Graphic User Interface Design for Electricity Load Prediction of New York State

Amanda Yilan Zhu Advisor: Dr. Junhui Zhao, Dr. Bijan Karimi
Department of Electrical and Computer Engineering, Tagliatela College of Engineering
University of New Haven

INTRODUCTION

With the growth of exponential electricity consumption, it has become necessary from households to energy distribution industries to monitor and forecasts for effective future planning and operations.

In this project, a non-linear regression model is constructed to forecast the day electricity load for a given date, based on provided historical hourly loads of the entire New York state, divided into 6 parts. As weather is one of the prime factor that affects electricity usage, 6 corresponding weather observation stations were selected to apply the hourly temperature as another set of training dataset. Neural Network is the main algorithm used for this prediction model. The prediction network is trained with annual hourly data, up to from 2008 to 2014 and tested on real data from 2015. The models are shown to produce highly accuracy load forecasts with average accuracy up to 90%.

FEATURES

Temperature

Climate change is one of the leading factors that effects on electricity consumption behavior and pattern. Meanwhile, the load is also correlated with the use of heating and cooling devices in large scale .

Data Normalization

Data normalization is the process of extracting values from a single column and placing them in a single reference table. As the initial step of the project, the obtained data in various formats could only be utilized after this step is complete.

Artificial Neural Network

Artificial Neural Networks (ANN) are important algorithms used for forecasting calculations. ANN perform accurate classification and even discover new trends or patterns in data. Basic Neural Network is usually composed of three layers: input, output, and hidden layer. Each layer would have nodes and nodes from input layer are connected to the nodes from hidden layer. Nodes from hidden layer are connected to the nodes from output layer. Those connections represent weights between nodes[2].

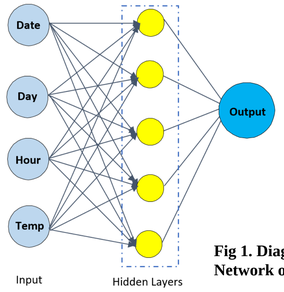

Fig 1. Diagram of Neural Network of forecasting system

Hadoop

MapReduce is a programming technique for analyzing data sets that do not fit in the memory. In this project Mapreduce is utilized for normalizing the extracted temperature data to a range of [-1,1] to improve the accuracy and the training speed of the neural network.

METHODOLOGY

Data Acquisition

Electricity load hourly data is obtained from New York Independent System Operator (NYISO). Records of the dataset have been collected and downloaded from 2008 January to December 2015. This contains detailed data of the hourly consumption of eleven primary substations with corresponding electrical transmission owners[3] .

Air Temperature data is acquired from Mesowest API and loaded directly to the MATLAB as training data. This is an online database for a collection of all the accessible weather observation stations throughout the United States by the University of Utah Department of Atmospheric Sciences. Mesowest's API is web services that provides data in response to specific requests (query string parameters). Weather Observation on the geographic center of the load zone are selected for weather information.

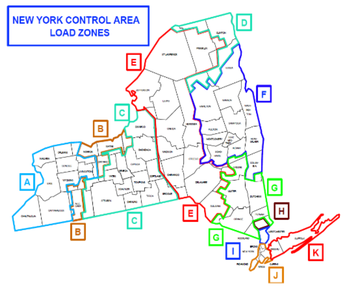

Fig. 2 New York Control Area Load Zones from NYISO

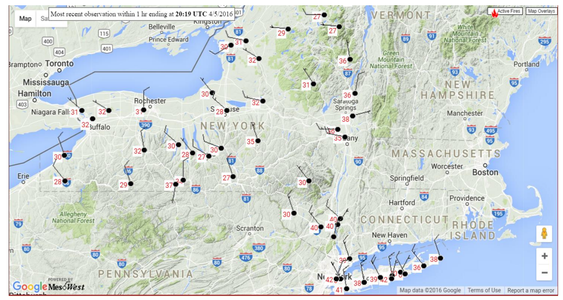

Fig 3. New York State with multiple weather observation stations from Mesowest

Basic structure of forecasting model

Data collection.
Arrays that contain month, date, day, hour and air temperature are formatted and normalized as the training set. Hourly electricity load values are set into arrays of the same size (number of rows) as the training sets.
Initialization and configuration of the network.
Determine the type of ANN algorithm for evaluation and have the parameters set for best network performance.
Network Training.
The network would be trained multiple times using every year's training dataset as the input and the corresponding hourly electricity load as the target.
Network Validation.
The trained network would predict a year's length of hourly electricity load. By comparing the predicted values with the real numbers of 2015, we could calculate the mean absolute percentage error (MAPE), which would be the main criterion for testing the accuracy of the prediction system.
Load Prediction.
Based on input information (area, the number of years for training and date for prediction) by the user, forecast and the effectiveness of the system is presented.

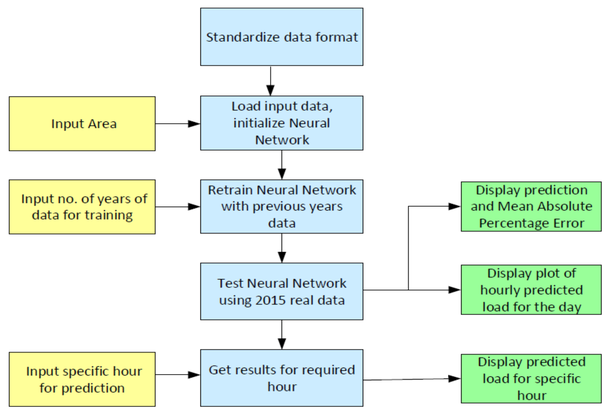

Fig 4. Basic structure of forecasting system

RESULTS

The final predicted results have an average accuracy of 93%-92%. The below are the plots comparing the forecasted hourly loads of 2015 comparing with the real electricity load of 2015.

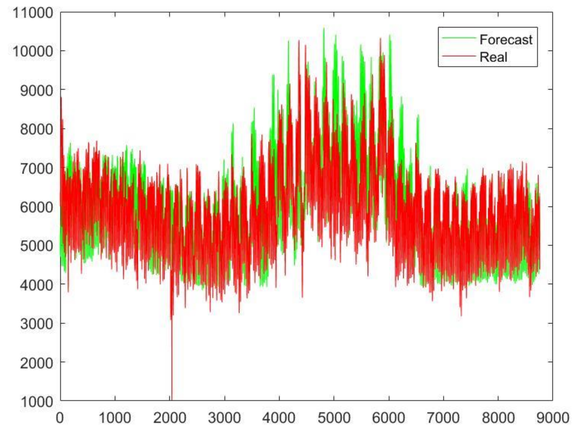

Fig 5. Plot of hourly forecast electricity load (green) vs. real load (red) of 2015

According to the information input by the user on the GUI, load forecast for a specific date would be produced.

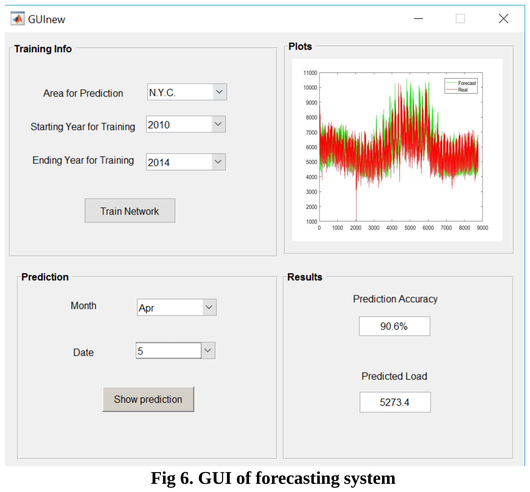

Fig 6. GUI of forecasting system

REFERENCES

1. Ali Akbar Salehizade, Mohammad Rahmanian. "Analysis of Temperature Changes on Electricity Consumption in Fars Province". Mediterranean Journal of Social Sciences. Rome-Italy, May 2015 Vol 6 No 3 S2.
2. Mirza Cilimkovic, "Neural Networks and Back Propagation Algorithm", Institute of Technology Blanchardstown. Blanchardstown Road North. Dublin 15. Ireland.
3. New York Independent System Operations. Federal Energy Regulatory Commission. Website: http://www.nyiso.com/public/markets_operations/market_data/load_data/index.jsp