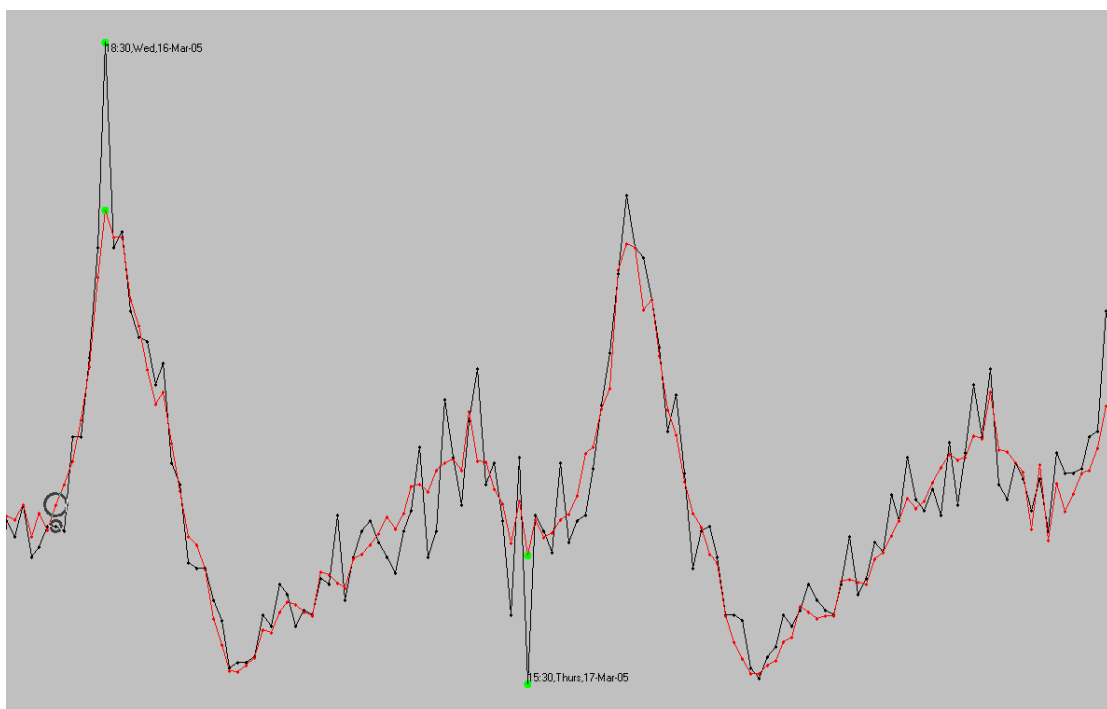
The input (time) was broken down using sine and cosine encoding to represent day of the week and hour of the day. The weekends were removed from the training data as the validation period did not include a weekend and they are obviously different.

Initial models were built and the errors examined. Three periods were discovered that contained suspect data that did not follow the norm. As information was not available as to what the reasons may be for these anomalies, the points were removed, rather than modelled in.

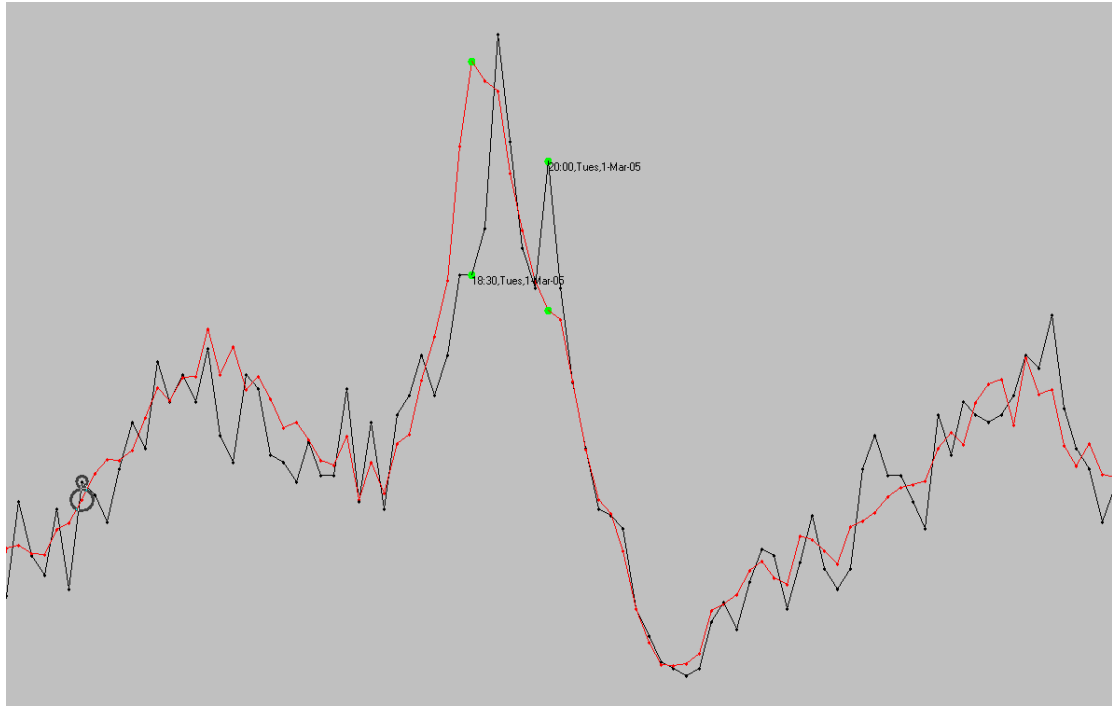
The neural network was trained to minimise the MAE, as this is the objective function on which the predictions will be measured. The MAE achieved on the training data (excluding the weekends and the anomalous points) was 4.12.

The Suspect Periods:

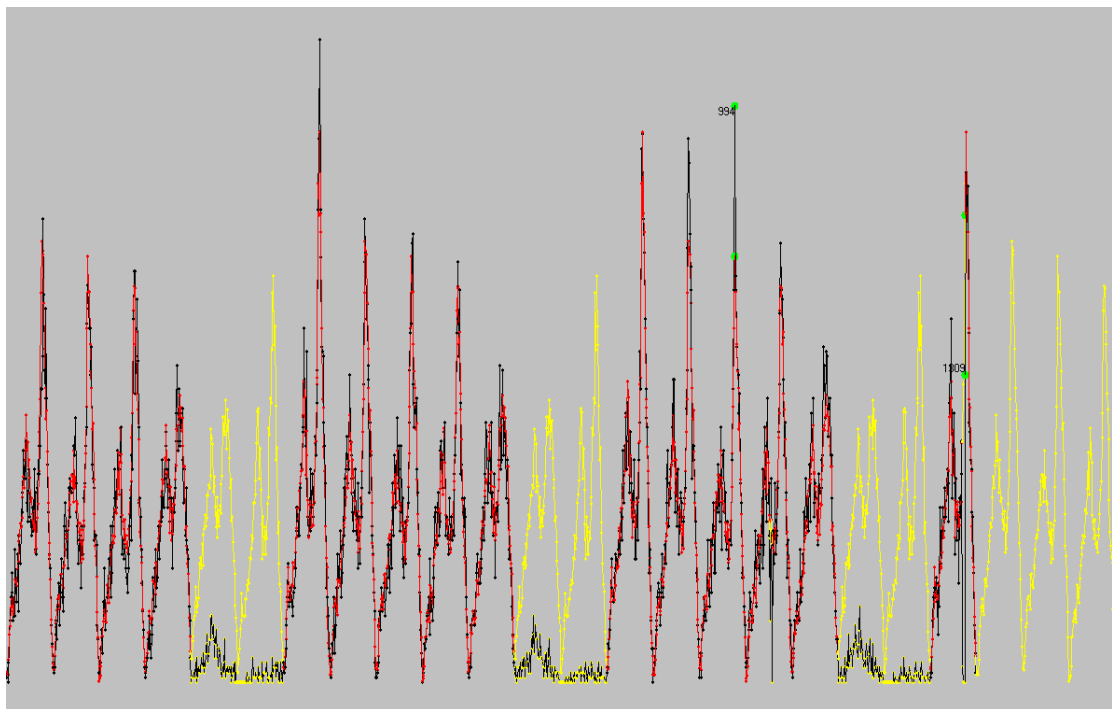
1) – there was no demand from 17:30 to 18:15 on Monday 21st March, and significantly reduced demand at 18:30. There will be a valid explanation for this, but the data cannot be used unless the reason is known and can be predicted in the future.



2) No demand at 17:30 on Thursday 17th March.



3) A 'delayed' demand on Tuesday 1st of March.



The final model, showing the weekends and the validation data in yellow.