Predicting Electricity Consumption using Deep Recurrent Neural Networks

Anupiya Nugaliyadde, Upeka Somaratne, and Kok Wai Wong Murdoch University, Perth, Australia, {a.nugaliyadde, upeka.somaratne, k.wong}@murdoch.edu.au

Summary

The ability to forecast the demand for electricity in the future will be very helpful for the electricity distributor. In this paper, two approaches to predicting electric consumption has been discussed. One approach is using a Recurrent Neural Network (RNN) and another approach is using a Long Short Term Memory (LSTM) network. Both these approaches are dependent purely on historical electricity consumption data. These models were tested on an individual house and also for a block of houses for long, short and medium sized time frames.

Predicting electricity consumption helps to decide the distribution of power from the national grid and therefore helps in preventing unplanned electricity distribution disruptions. An Advanced Metering Infrastructure was used to collect the electricity consumption data. The demand for electricity is based on several factors such as occupancy, variety of appliances used, weather, etc. However, taking into account every single factor that influences the consumption of electricity would make the model really complex, unstable and unreliable for future forecasting. As a result, approaches that make use of the time series data to capture time-dependent variations have been adopted in building prediction models. Past research shows that analysing time-series data using conventional approaches and AI based approaches such as ANN, ARIMA, SVM and Fuzzy based techniques are good at making short term predictions but are poor in making mid-term and long-term predictions.

Therefore, this paper talks about the use of RNN and LSTM to forecast electricity consumption over short-term, mid-term and long-term time intervals. Both these models appear to minimize the root mean square error compared to the other models based on AI and conventional approaches. The RNN and LSTM models have achieved a Root Mean Square Error of 0.1 for all the cases on an avearge.

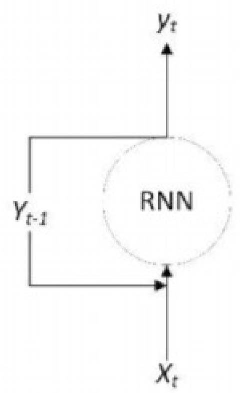
The most commonly used models in predicting electricity consumption are Support Vector Machines (SVM), Autoregressive integrated moving average (ARIMA), Linear Regression, and Artificial Neural Networks (ANN). It was observed that the ARIMA models perform extremely well for short-term forecasting. It is also observed that non-linear models achieve better results for short-term forecasting. Among Multiple Regression, ANN, Genetic Programming (GP), DNN and SVM, ANN happens to perform the best. Deep learning models are highly capable of predicting time-series data as they require very less feature selections. But, all these models that have been mentioned are not good at mid-term and long-term prediction.

RNN is a deep learning model model (deep neural network) that has a feedback loop from its past inputs to learn from its earlier sequences of data. A RNN adds the past output (y_{t-1}) to its current input (x_t) in a recursive manner and the current output (y_t) learns from the past sequence, using y_{t-1}. The past sequences of data carries on the past results with a combination. As a result, at every step the output is influenced by its past sequence and this happens throughout the sequence. Therefore, the RNN's equation is

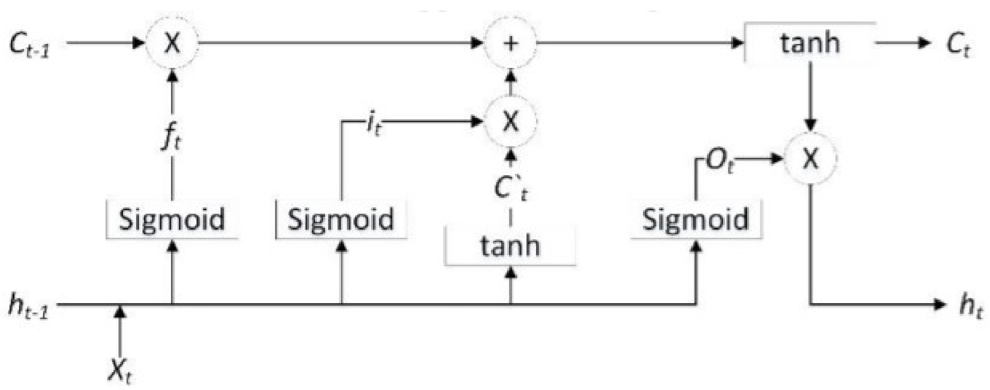
$$y_t = \tanh(W_r y_{t-1} + W_n x_t)$$

where W_r = weight given to the past output (y_{t-1})
W_n = weight given to the current input (x_t)

The RNN here has 100 hidden units sequentially connected to each other. It was trained for 300 epochs using a batch size of 20. Adam optimizer has been employed for obtaining high accuracy.



An **LSTM** when compared to RNN has a complex architecture. It is composed of gates that are used to filter and carry forward the past data. The current input x_t is added to h_{t-1} (past output). The f_t gate decides if C_{t-1} (cell state) is carried forward to generate the current output (h_t). Unlike the RNN which carries on y_{t-1}, the LSTM decides as to which data is carried forward through the cell state (C_{t-1}). The output gate creates the final output. The cell state C propagates the h_{t-1} to the next time stamp. But this C is calculated by the LSTM and therefore, the LSTM has a contol over C and h_t instead of directly generating the output, as in the case of RNN. The LSTM here has 100 LSTM units and the last layer has one dense layer before providing the output. It was trained for 300 epochs using a batch size of 20. Adam optimizer has been employed for obtaining high accuracy.



It is observed that RNN and LSTM perform as good as ARIMA for short-term forecasting and it outperformed all the other models for mid-term and long-term predictions with high accuracy.