SENIOR PROJECT FINAL REPORT

Project Subject: Electrical Load Forecasting

Name: Hüseyin Eryılmaz **ID: 20160808058**

The Advisor: Asst.Prof.Dr. Mustafa Berkay Yılmaz

1. Introduction

Electric load forecasting is very difficult process nowadays. The market has not a constant demand and it varies from hour to hour. Therefore, it is not easy to forecast to energy demand. If it is an industrial region then industrial plan and production effect the energy consuming. If it is a population area then weather conditions and other specific actions can be effect energy consuming. So, why electric load forecasting is important? Short-term electric load forecasting supply an uninterrupted power to home and industrial regions. Also it decrease the power loss.

2. Electric Load Forecasting

2.1. Problem Definition

Load forecasting is a technique used by power or energy-providing companies to predict the power/energy needed to meet the demand and supply equilibrium. Load forecasting very helpful for companies. Companies have good advantage on management. Companies increase the efficiency with Electric Load Forecasting. Because they can predict the needed load and they can make a good plan for distribute the electric. They can work with full capacity.

There are three different load forecasting, they are short-term, medium-term and long-term. Short-term predicts a few hour, medium-term predicts to a year and long-term predicts more than a year. The most useful approaches are medium-term and long-term approaches because they supply more economic advantages.

2.2. Good Sides of Electric Load Forecasting

- Company can make a good future plan for electric consumption.

-Long term prediction reduce the risk because it have an idea to future. Certainly, it can’t predict true always but at least it can give a rate. Company can make economic plans according to this prediction and can make economic investments.

-Power companies can determine needed resources to future. It will show the future demand. So company can find the needed resources. By the way they can plan short-term, medium-term and long-term and they can supply needed power.

-Company can plan new energy production stations with electric load prediction because company can see the future demand before. They can construct new energy stations and new distribution infrastructures.

-Maintenance of power systems can be plan before. They can make the maintenance when demand is low. Bad effects of maintenance are can be decrease with this way. Because people not need much energy at that time.

2.3. Difficulties of Electric Load Forecasting

-We use a lot of variables while prediction, one of this is weather conditions. If there is a storm or it is too hot then we increase the power demand. Weather conditions can change. That changes also effect our short-term load prediction. For example, temperature start the decrease while winter is coming. If winter colds come 10 days earlier than our prediction will belie us.

-One approach is manual approach to forecasting. Companies find forecasters and they predict it manually. Forecasters use upcoming events and a some dataset that obtain from previous events. This approach can't acceptable because there is a lot of factor that effect the prediction. I mean there is a kaotic field. Instead of manual forecasting, companies have to try find technological approaches. They have to fire forecasters and make investment to science.

-Another difficultly is different meter types. There are two type of meters. One of is smart meter and other one is traditional meter. Also there is different tariffs. Meter types and tariffs effect the power consumption. The software have to include all these differences and it has to process them.

-Price changes always effect the usage. People may try to save their money sometimes. By the way, we can't obtain an accurate data.

-Some similar seasons can give different accuracies because nature can change. When this happens we can't obtain a specific data and we cant give right prediction.

-A lot of factor effect the load prediction. If we produce a well worked software then we need to process all of these factors. For example, humidity, pressure, wind direction, wind speed can effect consumption.

3. Electric Load Forecasting Techniques

3.1. Regression

Regression is a statistical technique used in load forecasting in creating a model of a relationship that exists between consumption and variables such as consumer age, class, weather and date **[1]**

Statistical regression technique can be further classified into the simple linear regression model and the multiple regression models. In the simple linear regression model, the forecast and predictor variables are assumed to be related linearly with each other. As an example, let consider certain data as given in model is shown in Fig. 2 below and expressed through simple relation as equation:

$$y = \beta_o + \beta_1 x + \varepsilon_1 \tag{1}$$

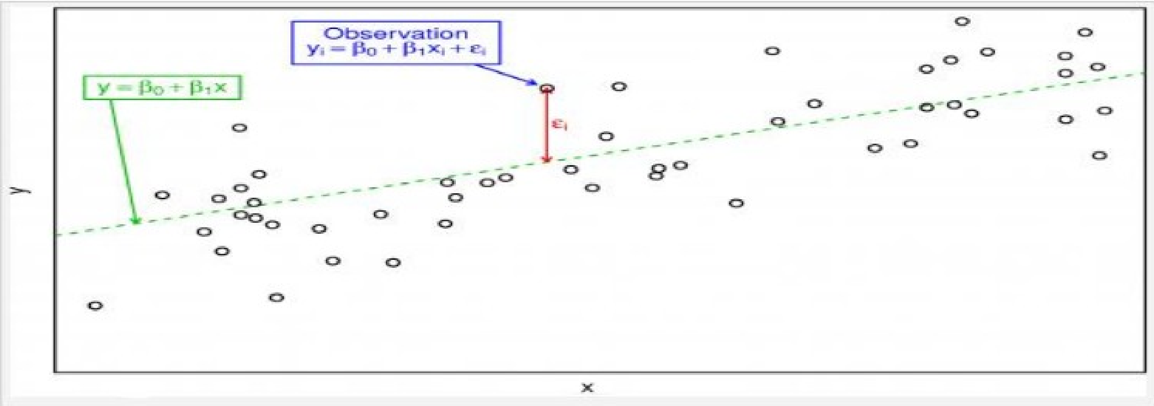


Figure 1.

In relation(1) x and y are variables quantities. The x quantity also called the regressor, independent or explanatory variables whereas the y acts as forecast variable and termed as the regress and, dependent or explained variable. The parameters β0 and β1 determine the intercept and the slope of the line respectively. The 1 represents the deviation from the underlying straight line model assumed as the random “error”. In above simple linear regression model evaluating the variable x through available data the forecast value y can be predicted with a margin of 1 **[2]**.

The multiple-variable regression model is similar to simple linear model yet it involves multiple variables. The general form of multiple variables regression is given in relation (2) below.

$$y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots \beta_k x_k + e_i \qquad (2)$$

In above relation yi is the variable to be forecast and x1,...,xk are the k predictor variables whereas the coefficients β1,...,βk gives the effect of each predictor after taking account of the effect of all other predictors in the model. Similar to simple linear regression model ei represents the error involved during forecasting. The evaluation of variables involved in the multiple variable regression are carried out through the application of various available statistical tools and predictors **[2]**.

3.2. The Time Series Method

This technique is based on the concept that information obtained on electricity demand usually has an internal structure, which might be in the form of an auto correlation, seasonal variation or trend. These methods explore and detect the internal structure and are utilized in such fields as digital signaling, processing, electric load forecasting, as well as economics. For instance, Autoregressive Moving Average(ARMA), Autoregressive Integrated Moving Average (ARIMA) and Autoregressive Moving Average with Exogenous Variables (ARMA-EOV) are examples of time series techniques.

In this case, ARMA models are at most times utilized for stationary procedures and are one of the popular linear models in time series forecasting during the past three to four decades. ARIMA is an addition to ARMA, but it also involves non-stationary processes. It is vital to note that both procedures utilize load and time as input parameters. It can be observed that ARIMAX is the most efficient and accurate tool to use in load forecasting among these models because load usually depends on time and weather, which the method applies **[2]**. However, the main disadvantage of this method is that the simple extrapolation of past trends may produce inaccurate results when definite events disturb past arrangements of the series **[3]**. This implies that there are circumstances that tend to have effects on the historic trends thereby causing the results to be imprecise.

The major components of time series methods are a base level, a trend and a cyclic fluctuations. Based on these components various time series methods such as the Moving-Average Model (MAM), the Exponential Smoothing Model (ESM) and Holt's Method for forecasting have been introduced and used in forecasting. The MAM model is based on;

-the observations recorded in period t

-the necessary calculations in period t

-application of observation and calculation to forecast for period (t + 1).

Mathematically it can be written as;

$$A_t = (x_t + x_{t-1} + x_{t-2} + \cdots + x_{t-n} + x_{t-n+1})/n \qquad (3)$$

where xt represents the observation made in period t, At denotes the moving average calculated after making the observation in period t. The ESM weighs recent observations more than older ones and be written in following mathematical form.

$$forecast\ F_t = \ S_t = \alpha x_t + (1 - \alpha)S_{t-1} \qquad (4)$$

where α is the smoothing constant having values between zero and one whereas St is the smoothed value of the observations based on our "best guess" as to the value of the mean **[4].**

3.3. End-use Models

The other load forecasting method is called End-use model, which estimates the energy consumed by utilizing information obtained from its end users. This information might be in terms of consumers' use, size of houses, age of consumers, and applications, among many others. The technique relies on statistical information from consumers accompanied by changes, which form the basis of the forecast. This implies that end-use models will focus mainly on the various uses of electricity in commercial, industrial and residential sectors. This method is based on the concept that electricity demand will always depend on the end user demand for refrigeration, cooling and lighting **[5]**.

Mathematically the end-use method is a set of equations designed to disaggregate end user's household's total annual energy consumption generally through Residential Energy Consumption Survey (RECS) and based on these disaggregated values, particular weight has been given to produce population estimates of total and average energy end uses at various levels of geography. A visualization of end-use modes is given in Figures 2 and 3 below.

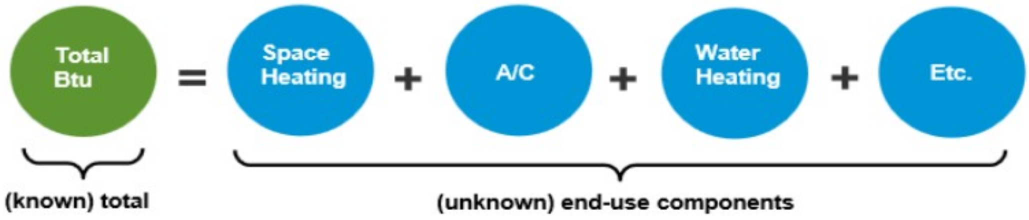


Figure 2. Total energy consumption for End-use Models.

The method is extremely accurate as the models explain electricity demanded by putting it as a function of its various applications. Nevertheless, the accuracy of this model relies entirely on the quality and amount of information obtained from end users. For instance, in some circumstances, information obtained on consumer age will be more relevant as compared to the size of houses. This is because there are definite areas that contain many youths who tend to overuse electricity through their various electronics such as heaters, fridges, video games, and computers, among many others. This implies that the size of a house occupied by an individual might be small, but he or she might be consuming a lot of electricity. It will for this reason be crucial for the forecaster to identify the appropriate variables to utilize by analyzing factors such as the method of distribution, as well as demand of the end users.

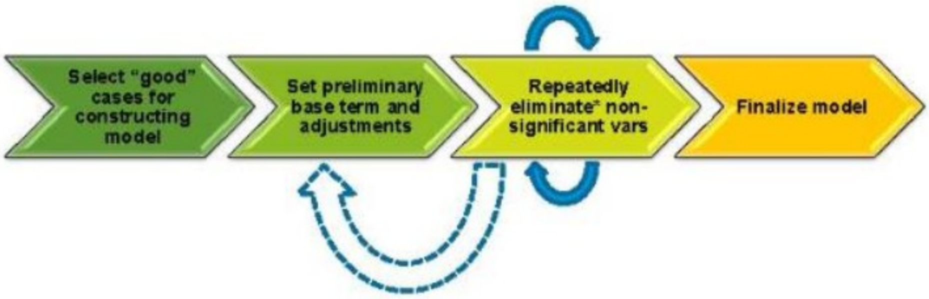


Figure 3. Flow diagram for the parameters of End-use Model for forecasting.

Apart from this, End-use forecasts usually demand more information on end users and less on historical data. The end-use approach can be integrated with the econometric method in introducing behavioral aspects into the equation making the procedure much-more accurate.

3.4. Artificial Neural Networks(ANNs)

Apart from the previously examined methods, there are other accurate techniques, which are based on artificial intelligence. Artificial neural networks are good examples of the advanced methods because they have succeeded in solving power system issues relating to load forecasting, design, protection, planning and control of electricity **[6]**. The main advantage of utilizing the neural networks is that they have the ability to map complex linear relationships, which exist between energy consumed and factors affecting consumption **[7]**.

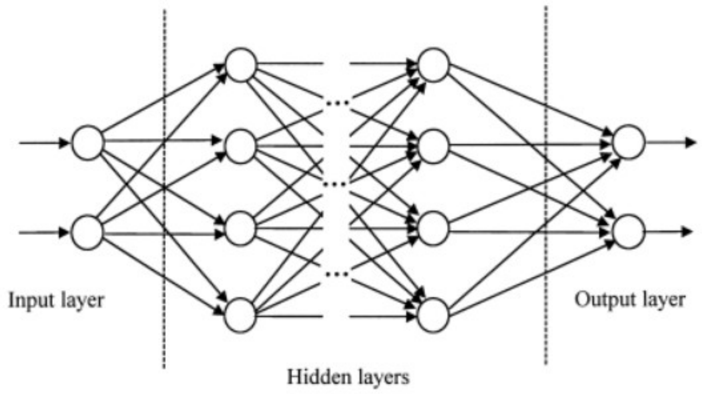


Figure 4. A Neural Network Model

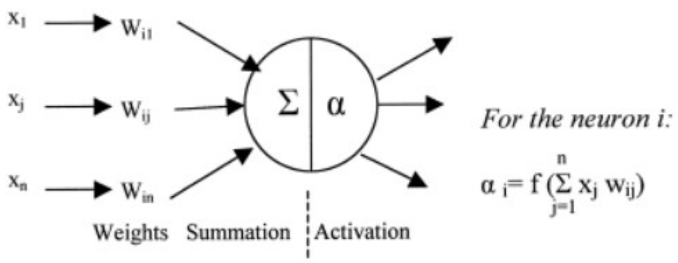


Figure 5. A Neuron in feed forward.

A schematic diagram of a typical multilayer feed-forward neural-network architecture is shown in Figure 4 comprising of an input layer, some hidden layers and an output layer. The information received through input nodes at input layer is processed and forward to hidden layers nodes where information is further processed stored processed (summation). The processed information is forward to output nodes at output layer to predict the future energy consumption **[8]**. The information processing in ANN is shown in Figure 5.

Artificial neural networks are mostly applied in the short-run period, which runs between one hour and one week. In order to make this approach more accurate, forecasters should make decisions on the size, as well as the number and neural utilized while applying various degrees of freedom. In this case, it will be critical to analyze various facets of artificial neural networks in order to develop suitable models. These aspects include the methods of training and network architecture **[9]**.

3.5. Genetic Algorithms (Long-term Forecasts)

The other artificial intelligence-based technique involves the use of genetic algorithms, and it deals with the utilization of genetic programming in predicting long-term consumption of electricity. The results obtained from this procedure are reliable as the method, if successfully undertaken, usually has a low chance of errors. This approach also uses a genetic neural network model that can adapt and learn, which means that it is robust. At this point, the model includes the ideas and experiences of experts in the field in order to create a comprehensive effect whereby, all variables that affect power load are reflected. This examination makes it apparent that the technique uses a numerical optimization approach. The main advantage of using genetic algorithms is that they include all aspects of computing such as imprecision, non-linearity, robustness, as well as uncertainty.

3.6 Support Vector Machine(SVM)

Support Vector Machine is a method used for classifying information or data, to obtain good generalization on a limited number of learning patterns **[9]**. It is a technique based on a concept known as structural risk minimization. For instance, recurrent neural network is part of this technique and runs under the idea that every unit is an output of the entire network. There are two main categories for SMVs namely Support Vector Classification (SVC) and Support Vector Regression(SVR). SVM is alearning system that applies a high dimensional feature space and yields prediction or forecasting functions that are expanded on a subset of support vectors. SVM can generalize complicated gray level structures with only a very few support vectors and thus provides a new mechanism for image compression as pointed out by Debasish Basak **[10].**

3.7. Fuzzy Logic

The fuzzy logic model is a rule-based system that applies fuzzy rules, which are control decision mechanisms. The rules usually adjust the impacts that definite stimuli have on energy demand as a factor. The model maps on a set of input variablesto a set of output variable simply using the IF-THEN logic statement to predict and forecasting. These types of mappings allow the incorporation of expert knowledge with Fuzzy Logic models. These modelscan provide algorithms that transform linguistic strategy obtained from experts into automatic strategies. Fuzzy rules are often integrated with neural networks so as to train artificial neural networks, as well as acquire accurate results in terms of load demand forecasting. The advantage of this hybrid system is that it utilizes the benefits of both artificial neural networks and fuzzy inferences thereby, formalizing and handling knowledge and experience of expert forecasters **[11]**.

3.8. Recurrent Neural Networks(RNNs)

A Recurrent Neural Network feed the current input with outputs from previous section. We discussed ANNs in section 3.4, all inputs and outputs are independent in a traditional ANN. If we need to predict the next step as time series then we need to use previous sections so RNN can be useful such like this situation. RNN has a memory and it remembers the previous situations. It feeds the hidden layer with previous data and it calculates the predictions.

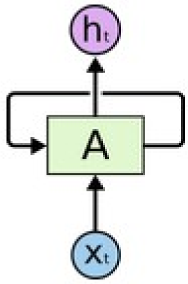


Figure 6. A cell of RNN.

3.8.1. Long Short-Term Memory(LSTM)

An LSTM network is a special variant of an RNN. LSTM capable of learning long-term dependencies. LSMT introduced by Hochreiter &Schmidhuber then a lot of people contribute to LSTM. Today people can use LSTM as an effective machine learning method. The ability of remembering is their default skill. It supplies an efficient learning to long period works. All RNN architecture have a repeating chain form. A single tan(h) layer supply this chain. Figure 7 shows to us this chain form.

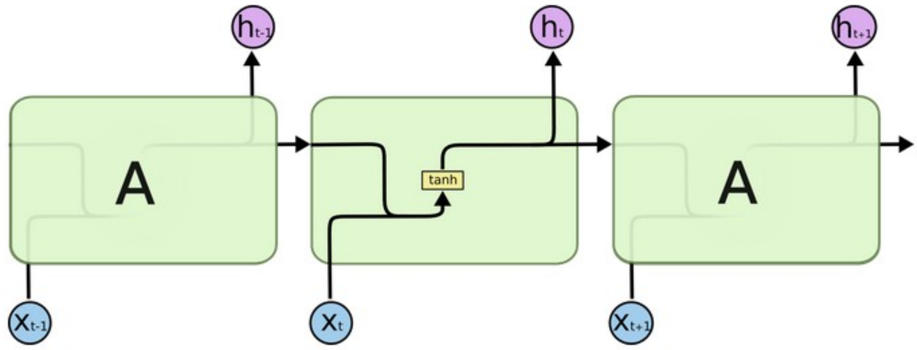


Figure 7. A standard Recurrent Neural Network with single tanh layer.

LSTM present 4 layer neural network instead of a single one. That is the difference between a traditional RNN and LSTM. Figure 8 shows to LSTM architecture.

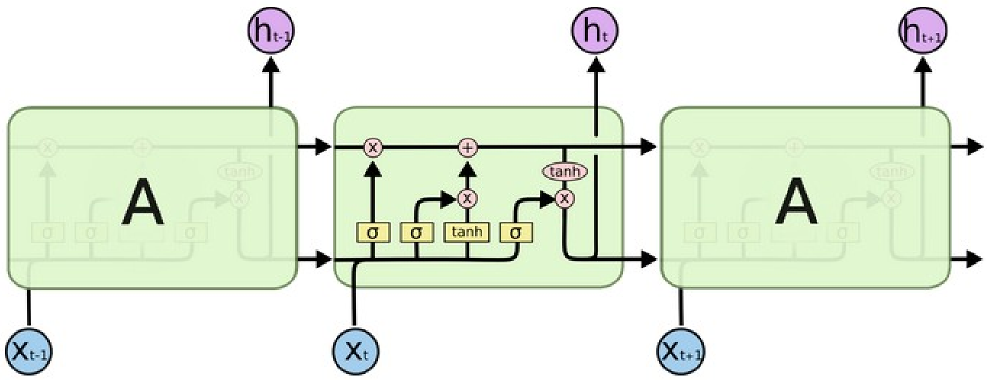
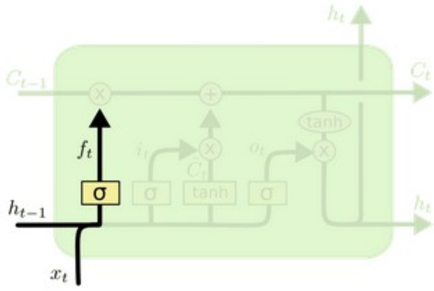


Figure 8. An LSTM architecture.

3.8.1.1 Long Short-Term Memory(LSTM)'s Steps

– 1st Step

LSTM decide to which data will be keeped and which data will be removed in first step. A sigmoid layer made this decision. We call this layer like "Forget Gate". It looks like $h(t-1)$ and $x(t)$ in our simulation(Figure 9) , and outputs a number between 0 and 1 for each number in the cell state $C(t-1)$. If it return 1 then output will completely keep, otherwise if it 0 then it will completely removed. Figure 9 simulates the first step.

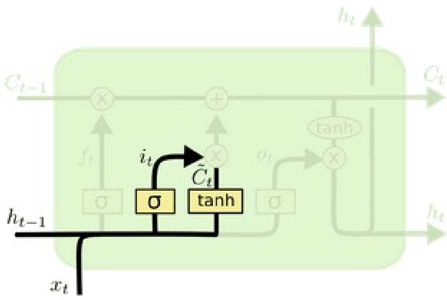


$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] \ + \ b_f\right)$$

Figure 9. 1st step of LSTM

– 2nd Step

LSTM decide to which information will store in the cell. This process has two section. First a sigmoid layer decides which values will have to update. After then, tanh layer generates a new vector called as *C(t)* and it include the state. Finally, we combine these two and we make an update.

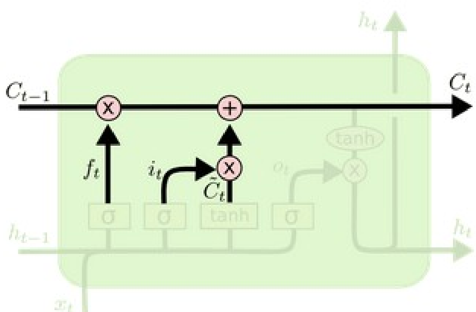


$$i_t = \sigma\left(W_i \cdot [h_{t-1}, x_t] \ + \ b_i\right)$$
$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] \ + \ b_C)$$

Figure 10. 2nd step of LSTM

– 3rd Step

We update the cell in 3rd step. C(t) shows the updated new cell state. C(t)-1 shows the old state. We multiply the forgetting information f(t) with old state C(t)-1. Then we sum with new candidate values *i(t)∗C(t)*. The process showed at Figure 11.

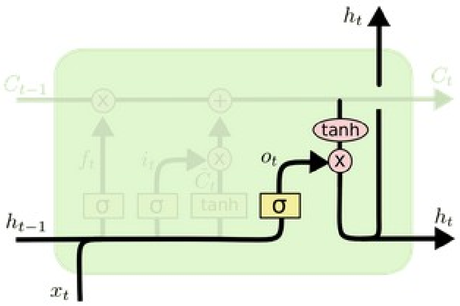


$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t$$

Figure 11. 3rd step of LSTM

– 4th Step

We have to decide the output of LSTM in last step. The output will be a filtered version of cell state. A sigmoid layer will decides the output of cell state. Then we multiply with output and cell state.//// Figure 12 shows that process.



$$o_t = \sigma\left(W_o\ [h_{t-1}, x_t]\ +\ b_o\right)$$
$$h_t = o_t * \tanh\left(C_t\right)$$

Figure 12. 4th step of LSTM

3.9 Convolutional Neural Network(CNNs)

Convolutional Neural Network is a feed-forward neural network. Like the traditional architecture of a neural network including input layers, hidden layers and output layers, convolutional neural network also contains these features and the input of the layer of convolution are the output of the previous layer of convolution or pooling. Of course, they still have some unique features such as pooling layers, full connection layers, etc. The number of hidden layers in a convolutional neutral network is more than that in a traditional neural network, which, to some extent, shows that the capability of the neural network. The more the hidden layers are, the higher feature it can extract and recognize from the input. People always use convolutional neural network in computer vision, such as face recognition, image classification**[12]**.

3.9.1 Convolution Function for 1-Dimensional Data

Usually, by means of Google's TensorFlow, we can directly use the function "conv2d" to do the computation of convolution in a CNN model which usually accepts a 2D image as its input. But, as our data of the electric load belongs to 1D time series data, it is necessary to change the function to help us do the computation. The new function we adopt in the model accepts a number of electric load data and other properties such as the number of the filters, width of the filters and stride as its input. Note that the height of the kernel here is not applicable. Then we use a Gaussian distribution to initiate the value of the filters and zero to initiate the biases. The outputs of our function are a number of matrices, and the specific number is decided by the number of the filters. CNN model will extract some features from them and they also will be the input of the next pooling layer after the computation of activation function**[12]**.

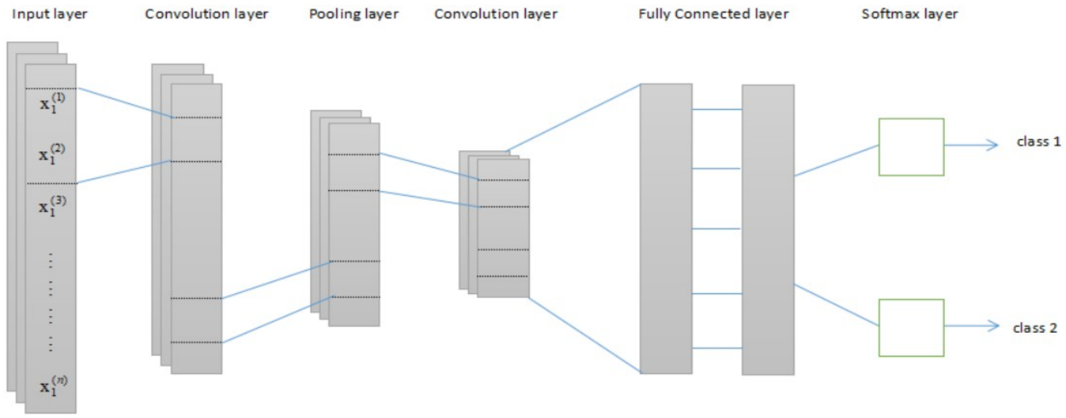


Figure 13. Architecture for 1-Dimensional Data

4. Trials and Observations

4.1. Long Short-Term Memory(LSTM) to Electic Load Forecasting

I used Python3 programming language and I tried an LSTM model to predict electric load forecasting.

The used libraries are:

- math (to basic mathematical operations)

- numpy (to matrix operations)

- pandas (to read data file and save results)

- sklearn (to data scale)

- keras (to build LSTM model)

- matplotlib (to visualization data and results).

Here is the model that I used ;

```
#Build the LSTM model

model = Sequential()
model.add(LSTM(50, return_sequences=True, input_shape= (x_train.shape[1], 1)))
model.add(LSTM(50, return_sequences= False))
model.add(Dense(25))
model.add(Dense(1))

#Compile the model
model.compile(optimizer='adam', loss='mean_squared_error')

#Train the model
model.fit(x_train, y_train, batch_size=1, epochs=1)
```

Figuere 14. The LSTM model

I use the same train-test data ratio in all trails. Train data ratio is 80% and test data ratio is 20%.

4.1.1. Dataset

I found a dataset from kaggle.com. It contains American Electric Power Company Inc.'s hourly electric consuming data of some USA regions between 2005-2018.

Here is the data source;

https://www.kaggle.com/robikscube/hourly-energy-consumption#AEP_hourly.csv

I split data to month, year, 2 years, 3 years, 4 years and 8 years. By the way, I can observe the difference between dataset and results.

4.1.2. Monthly Dataset

I take April, May and June data. April data contains;

April 2005, April 2006, April 2007, … , April 2018 and other months also just like April.

Here April data visualization;

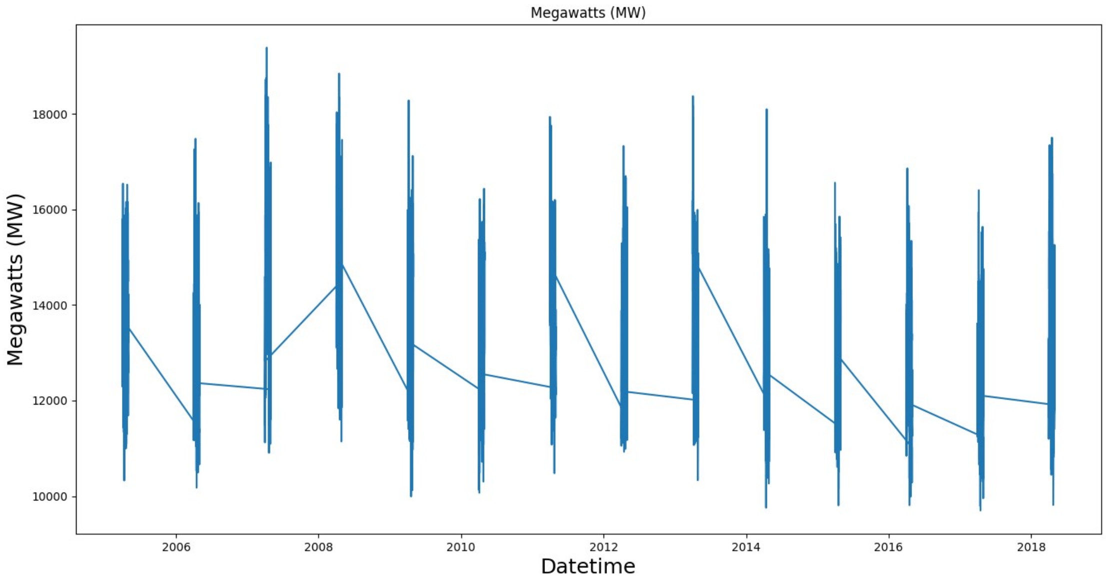


Figure 15. Aprils data

Here April predictions visualization;

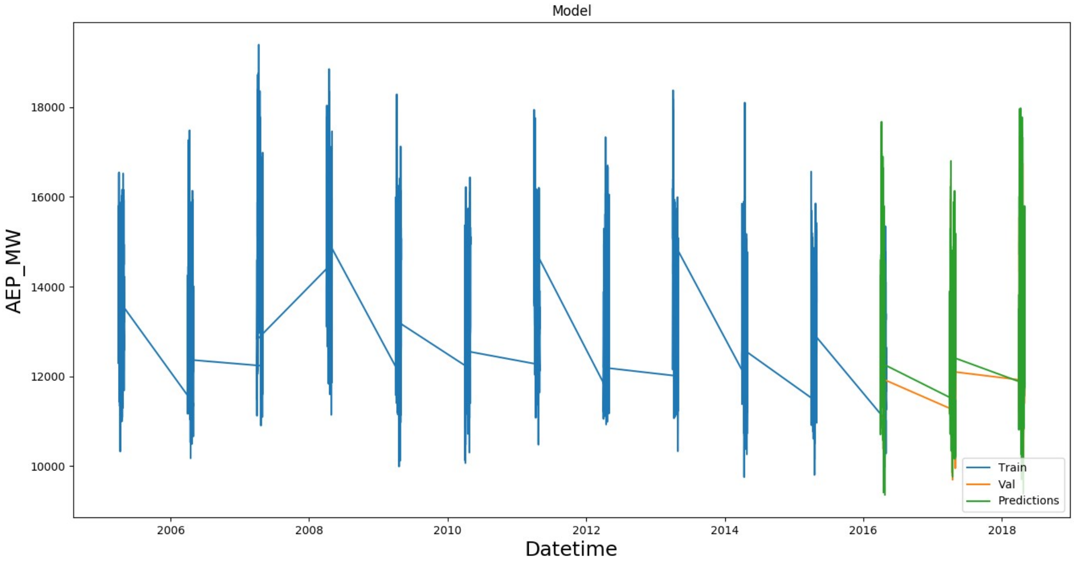


Figure 16. April Predictions

Here April, May and June predictions and accuracy ;

Months	Accuracy for Threshold 1000 Megawatt	Accuracy for Threshold 750 Megawatt	Accuracy for Threshold 500 Megawatt	Accuracy for Threshold 250 Megawatt	Accuracy for Threshold 100 Megawatt
April	0.98	0.96	0.82	0.45	0.17
May	0.98	0.92	0.72	0.45	0.21
June	0.95	.0.87	0.68	0.30	0.12

Table 1.

I need to select a threshold for accuracy so I select 5 different threshold because it can’t predict the same number in real data. Also we can observe the algorithm success with 5 different threshold. For example, “Accuracy for Threshold 1000 Megawatt” takes difference between real data and prediction. If difference smaller than 1000 Megawatt than algorithm accepts result as success, otherwise it accepts the result as fail.

4.1.3. Yearly Dataset

I use 2005,2006,2007 and 2008 years.

Here visualization of 2005;

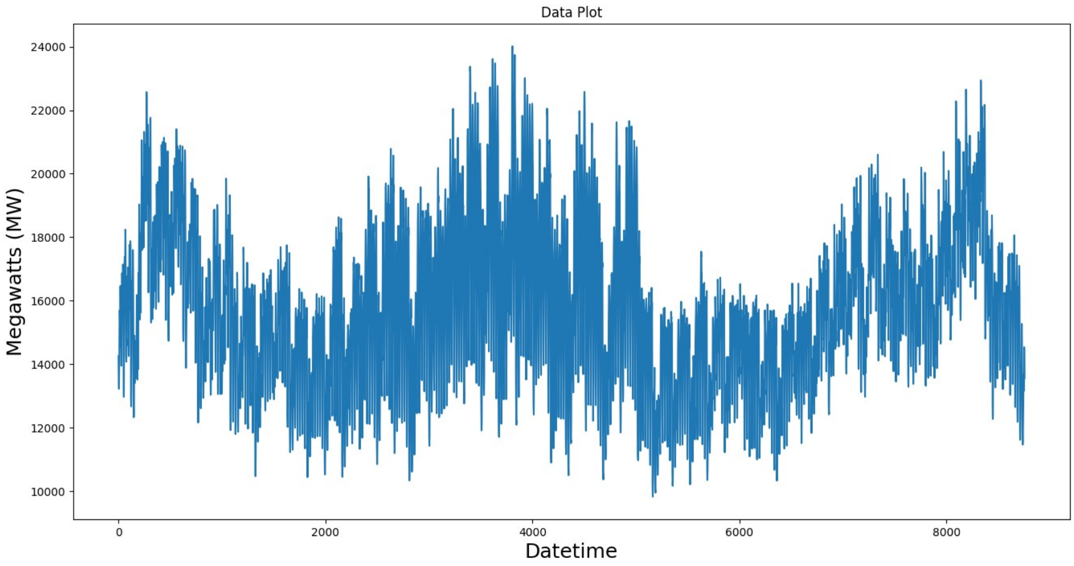


Figure 17. Electric load of 2005.

Here visualization of prediction;

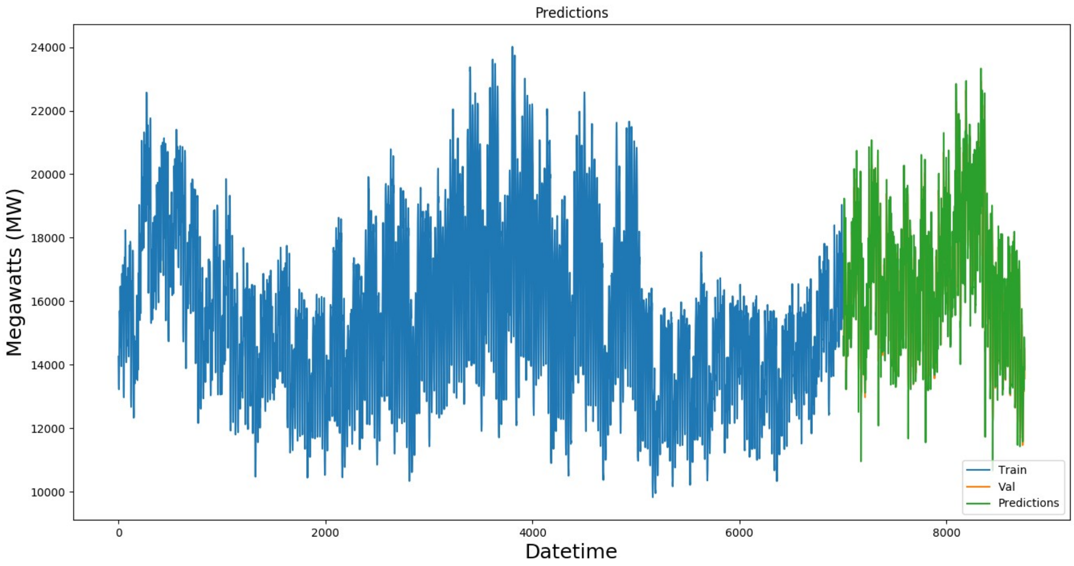


Figure 18. Predictions of 2005.

Here 2005,2006,2007 and 2008 accuracy;

Years	Accuracy for Threshold 1000 Megawatt	Accuracy for Threshold 750 Megawatt	Accuracy for Threshold 500 Megawatt	Accuracy for Threshold 250 Megawatt	Accuracy for Threshold 100 Megawatt
2005	0.95	0.90	0.79	0.54	0.26
2006	0.91	0.83	0.66	0.36	0.15
2007	0.89	0.81	0.59	0.31	0.12
2008	0.93	0.87	0.79	0.56	0.27

Table 2.

4.1.4. 2-Years Dataset

This section contains 3 dataset with consist of 2 years data. They are “2005-2006”, “2006-2007” and “2007-2008”. For example, “2005-2006” starts 1 January 2005 and ends with 31 December 2006.

Here visualization of “2005-2006” dataset;

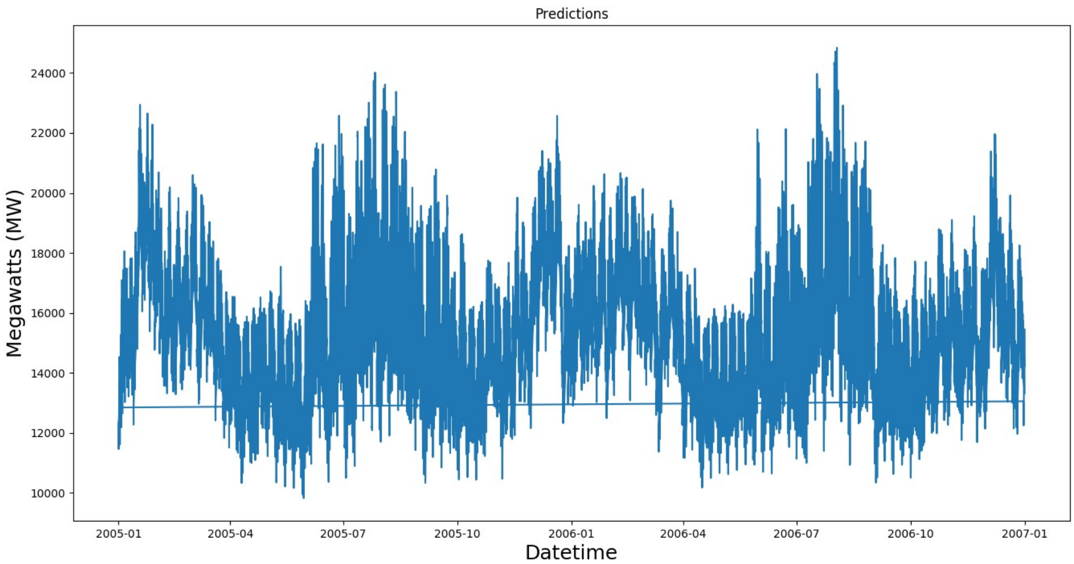


Figure 19.

Here visualization of “2005-2006”;

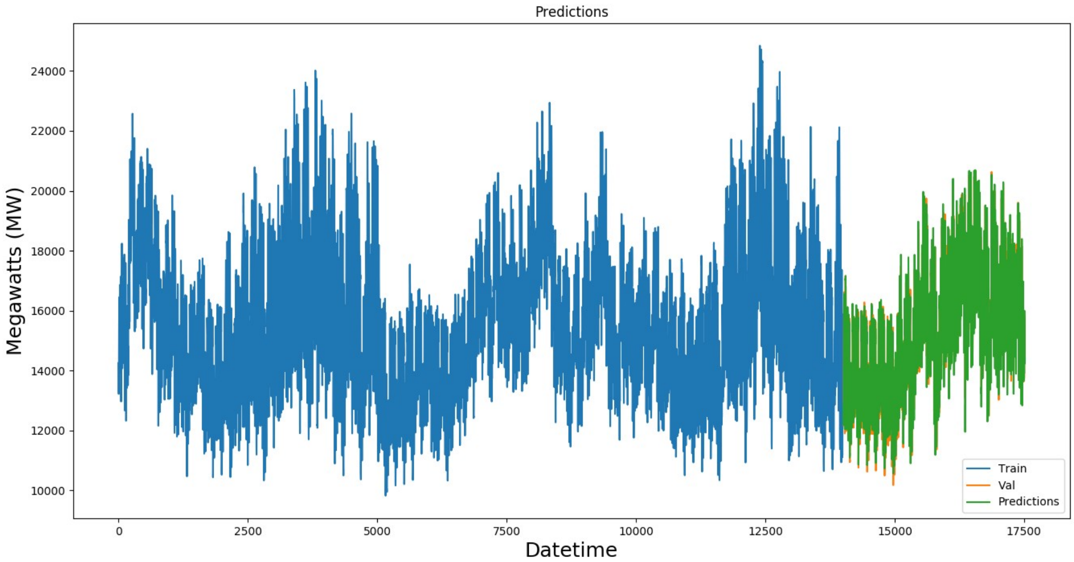


Figure 20.

Here the accuracy of “2005-2006”, “2006-2007” and “2007-2008” datasets;

Years	Accuracy for Threshold 1000 Megawatt	Accuracy for Threshold 750 Megawatt	Accuracy for Threshold 500 Megawatt	Accuracy for Threshold 250 Megawatt	Accuracy for Threshold 100 Megawatt
“2005-2006”	0.97	0.95	0.87	0.60	0.26
“2006-2007”	0.96	0.91	0.78	0.45	0.17
“2007-2008”	0.95	0.88	0.69	0.28	0.10

Table 3.

4.1.4. 3-Years Dataset

This section contains 2 dataset with consist of 3 years data. They are “Between 2005-2007” and “Between 2006-2008”. For example, “Between 2005-2007” starts 1 January 2005 and ends with 31 December 2007.

The visualization of “Between 2005-2007” dataset;

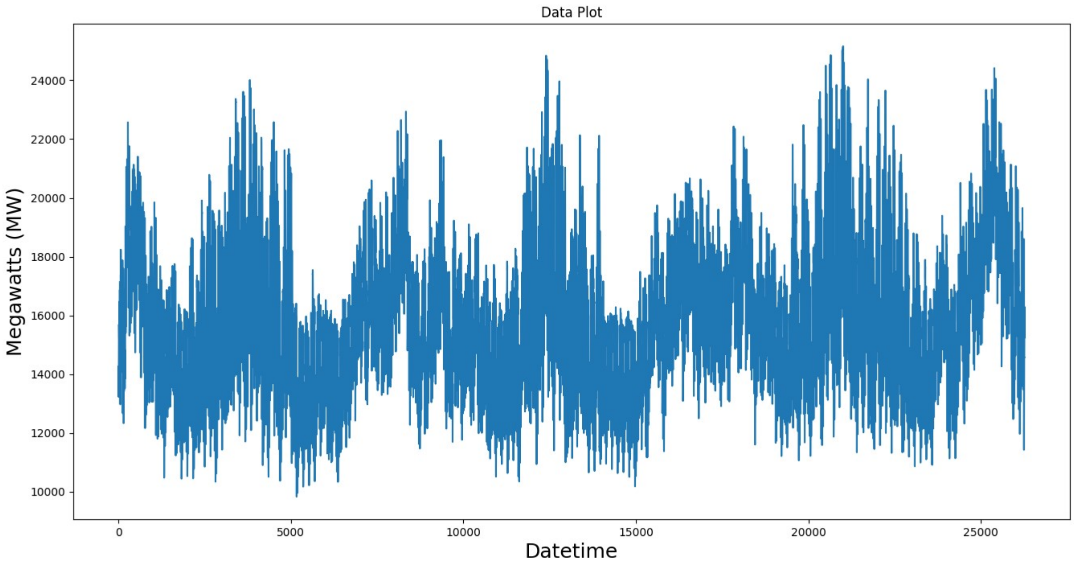


Figure 21.

The prediction of “Between 2005-2007”;

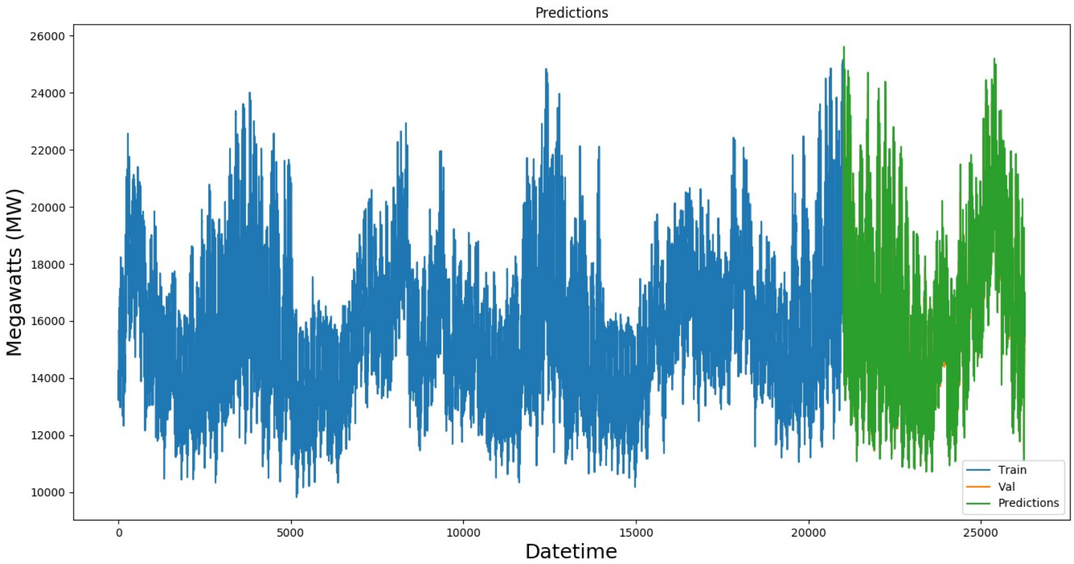


Figure 22.

Here the accuracy of “Between 2005-2007” and “Between 2006-2008” datasets;

Years	Accuracy for Threshold 1000 Megawatt	Accuracy for Threshold 750 Megawatt	Accuracy for Threshold 500 Megawatt	Accuracy for Threshold 250 Megawatt	Accuracy for Threshold 100 Megawatt
“Between 2005-2007”	0.96	0.91	0.74	0.39	0.15
“Between 2006-2008”	0.97	0.94	0.86	0.61	0.29

Table 4.

4.1.5. 4-Years Dataset

I have a single dataset here. It starts 1 January 2005 and ends with 31 December 2008.

Here the visualization of “Between 2005-2008” dataset;

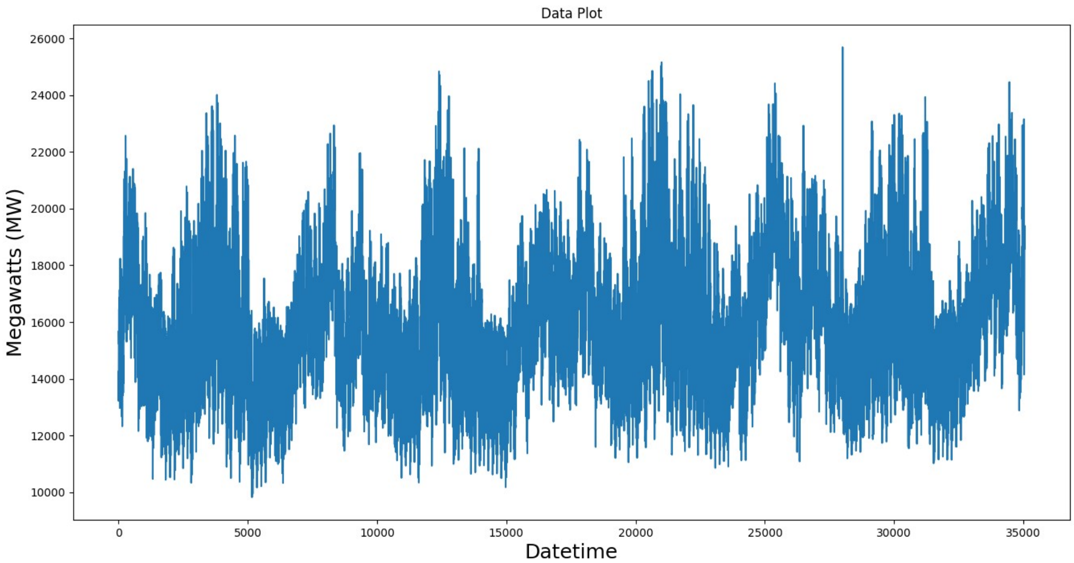


Figure 23.

Here the predictions of "Between 2006-2008" dataset;

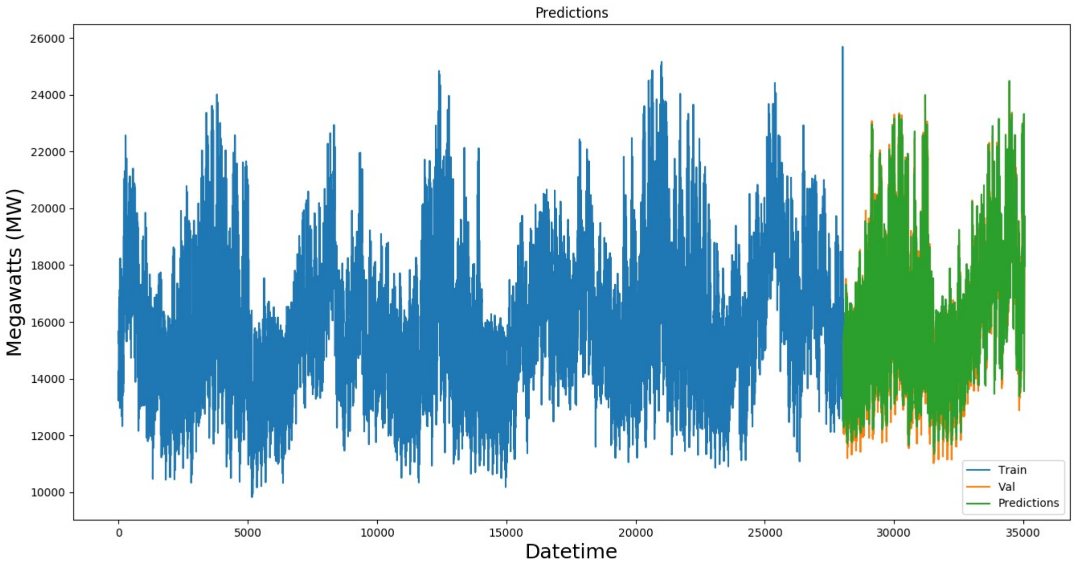


Figure 24.

Here accuracy of "Between 2005-2008" dataset;

Years	Accuracy for Threshold 1000 Megawatt	Accuracy for Threshold 750 Megawatt	Accuracy for Threshold 500 Megawatt	Accuracy for Threshold 250 Megawatt	Accuracy for Threshold 100 Megawatt
"Between 2005-2008"	0.97	0.94	0.85	0.58	0.26

Table 5.

4.1.6. 8-Years Dataset

This dataset starts 1 January 2005 and ends with 31 December 2012. It contains 8 years data.

The visualization of "Between 2005-2012" dataset;

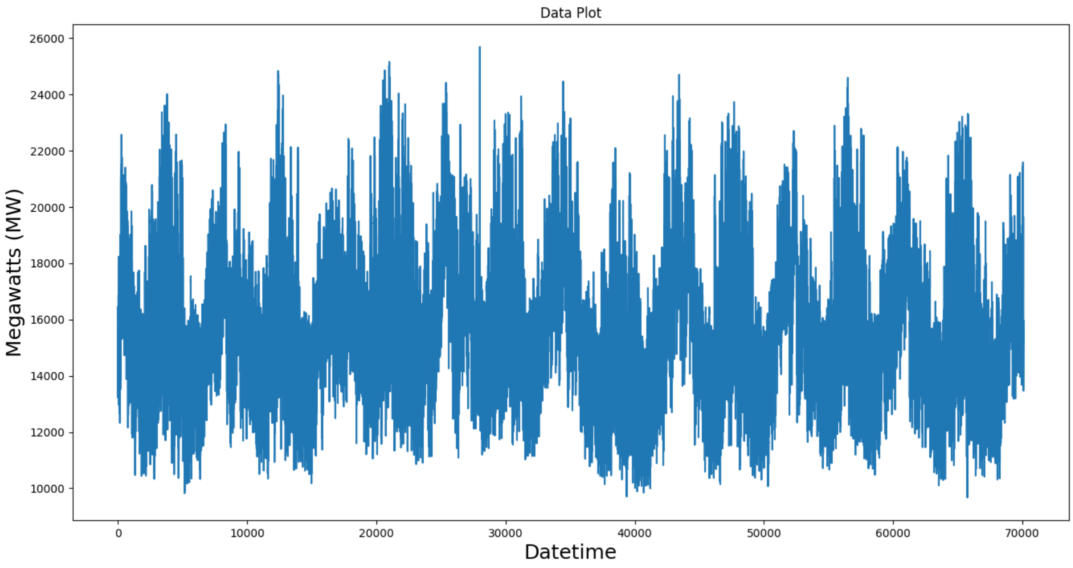


Figure 25.

Here the predictions of “Between 2005-2012” dataset;

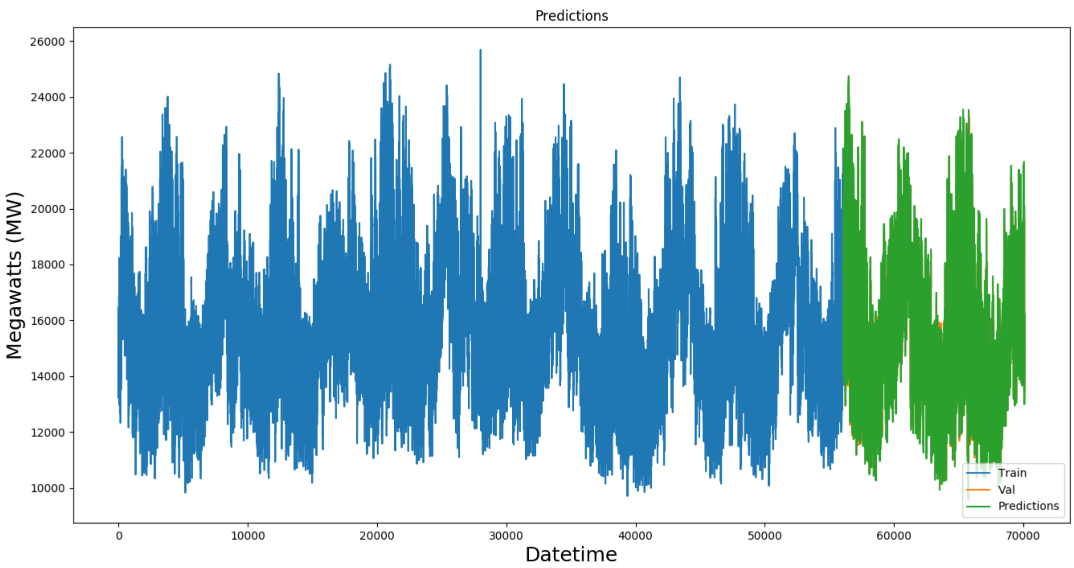


Figure 26.

Here the accuracy of “Between 2005-2008” dataset;

Years	Accuracy for Threshold 1000 Megawatt	Accuracy for Threshold 750 Megawatt	Accuracy for Threshold 500 Megawatt	Accuracy for Threshold 250 Megawatt	Accuracy for Threshold 100 Megawatt
“Between 2005-2012”	0.98	0.95	0.89	0.67	0.33

Table 6

4.1.7. 14-Years Dataset

This is all dataset that I have. It stars 1 January 2005 and ends with 31 December 2018.

The visualization of dataset;

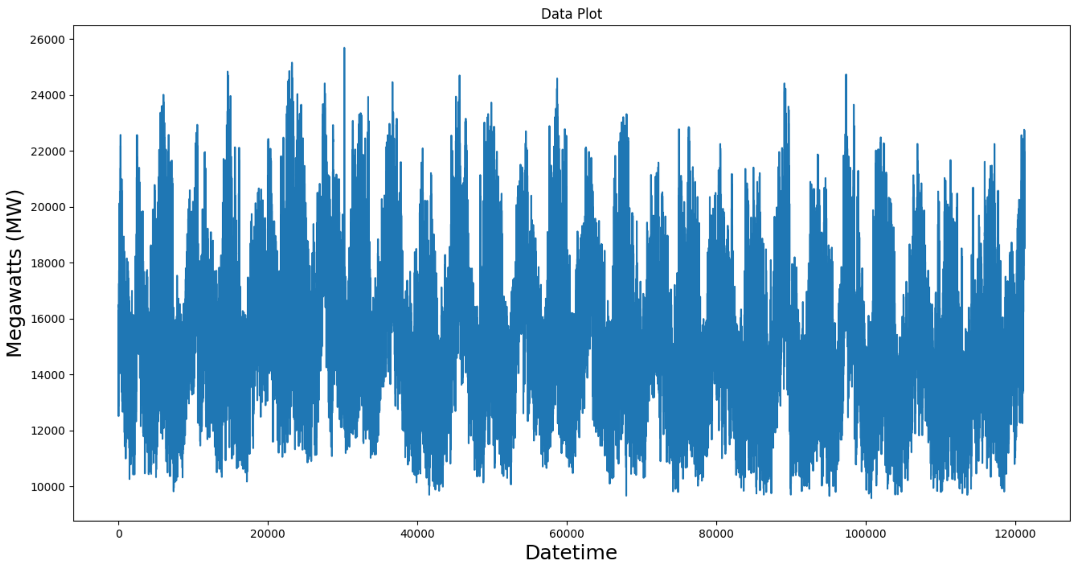


Figure 27.

Here the predictions of “Between 2005-2018” dataset;

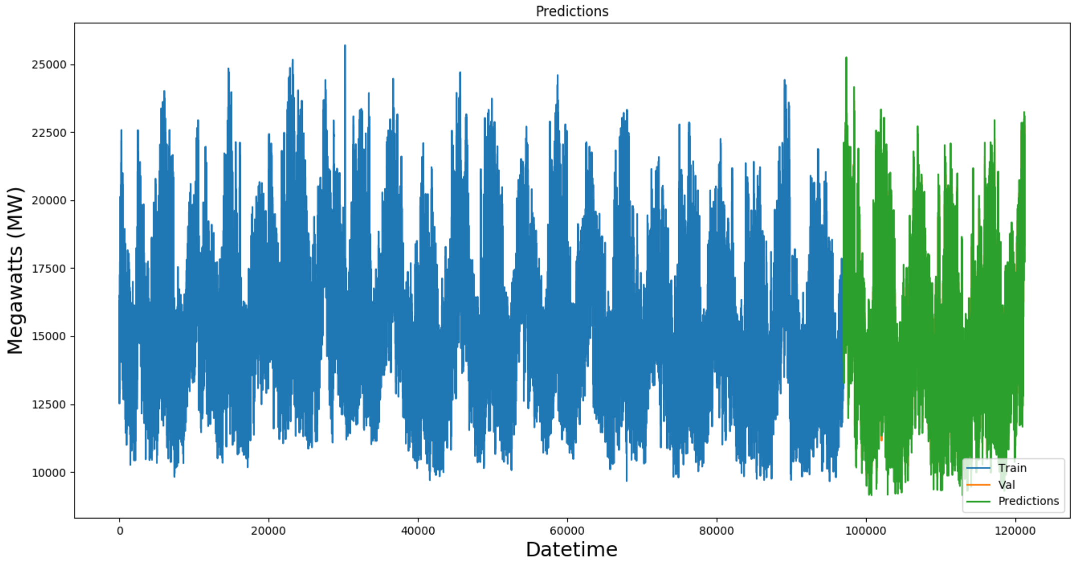


Figure 28.

Here the accuracy table

Years	Accuracy for Threshold 1000 Megawatt	Accuracy for Threshold 750 Megawatt	Accuracy for Threshold 500 Megawatt	Accuracy for Threshold 250 Megawatt	Accuracy for Threshold 100 Megawatt
“Between 2005-2018”	0.98	0.96	0.89	0.60	0.26

Table 7.

4.2. Support Vector Machine to Electric Load Forecasting

The learning model was build on python3. It uses sklearn’s svm model.

The used libraries are:

- math (to basic mathematical operations)
- numpy (to matrix operations)
- pandas (to read data file and save results)
- sklearn (to data scale and build SVM model)
- matplotlib (to visualization data and results).

Here is the model that I used ;

```
#build the SVM Model
model = svm.SVR(kernel='linear')
model.fit(x_train, y_train) #train

#Get the models predicted price values
predictions = model.predict(x_test)
predictions = np.reshape(predictions, (x_test.shape[0], 1 ))
predictions = scaler.inverse_transform(predictions)
```

Figure 29. The SVM Model

I use the same train-test data ratio in all trails. Train data ratio is 80% and test data ratio is 20%.

4.2.1. Dataset

It is the same dataset that I used section 4.1. Here is the link,

https://www.kaggle.com/robikscube/hourly-energy-consumption#AEP_hourly.csv

I split data to 2 years, 3 years, 4 years, 8 years and 14 years. By the way, I can observe the difference between dataset and results.

4.2.2. 2-Years Dataset

This section contains 3 dataset with consist of 2 years data. They are “2005-2006”, “2006-2007” and “2007-2008”. For example, “2005-2006” starts 1 January 2005 and ends with 31 December 2006.

Here visualization of “2005-2006” dataset;

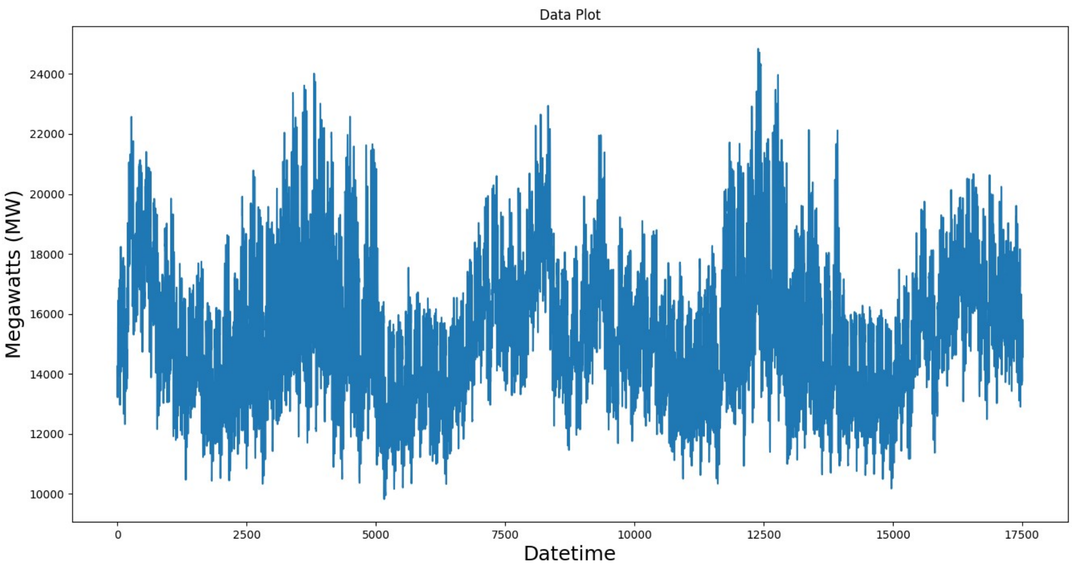


Figure 30. 2005-2006 data visualization

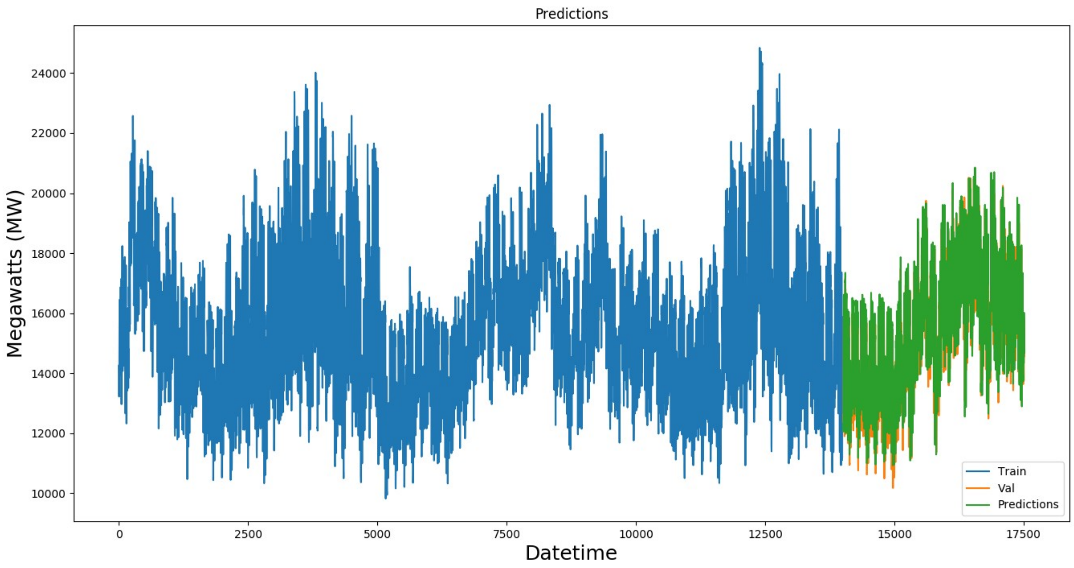


Figure 31. 2005-2006 predictions with SVM

Years	Accuracy for Threshold 1000 Megawatt	Accuracy for Threshold 750 Megawatt	Accuracy for Threshold 500 Megawatt	Accuracy for Threshold 250 Megawatt	Accuracy for Threshold 100 Megawatt
“2005-2006”	0.95	0.84	0.64	0.35	0.14
“2006-2007”	0.92	0.82	0.61	0.33	0.13
“2007-2008”	0.93	0.83	0.63	0.34	0.14

Table 8.

4.2.3. 3-Years Dataset

This section contains 2 dataset with consist of 3 years data. They are “Between 2005-2007” and “Between 2006-2008”. For example, “Between 2005-2007” starts 1 January 2005 and ends with 31 December 2007.

The visualization of “Between 2005-2007” dataset;

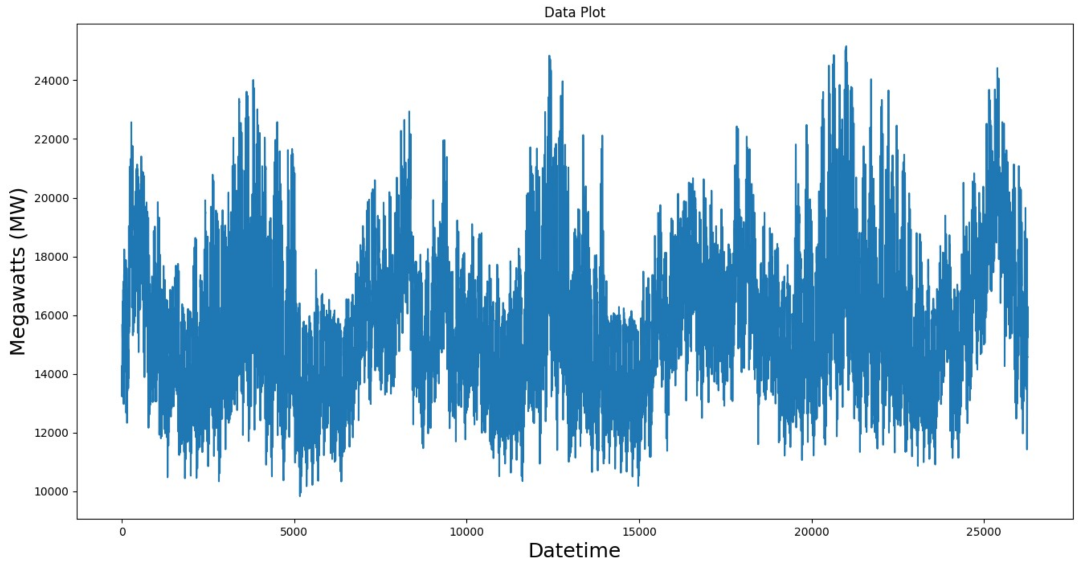


Figure 32. 2005-2007 data visualization

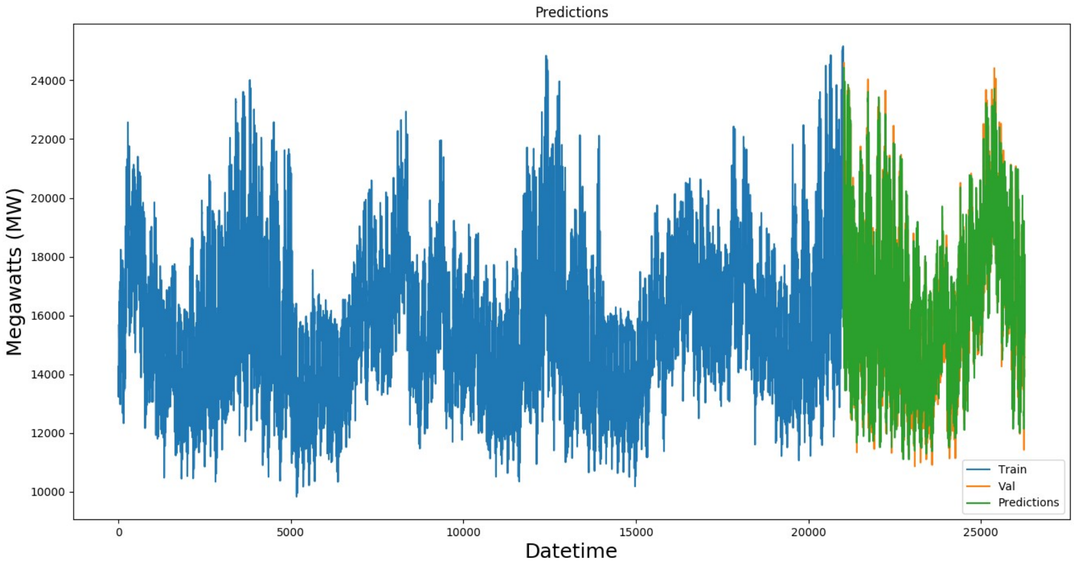


Figure 33. 2005-2007 predictions with SVM

Years	Accuracy for Threshold 1000 Megawatt	Accuracy for Threshold 750 Megawatt	Accuracy for Threshold 500 Megawatt	Accuracy for Threshold 250 Megawatt	Accuracy for Threshold 100 Megawatt
“2005-2007”	0.92	0.82	0.64	0.36	0.15
“2006-2008”	0.93	0.84	0.65	0.36	0.15

Table 9.

4.2.4. 4-Years Dataset

I have a single dataset here. It starts 1 January 2005 and ends with 31 December 2008.

Here the visualization of “Between 2005-2008” dataset;

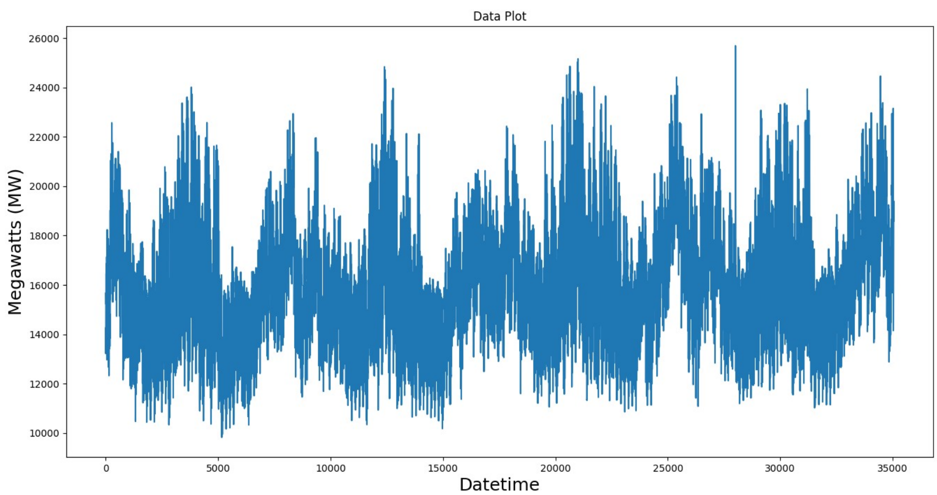


Figure 34. 2005-2008 data visualization

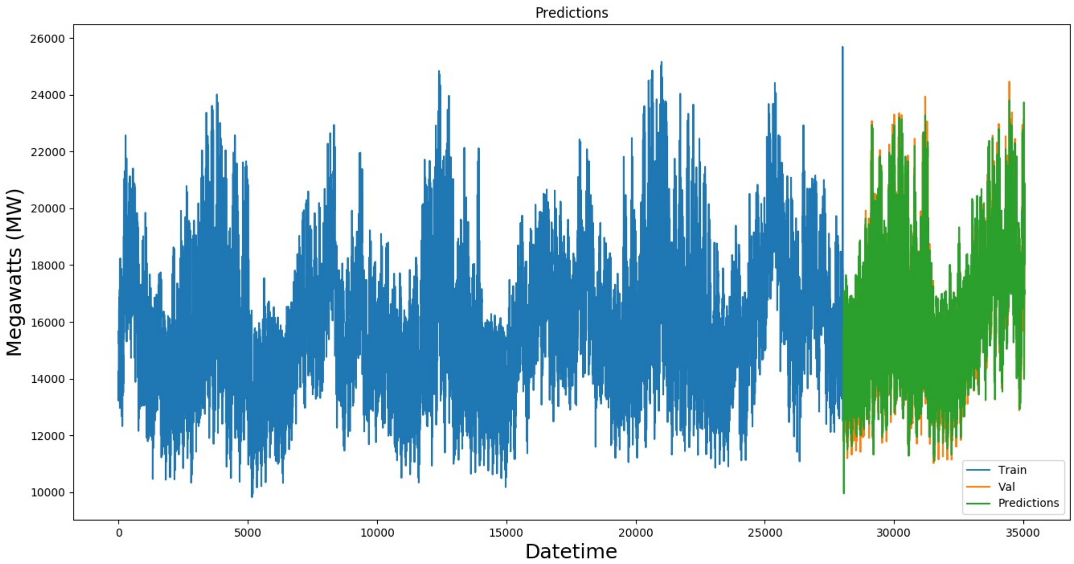


Figure 35. 2005-2008 predictions with SVM

Years	Accuracy for Threshold 1000 Megawatt	Accuracy for Threshold 750 Megawatt	Accuracy for Threshold 500 Megawatt	Accuracy for Threshold 250 Megawatt	Accuracy for Threshold 100 Megawatt
"2005-2008"	0.94	0.84	0.65	0.36	0.15

Table 10.

4.2.5. 8-Years Dataset

This dataset starts 1 January 2005 and ends with 31 December 2012. It contains 8 years data.

The visualization of "Between 2005-2012" dataset;

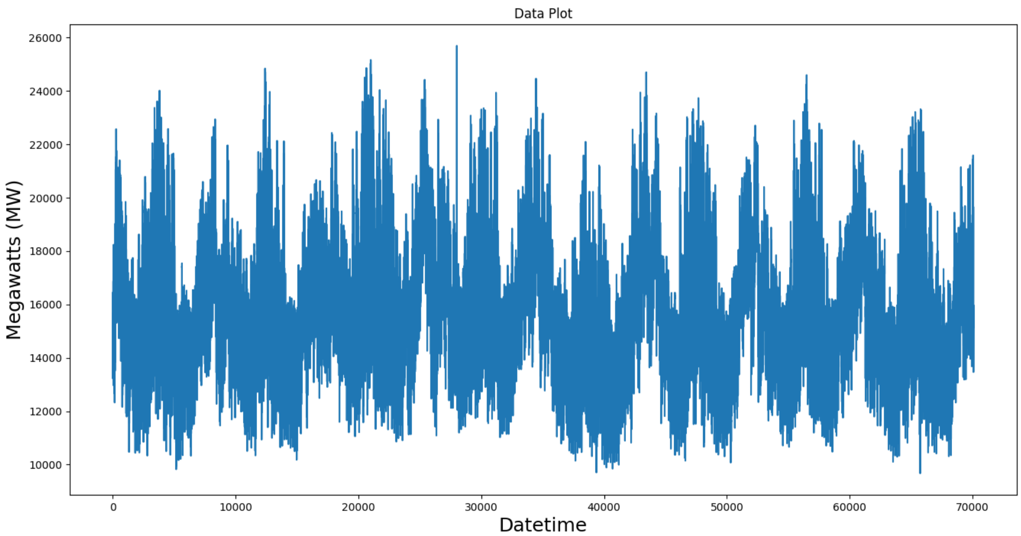


Figure 37. 2005-2012 data visualization

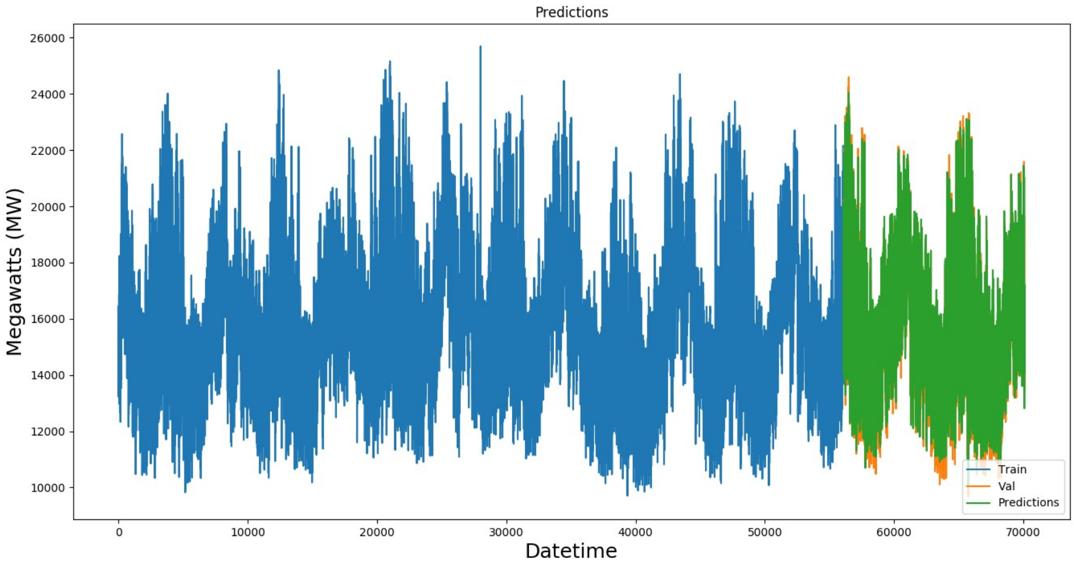


Figure 38. 2005-2012 predictions with SVM

Years	Accuracy for Threshold 1000 Megawatt	Accuracy for Threshold 750 Megawatt	Accuracy for Threshold 500 Megawatt	Accuracy for Threshold 250 Megawatt	Accuracy for Threshold 100 Megawatt
“2005-2012”	0.93	0.83	0.65	0.36	0.15

Table 11.

4.2.6. 14-Years Dataset

This is all dataset that I have. It stars 1 January 2005 and ends with 31 December 2018.

The visualization of dataset;

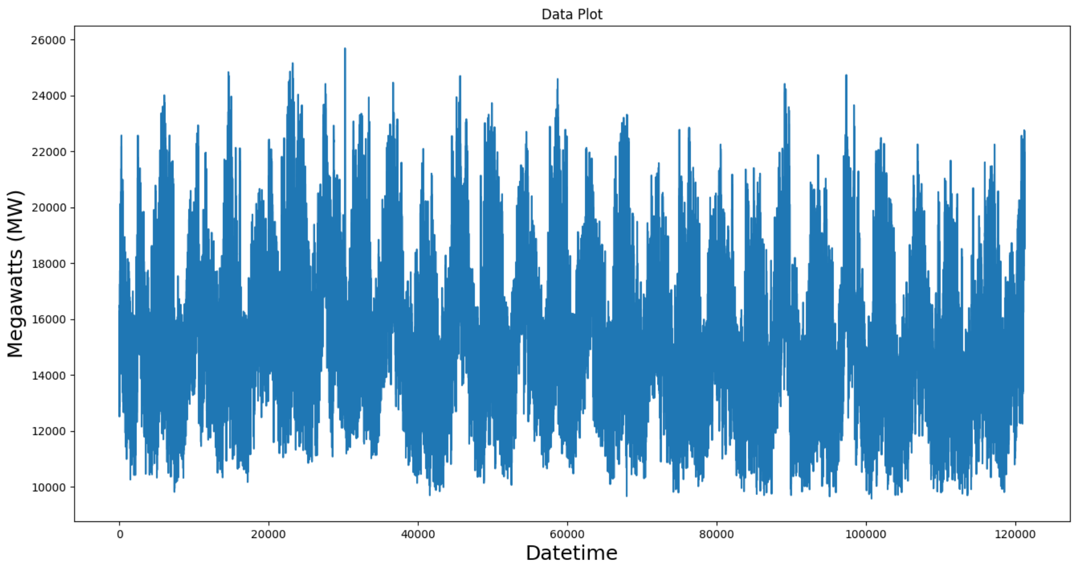


Figure 39. 2005-2018 data visualization

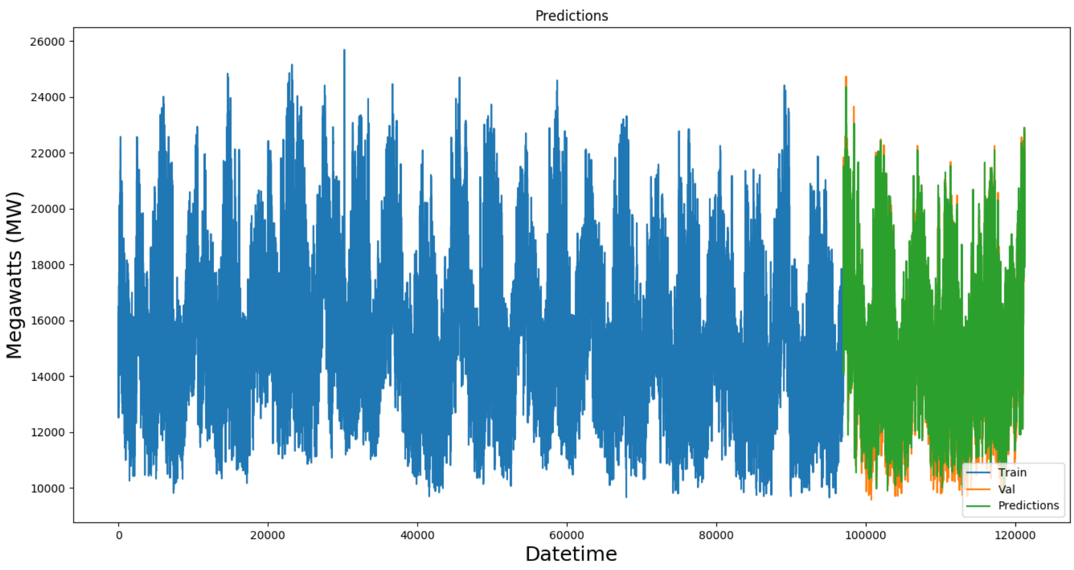


Figure 40. 2005-2018 predictions with SVM

Years	Accuracy for Threshold 1000 Megawatt	Accuracy for Threshold 750 Megawatt	Accuracy for Threshold 500 Megawatt	Accuracy for Threshold 250 Megawatt	Accuracy for Threshold 100 Megawatt
“2005-2018”	0.93	0.81	0.58	0.30	0.12

Table 12.

4.3. LSTM and SVM Comparison

Years	Model	Accuracy for Threshold 1000 Megawatt	Accuracy for Threshold 750 Megawatt	Accuracy for Threshold 500 Megawatt	Accuracy for Threshold 250 Megawatt	Accuracy for Threshold 100 Megawatt
“2005-2006”	SVM	0.95	0.84	0.64	0.35	0.14
	LSTM	0.97	0.95	0.87	0.60	0.26
“2005-2007”	SVM	0.92	0.82	0.64	0.36	0.15
	LSTM	0.96	0.91	0.74	0.39	0.15
“2005-2008”	SVM	0.94	0.84	0.65	0.36	0.15
	LSTM	0.97	0.94	0.85	0.58	0.26
“2005-2012”	SVM	0.93	0.83	0.65	0.36	0.15
	LSTM	0.98	0.95	0.89	0.67	0.33
“2005-2018”	SVM	0.93	0.81	0.58	0.30	0.12
	LSTM	0.98	0.96	0.89	0.60	0.26

Table 13.

LSTM learning model supplies more efficient learning for this dataset. Instead of higher accuracy it takes 45 minutes to “2005-2018” dataset. SVM model takes just a few seconds. If we will work with larger datasets then time will be a problem. So I can advice that LSTM model can be use in small datasets to take higher accuracy.

5. Related Works

I checked GitHub and searched related works. Electric load forecasting prone to time series and RNN approaches. LSTM model is a submodel of RNN. I check some projects in GitHub and I can say that the one of the most common used method is LSTM. The one of the project that I checked from GitHub is here,

https://github.com/Yifeng-He/Electric-Power-Hourly-Load-Forecasting-using-Recurrent-Neural-Networks/blob/master/load_forecasting.py

It used 3 layer LSTM model. Here the LSTM model,

```
# build the model
model = Sequential()
# layer 1: LSTM
model.add(LSTM( input_dim=1, output_dim=50, return_sequences=True))
model.add(Dropout(0.2))
# layer 2: LSTM
model.add(LSTM(output_dim=100, return_sequences=False))
model.add(Dropout(0.2))
# layer 3: dense
# linear activation: a(x) = x
model.add(Dense(output_dim=1, activation='linear'))
# compile the model
model.compile(loss="mse", optimizer="rmsprop")

# train the model
model.fit(X_train, y_train, batch_size=512, nb_epoch=50, validation_split=0.05, verbose=1)
```

Figure 41.

The output of model very close the normal values. Here the outputs,

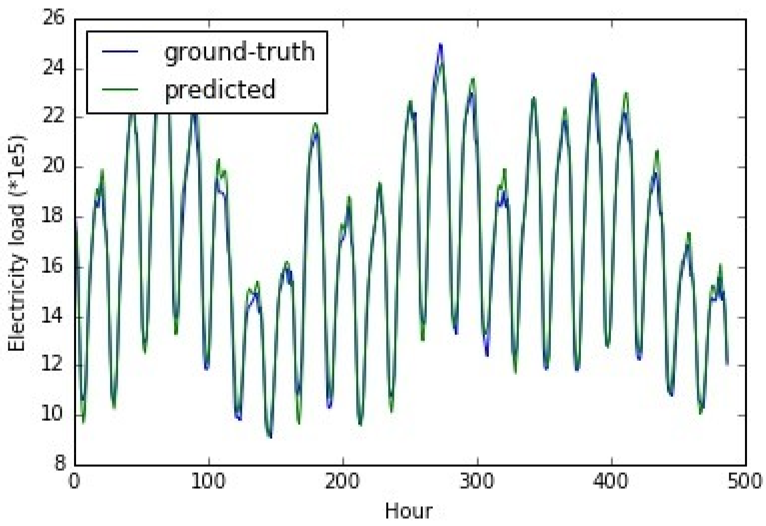


Figure 42.

6. Discussions

The correct algorithm and methods brings correct results and high accuracy. We also need clean and accurate data. Some data have date deviations and cause errors during training. The dataset should be chosen correctly and used correctly. Two different algorithms were tried. They are SVM and LSTM. LSTM has high accuracy but it takes more time in training. Although SVM trains in a few seconds, it offers lower accuracy.

7. Video Presentation Link

https://drive.google.com/file/d/1Xs2FB0LFsIgGhYSs0LMUjxJR2OgyIRt8/view?usp=sharing

References

– **[11]** Alfares, H. & Nazeeruddin, M. (2002). Electric Load Forecasting: Literature Survey and Classification of Methods. International Journal of Systems Science.

– **[10]** Basak, D., Pal, B. S. & Patranabis, D. C. Support Vector Regression. Neural Information Processing Letters and Reviews.

– **[5]** Feinberg, E. & Genethliou, Dora. (2005). Load Forecasting. Applied Mathematics for Power Systems.

– **[1]** Feinberg, E. A., & Genethliou, D. (2005). Applied Mathematics for Restructured Electric Power Systems. New York: Springer Science & Business Media.

– **[9]** Ghods, L and Kalantar, M. (2011). Different Methods of Long-Term Electric Load Demand Forecasting; A Comprehensive Review. Iranian Journal of Electrical & Electronic Engineering.

– [2] Hyndman R. J & Athanasopoulos, G. (2013). Forecasting: principles and practice. Otext Publisher.

– **[7]** Islam, B. (2011). Comparison of Conventional and Modern Load Forecasting Techniques Based on Artificial Intelligence and Expert Systems. IJCSI International Journal of Computer Science Issues.

– **[8]** Kalogirou, S. A. (2000). Applications of artificial neural-networks for energy systems. Applied Energy.

– **[3]** Kandil, M. S., El-Debeiky, S. M., & Hasanien, N. E. (2000). Overview and Comparison of Long-Term Forecasting Techniques for a Fast Developing Utility, Part I. New York: Elsevier Publishers.

– **[6]** Petridis, V., Kahagias, A., Bakirtzis, A. & Kiartzis, S. (2003). Short Term Load Forecasting Using Predictive Modular Neutral Networks. Aristotle University of Thessaloniki, Greece.

– **[4]** Powell, S. G. & Baker, K. R. (2010). Management Science: The Art of Modeling with Spreadsheets. New York: John Wiley & Sons.

- **[12]** Sheng Chen, Hongxiang He. "Stock Prediction Using Convolutional Neural Network", IOP Conference Series: Materials Science and Engineering, 2018