Prediction of Electricity Demand in Ontario Province

Tolulope Olugbenga
University of New Brunswick
3 Bailey Drive
t.boy@unb.ca

Md. Mahbubur Rahman
University of New Brunswick
3 Bailey Drive
mrahma15@unb.ca

ABSTRACT

Proper electricity demand forecasting is an important area for power supply companies because it promotes better scheduling and efficiency management. In terms of power supply and demand, For the stable supply of electricity, the reserve power must be prepared [1]. For this reason, power companies have to find better models that can forecast and help with the planning of energy usage. It is essential to forecast power demand because electricity is difficult to store.

In this paper, we got the previous hourly power demand from 2003 until early 2019 from an Independent Electricity System Operator in Ontario, and the hourly weather conditions from a weather station in Ontario's most populated city Toronto [2], [3]. We classified the dataset to 85% for training our model and 15% for testing. The results were compared with the actual demand. We compared the results on charts using the hourly, daily and monthly averages, we got the Correlation Coefficient, Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), Mean Squared Error (MSE), and Root Mean Squared Error (RMSE) for each case. We ran the dataset on three algorithms: Decision Tree, GaussianNB, and Random Forest. Based on the MAPE, the random forest algorithm performed better with a MAPE ranging from 1.77 to 2.23 for both Market and Ontario Demand.

KEYWORDS: power demand forecasting, decision tree, gaussiannb, random forest, artificial intelligence

1. INTRODUCTION

Electricity is one of the driving forces of economic development and is essential to our daily life and wellbeing. Power demand forecasting is a difficult task due to the number of the different random variables that needs to be taken into consideration in order to predict human behavior. People often use electricity at any time that suits their lifestyle, but for the most part we all happen to use electricity at the same time. Most people share a similar lifestyle pattern, from when we wake up, to having a shower, making some breakfast, leaving for work, coming back at night, going to bed, doing our laundry on weekends and so on.

In this paper, we got the previous hourly power demand from 2003 until early 2019 from an Independent Electricity System Operator in Ontario, and the hourly weather conditions from a weather station in Ontario's most populated city Toronto [2], [3]. Weather plays a huge role in electricity usage because, in warmer climate people would use more electricity for air conditioning while during the colder times people would use more electricity for heating. We were able to predict the power demand in the province of Ontario with a lot of uncertainties and missing data values. We implemented and compared three algorithms: Decision Tree, GaussianNB, and Random Forest. Based on the MAPE, the random forest algorithm performed better with a MAPE ranging from 1.77 to 2.23 for both Market and Ontario Demand.

2. RELATED WORKS

Machine learning is a subfield of Artificial Intelligence [4]. Machine learning has been used to solve different problems in our society today, it has a lot of daily use utilizations like in google searches, text editors, iPhone Siri, etc. These machine learning algorithms are often used for classification problems and pattern detection, they work best with a large dataset. There are some reports that used machine learning to predict power usage and other similar works are highlighted here.

Williams and Gomez performed a study to predict energy consumption for the next month, they implemented three algorithms: Linear Regression, Regression Trees and Multi-variate Adaptive Regression Spline (MARS). The authors used data from homes in Bexar County, Texas containing four years of monthly consumption. After evaluation, the authors achieved a residual mean squared error (RMSE) of 99.803 ± 3.057, 100.435 ± 3.441 and 94.286 ± 3.238 kBtu/day for the linear regression, regression trees and MARS, respectively. The RMSE is calculated by comparing the original average daily energy (kBtu/day) with the predicted average daily energy (kBtu/day). The authors also did an aggregation on daily consumption predictions, to predict the monthly consumption for groups of homes; in this case, the RMSE turned out to be 25.743±1.097, 18.277±1.156 and 19.831±1.187 kBtu/day for the linear regression, regression trees and MARS, respectively [5][6].

Rodrigues and his team wanted to forecast energy consumption of households using an Artificial Neural Network (ANN). The authors collected data from Lisbon, Portugal containing data of 93 households, the data contained hourly energy consumption during 6 to 8 weeks resulting in a total of 93,744 records. Two methods were proposed by the authors: one is used to forecast daily consumption and the other is used to forecast hourly consumption, they both used an ANN with one hidden layer and 20 neurons. After evaluation, the Mean Absolute Percent Error (MAPE) of the average consumption was 4.2% and the maximum consumption was 18.1%, the results were reported in regards to the daily energy consumption [6][7].

3. PROBLEM STATEMENT

Electricity consumption is growing at very rapid rate worldwide. This growth presents the need for better planning of energy usage, this includes planning of future electricity demand by electricity distribution companies. Demand response management (DRM) has been one of the major features in smart grid that helps with balancing the electricity demand with the supply available to outages. DRM controls the electricity consumption at the customer’s end, based on the customer’s preset parameters. To have a stable supply of electricity the reserve power has to be prepared. This means electricity companies need to be able to predict the future electricity demand. Power consumption prediction is an important part of the planning and operation of an electricity distribution company. Accurate demand forecasting is important for policy makers to formulate electricity supply policies.

4. OUR APPROACH

A typical workflow for demand forecasting is: a) getting the dataset from disparate sources, such as databases or spreadsheets; b) cleaning the data, removing outliers, noise and combining datasets; c) developing an accurate predictive model based on the aggregated data using forecasting techniques; d) deploying the model as an application in a production environment [8]. Our approach was we got the dataset, filtered and combined them, separated some for training and testing of our models, then analyze the results.

4.1 Datasets

This case study consists of two datasets: the past power usage for Ontario and the weather data from one of their major cities Toronto. The power data was gotten from an independent electricity system operator (ieso) [2] and the weather data was gotten from the government of Canada’s website [3]. The power data consists of hourly data from the 1st of May, 2002 and contained 3 headers namely: Date, Hour, Market Demand and Ontario Demand.

Market Demand represents the total energy that was supplied from the IESOAdministered Market; it is calculated by adding all output from generators registered in the market and all scheduled imports to the province; it is also equal to the sum of all load supplied from the market, exports from the province and all lines losses incurred on the IESO-controlled grid [9]. Ontario demand represents the total energy that was supplied from the IESO-Administered Market for the purpose of supplying load within Ontario; It is also equal to the sum of all loads within Ontario which is supplied from the Market, plus all line losses incurred on the IESO-controlled grid [9]. The Market Demand and Ontario Demand are both in Megawatts (MW).

The weather data consists of various weather information, the most important for this case study are Temperature (℃), Dew Point and Relative Humidity. These were the categories that had the most complete information.

4.2 Combining datasets

There was no weather data until the 4th of June, 2002 at 16:00, for this reason we considered data starting from 2003 still early 2019. All missing hourly weather data were replaced with the data from the cell in the previous hour until we had a fully complete dataset; we felt the temperature wouldn't change as much within two hours. We combined the columns of the power and weather data together. The final dataset was from the 1st of January, 2003 still the 2nd of April, 2019; and it contains 10 headers: - Date/Time, Year, Month, Day, Hours, Market Demand, Ontario Demand, Temperature, Dew Point and Relative Humidity, as seen in the figure below.

Date/Time	Year	Month	Day	Hours	Market Demand	Ontario Demand	Temp (Â°C)	Dew Point Temp (Â°C)	Rel Hum (%)
01-01-03 0:00	2003	1	1	1	15558	14745	1.4	-3.9	68
01-01-03 1:00	2003	1	1	2	15131	14280	1.5	-4.4	65
01-01-03 2:00	2003	1	1	3	15075	13821	1.5	-4.2	66
01-01-03 3:00	2003	1	1	4	14308	13239	1	-4.7	66
01-01-03 4:00	2003	1	1	5	14059	13236	0.7	-5.2	65
01-01-03 5:00	2003	1	1	6	14375	13504	0.7	-5.2	65
01-01-03 6:00	2003	1	1	7	14881	13814	0.3	-5.4	66
01-01-03 7:00	2003	1	1	8	14305	13544	0	-5.5	66
01-01-03 8:00	2003	1	1	9	14404	13837	0	-5.1	68
01-01-03 9:00	2003	1	1	10	15184	14623	0	-5.1	68
01-01-03 10:00	2003	1	1	11	15880	15319	-0.2	-6.1	64
01-01-03 11:00	2003	1	1	12	16577	15886	-0.2	-6.1	64
01-01-03 12:00	2003	1	1	13	16888	16222	-0.3	-6.2	64
01-01-03 13:00	2003	1	1	14	17250	16161	-0.2	-6.3	63
01-01-03 14:00	2003	1	1	15	16735	16094	-0.1	-7	60
01-01-03 15:00	2003	1	1	16	16965	16374	-0.4	-7.7	58
01-01-03 16:00	2003	1	1	17	17988	17097	-0.5	-7.8	58
01-01-03 17:00	2003	1	1	18	19665	18594	-0.9	-6.9	64
01-01-03 18:00	2003	1	1	19	19567	18496	-0.9	-6.9	64
01-01-03 19:00	2003	1	1	20	19290	18333	-2.4	-9.8	57
01-01-03 20:00	2003	1	1	21	18886	18100	-3.4	-11.9	52
01-01-03 21:00	2003	1	1	22	18781	17665	-4.3	-13	51
01-01-03 22:00	2003	1	1	23	18083	16921	-5.1	-13.3	53
01-01-03 23:00	2003	1	1	24	17524	15848	-6	-13.2	57

Figure 1: The figure shows the top part of the final dataset used.

5. IMPLEMENTATION

5.1 Design

We have implemented the algorithms with python and the ide we used is Sypder3. We did not develop any system or software. We implemented those algorithms to evaluate the result and compare the results to come to a conclusion that which algorithm works better and provides the most accurate predictions. There are several built-in libraries that we used to implement the algorithms.

5.2 Description of the code/script

We have implemented the algorithms using python and built in libraries of python. The first step was to explore and manipulate the data and for that we have used Pandas. The most important part of Pandas library is the Data Frame which holds the data as a table and Pandas has powerful methods to work with this type of data. NumPy is another important library we used for our implementation. NumPy provides support for large, multi-dimensional arrays and metrics along with a large collection of high-level mathematical functions to operate on these arrays. After exploring data, we selected our prediction target with the features we need to predict that target. Then we split our dataset for training and testing purpose. After the split, we created a model and fit that data into our model. Finally, we predicted and evaluated the predictions with the original dataset.

```python

@author: Maruf
"""

import pandas as pd
import numpy as np
from sklearn.model_selection import train_test_split
#from sklearn.tree import DecisionTreeRegressor
#from sklearn.ensemble import RandomForestRegressor
from sklearn.naive_bayes import GaussianNB
from sklearn import metrics



# Path of the file to read
path = 'C:/Users/Maruf/Documents/Python Scripts/datasets/2003-2019.csv'

data = pd.read_csv(path)
# Create target object and call it y
y = data.OntarioDemand
# Create X
features = ['Year', 'Month', 'MarketDemand','Temp', 'DewPointTemp', 'RelHum']
X = data[features]

train_X, val_X, train_y, val_y = train_test_split(X, y, test_size=.15, random_state=0)

model = GaussianNB()

model.fit(train_X, train_y)

predictions = model.predict(val_X)
coef = metrics.matthews_corrcoef( predictions, val_y)
print("COEF:",coef)
mae=metrics.mean_absolute_error(predictions, val_y)
print("MAE:", mae)
val_mse = metrics.mean_squared_error( predictions, val_y)
print("MSE:",val_mse)
val_rmse = np.sqrt(metrics.mean_squared_error(predictions, val_y))
print("RMSE:",val_rmse)
y_true, y_pred = np.array(val_y), np.array(predictions)
mape = np.mean(np.abs((y_true - y_pred) / y_true)) * 100
print("MAPE:",mape)
```

Figure 2: Sample Code of Implementation

6. EVALUATION

6.1 Experimental setup

We have used Spyder3 as IDE and the specification of our device was core i7-4700HQ quad core processor, 8GB DDR3 ram, and 1TB SATA drive. The Spyder3 IDE has its own kernel to display the results. We divided the dataset into two part, one is for training purpose and other one for testing. We used 85% of the data for training the algorithms and 15% data to test. Moreover, we tried to use as many features as possible so that the prediction result becomes more accurate. We represented our predictive data and validation data using Excel sheet. We displayed our validation data in hourly, daily and monthly basis. We have represented the comparison results in graphical view for a better understanding.

6.2 Experimental results

We have tried to compare the results we got by implementing all the algorithms in tabular form as well as in graphical view. Figure 3 displays the daily predictions of Market Demand and Ontario Demand using the Decision Tree Regressor of the month February 2019.

Date	Days	Actu. Market Demand	Pred. Market Demand	Actu. Ontario Demand	Pred. Ontario Demand	Temp (°C)	Dew Point Temp (°C)
01-Feb-19	1	20837.125	17566	18651.29167	16251.125	-13.34583333	-20.125
02-Feb-19	2	19669.70833	18308.20833	16987.95833	16691.16667	-5.116666667	-8.904166667
03-Feb-19	3	18697.41667	18660.04167	15676.08333	16668.45833	4.179166667	1.666666667
04-Feb-19	4	18942.45833	18583.375	15975.29167	16078.20833	8.283333333	3.904166667
05-Feb-19	5	18697.375	19105.08333	16483.54167	16077.125	1.458333333	-2.854166667
06-Feb-19	6	18920.45833	18211.70833	17649.45833	15967.41667	-2.5625	-4.5625
07-Feb-19	7	18492.58333	18802.75	17059.95833	16715.83333	-0.741666667	-2.195833333
08-Feb-19	8	19180.16667	17120.375	16902.04167	17287.625	-2.170833333	-8.020833333
09-Feb-19	9	18897.54167	17450.58333	16606.20833	16339.375	-6.045833333	-12.79166667
10-Feb-19	10	18146.375	18021.625	16371.54167	15938.41667	-5.220833333	-10.8875
11-Feb-19	11	19456.91667	18599.79167	17205.625	16448.33333	-3.929166667	-8.895833333
12-Feb-19	12	19878.41667	18318.83333	17780.58333	16479.20833	-3.095833333	-5.433333333
13-Feb-19	13	19834	17916.33333	17185.91667	17419.54167	-1.9375	-5.341666667
14-Feb-19	14	19080.25	17812.45833	16921.625	16073.58333	-1.079166667	-5.141666667
15-Feb-19	15	18827.54167	17604	16439.20833	16290.91667	0.866666667	-3.891666667
16-Feb-19	16	17481.75	18379.875	15777.20833	16207.625	-4.141666667	-10.06666667
17-Feb-19	17	17573.70833	17581.33333	15858.41667	16403.875	-7.1	-12.4625
18-Feb-19	18	18304.41667	17893.08333	16141	15926.79167	-7.629166667	-13.12083333
19-Feb-19	19	19265.875	18012.83333	17488	15730.625	-7.229166667	-13.29166667
20-Feb-19	20	19622.70833	18496.91667	17481.45833	16764.29167	-2.583333333	-6.875
21-Feb-19	21	18811.54167	18473.125	16606.41667	16534.79167	2.354166667	-3.4625
22-Feb-19	22	18027.58333	17938.08333	16436.16667	15972.5	0.55	-5.683333333
23-Feb-19	23	17524.25	19074.41667	15206	16562.83333	0.629166667	-3.991666667
24-Feb-19	24	17965.04167	18917.45833	15352.75	15833.33333	2.904166667	-0.1875
25-Feb-19	25	19397.625	19528.16667	16848.29167	15793.25	-4.25	-10.12083333
26-Feb-19	26	19497.79167	17759.79167	17480.16667	16285.41667	-7.654166667	-15.30833333
27-Feb-19	27	19996	18320.04167	18301.95833	17712.375	-9.733333333	-13.21666667
28-Feb-19	28	19643.33333	18254.79167	17615.16667	16044.08333	-9.35	-13.93333333

Figure 3: Market Demand and Ontario Demand Prediction for Month February 2019

Figure 4 displays the comparison graph of average daily original Market Demand with predicted Market Demand as well as average daily original Ontario Demand with predicted Ontario Demand of the month of February 2019 of the results displayed above.

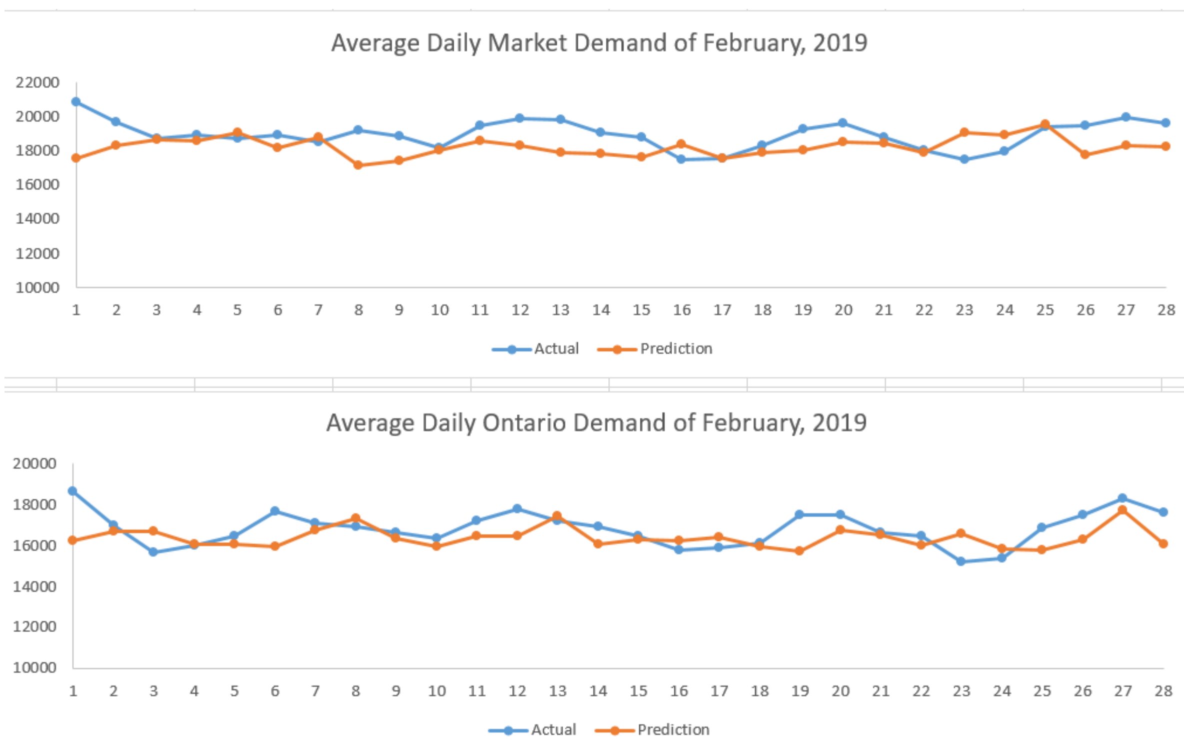


Figure 4: Average Daily Market Demand and Ontario Demand Comparison

Figure 5 displays a comparison graph of average daily temperature and dew point of the month February 2019.

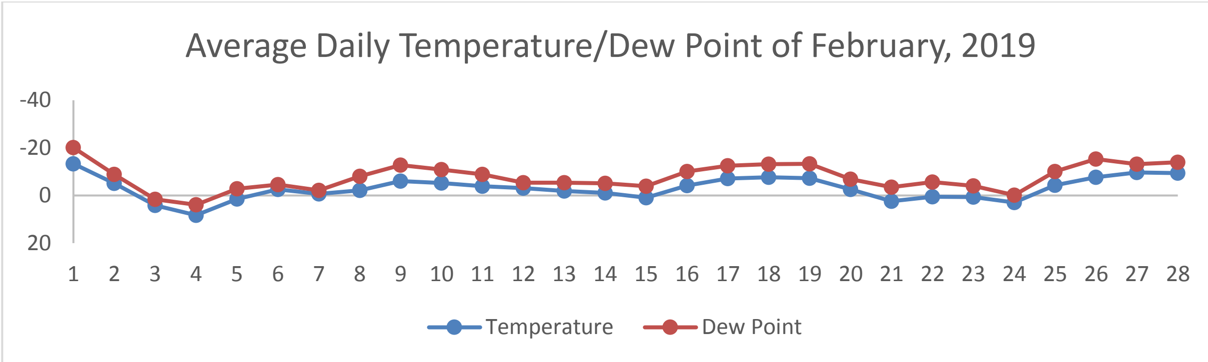


Figure 5: Average Daily Temperature Vs Dew Point

The Figure 6 displays the monthly predictions of Market Demand and Ontario Demand using the GaussianNB (Gaussian Naïve Bayes) of the year 2018.

Year	Month	Actu. Market Demand	Pred. Market Demand	Actu. Ontario Demand	Pred. Ontario Demand	Temp (°C)	Dew Point Temp (°C)
2018	January	19909.54032	18121.97849	17071.00269	16386.94086	-4.259005376	-9.086290323
2018	February	18394.73214	18106.58482	16296.22768	16434.7128	-0.43764881	-5.498363095
2018	March	17503.59677	18173.14919	15285.51075	16388.32258	0.682123656	-6.889919355
2018	April	16552.70556	18197.69722	14692.17778	16422.14722	4.243888889	-3.765138889
2018	May	16300.27016	17884.99462	14007.35081	16099.04704	17.02728495	6.603360215
2018	June	17218.30278	18044.42778	15171.45278	16327.88056	19.63819444	10.50416667
2018	July	18661.03629	18072.63978	17050.16263	16334.57392	23.66666667	14.1483871
2018	August	18741.95161	18167.13038	17054.83199	16401.31317	23.44879032	16.40591398
2018	September	17422.14028	18375.6	15482.375	16628.26667	19.38041667	13.35097222
2018	October	16601.38306	18133.66667	14405.03763	16397.71237	9.800537634	4.273924731
2018	November	17853.20972	18295.65278	15778.22361	16579.63194	2.717222222	-2.442361111
2018	December	18031.05914	18051.63038	15980.99059	16259.7379	1.130376344	-3.129435484

Figure 6: Market Demand and Ontario Demand Prediction for 2018

Figure 7 displays the comparison graph of average monthly original Market Demand with predicted Market Demand as well as average monthly original Ontario Demand with predicted Ontario Demand of the year of 2018 of the results displayed above.

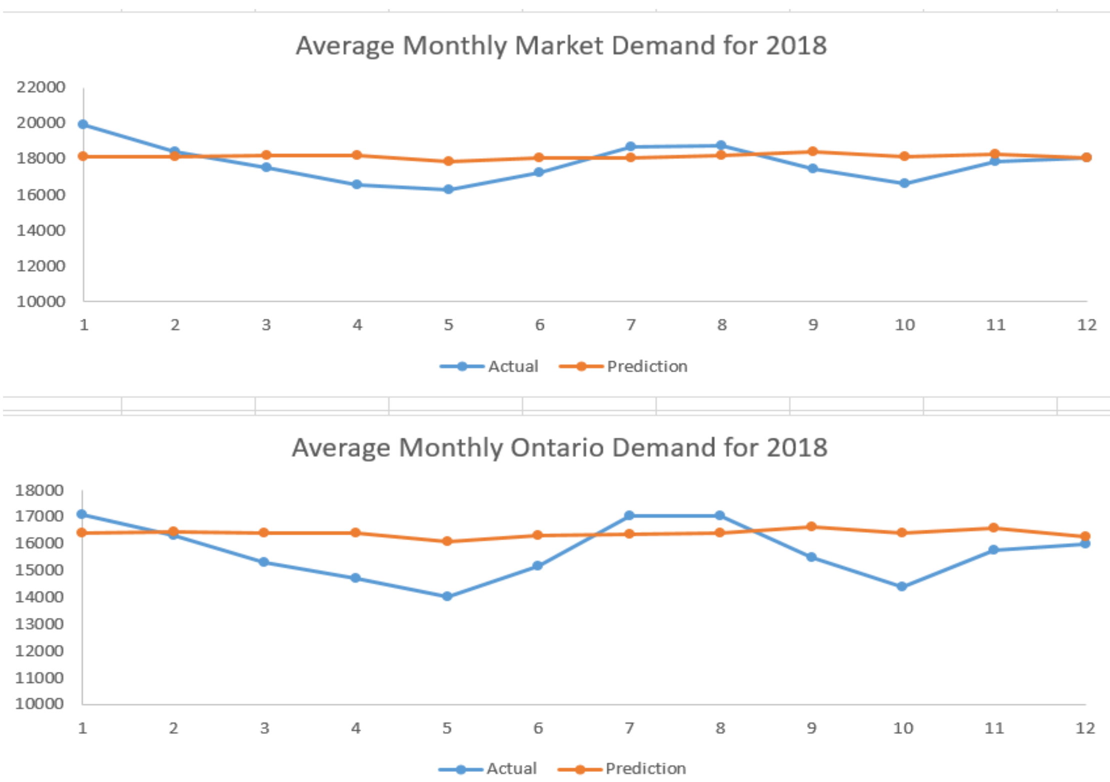


Figure 7: Average Monthly Market Demand and Ontario Demand Comparison

Figure 7 displays a comparison graph of average daily temperature and dew point of the month February 2019.

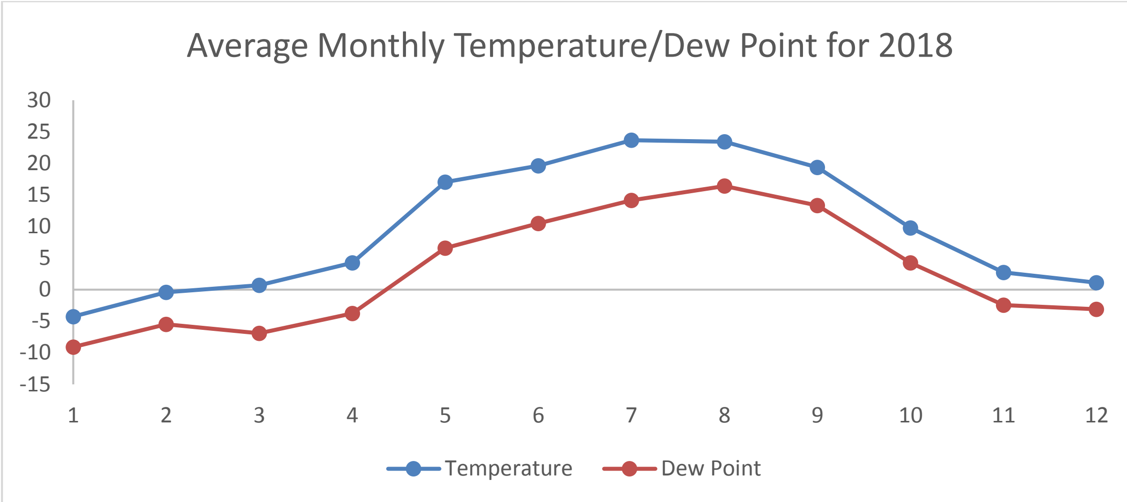


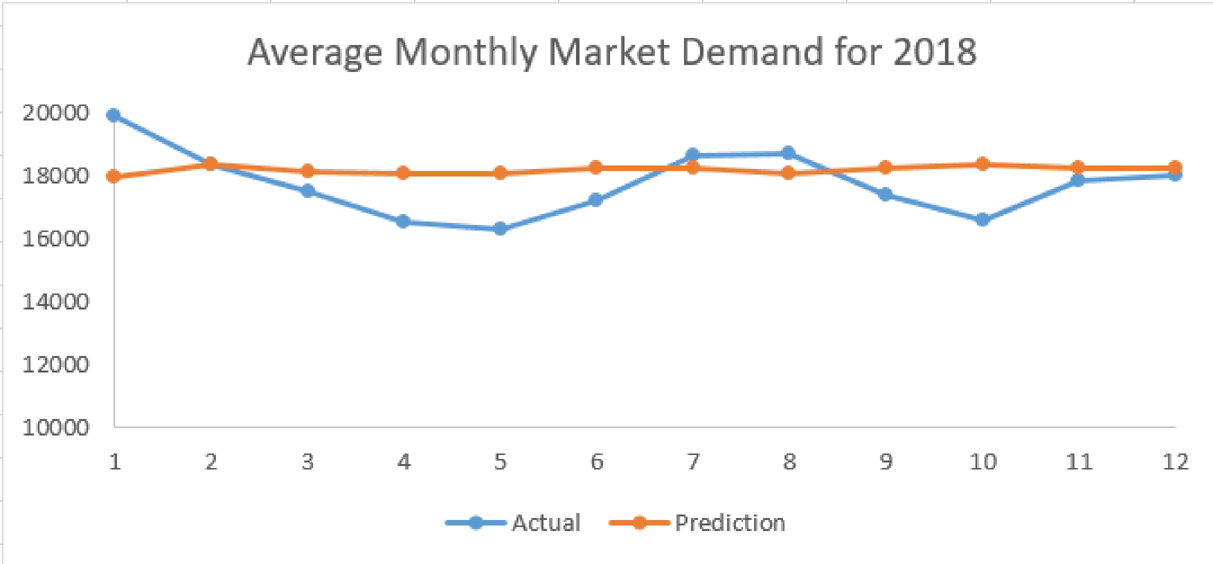
Figure 7: Average Monthly Temperature Vs Dew Point

The Figure 8 displays the monthly predictions of Market Demand and Ontario Demand using the Random Forest Regressor of the year 2018.

Year	Month	Actu. Market Demand	Pred. Market Demand	Actu. Ontario Demand	Pred. Ontario Demand	Temp (°C)	Dew Point Temp (°C)
2018	January	19909.54032	17992.07823	17071.00269	16487.1914	-4.259005376	-9.086290323
2018	February	18394.73214	18358.95655	16296.22768	16464.32961	-0.43764881	-5.498363095
2018	March	17503.59677	18118.00403	15285.51075	16392.40202	0.682123656	-6.889919355
2018	April	16552.70556	18096.17458	14692.17778	16534.76278	4.243888889	-3.765138889
2018	May	16300.27016	18081.0797	14007.35081	16320.49019	17.02728495	6.603360215
2018	June	17218.30278	18249.98694	15171.45278	16466.78458	19.63819444	10.50416667
2018	July	18661.03629	18260.48468	17050.16263	16351.38992	23.66666667	14.1483871
2018	August	18741.95161	18116.4297	17054.83199	16289.14032	23.44879032	16.40591398
2018	September	17422.14028	18246.92806	15482.375	16359.01778	19.38041667	13.35097222
2018	October	16601.38306	18392.81774	14405.03763	16552.95108	9.800537634	4.273924731
2018	November	17853.20972	18269.17875	15778.22361	16520.93319	2.717222222	-2.442361111
2018	December	18031.05914	18281.79113	15980.99059	16342.72984	1.130376344	-3.129435484

Figure 8: Market Demand and Ontario Demand Prediction for 2018

Figure 9 displays the comparison graph of average monthly original Market Demand with predicted Market Demand as well as average monthly original Ontario Demand with predicted Ontario Demand of the year of 2018 of the results displayed above.



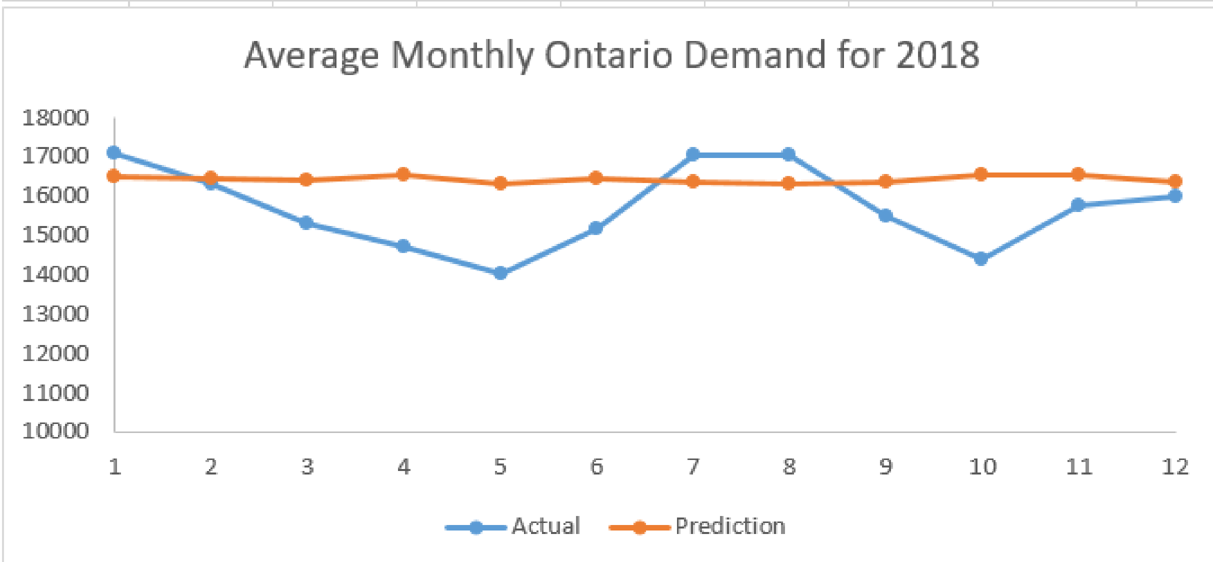


Figure 9: Average Monthly Market Demand and Ontario Demand Comparison

Figure 10 displays the comparison results of all the validation matrices (Correlation Coefficient, Mean Absolute Error, Mean Squared Error, Root Mean Squared Error, and Mean Absolute Percentage Error) on hourly, daily, and monthly basis.

Market Demand (Hourly)					
Algorithm	Correlation Coefficient	Mean Absolute Error	Mean Squared Error	Root Mean Squared Error	Mean Absolute Percentage Error
Decision Tree	0.970142035	0.211932616	91.69457183	9.575728266	2.282780406
GaussianNB	0.958115212	0.261675246	73.5620964	8.576834871	4.789975943
Random Forest	0.969706662	0.189953205	79.17490735	8.898028284	1.781301458

Ontario Demand (Hourly)					
Algorithm	Correlation Coefficient	Mean Absolute Error	Mean Squared Error	Root Mean Squared Error	Mean Absolute Percentage Error
Decision Tree	0.971733358	0.209358914	135.7144595	11.64965491	2.489260847
GaussianNB	0.959706662	0.248151614	65.38778662	8.08627149	4.553516131
Random Forest	0.961733358	0.180233973	82.20770426	9.066846434	1.914725389

Market Demand (Daily)					
Algorithm	Correlation Coefficient	Mean Absolute Error	Mean Squared Error	Root Mean Squared Error	Mean Absolute Percentage Error
Decision Tree	0.970703544	0.208890969	115.4757136	10.74596267	2.251068507
GaussianNB	0.960501957	0.259148339	73.55919513	8.576665735	5.005734615
Random Forest	0.970501957	0.186045859	84.23430136	9.177924676	1.772921163

Ontario Demand (Daily)					
Algorithm	Correlation Coefficient	Mean Absolute Error	Mean Squared Error	Root Mean Squared Error	Mean Absolute Percentage Error
Decision Tree	0.973464941	0.212915302	136.6512869	11.68979413	2.63126068
GaussianNB	0.961953018	0.245484324	65.38352831	8.086008181	4.748494771
Random Forest	0.971953018	0.145521759	48.49554703	6.963874427	2.058297425

Market Demand (Monthly)					
Algorithm	Correlation Coefficient	Mean Absolute Error	Mean Squared Error	Root Mean Squared Error	Mean Absolute Percentage Error
Decision Tree	0.972341601	0.204351895	115.3432382	10.73979693	2.494398823
GaussianNB	0.961437906	0.257838091	73.55563875	8.576458403	4.749776219
Random Forest	0.970501957	0.164085166	54.18714974	7.361192141	1.959117798

Ontario Demand (Monthly)					
Algorithm	Correlation Coefficient	Mean Absolute Error	Mean Squared Error	Root Mean Squared Error	Mean Absolute Percentage Error
Decision Tree	0.974541323	0.20706598	135.7134768	11.64961273	2.854045446
GaussianNB	0.959706662	0.248151614	65.38778662	8.08627149	4.553516131
Random Forest	0.145521759	0.141268133	43.70595929	6.611048275	2.229414277

Figure 10: Comparison of Validation Matrices

From the validation matrices we can come to a conclusion that overall Random Forest Regressor performed better that other algorithms. But in certain case like for Mean Squared Error and Root Mean Squared Error GaussianNB provided better result sometimes close to Random Forest Regressor. Decision Tree Regressor preformed worst among all the three algorithms. But for Correlation Coefficient Decision Tree Regressor performed close to the Random Forest Regressor.

7. CHALLENGES

During the whole project we have encountered some challenges. The first challenge was to collect the dataset. We needed a consistent large dataset with sufficient number of features. At first our plan was to implement our algorithms with UNB's (University of New Brunswick) dataset but the dataset we collected did not meet our requirements and that's why we experimented with Ontario dataset. Another challenge was to find out proper algorithms to compare which can provide us continuous values. The paper we followed has compared different algorithms than ours because at first, we wanted to just follow the steps of the paper but later on we wanted to experiment with some new algorithms. So, we had to spent time finding out the proper algorithms with our requirements. The final challenge was to manage the time and complete the project within the time frame. In the beginning we started individual project and only about two weeks ago we built the team and started working together. So, managing the time was one of the crucial challenges for us.

8. INDIVIDUAL CONTRIBUTION

We both contributed equally to this project. Tolulope searched and analyzed different datasets, he found and imported the dataset from ieso and the weather station [2], [3]; he filtered and cleaned the datasets, catered for empty or missing data; he classified and compared the results on charts and tables among the different algorithms that were implemented. So, he was mainly responsible for all the datasets handling and results evaluation.

Rahman experimented the possible algorithms for the project and chose three algorithms for implementation. He implemented all the algorithms with Python and extracted all the predicted results in the excel file. As well as he implemented all the validation matrices to compare the efficiency of all the algorithms. So, he mainly responsible for all the implementation required for the project.

We wrote the project report together.

9. CONCLUSIONS

In this project we mainly compared the results of different machine learning algorithms to evaluate their performance. Though we planned to implement algorithm to predict the future demand, for the time constraint we could not implement that. So, we tried to find out which algorithm would perform better to predict within a range of time and compare the predictions with the original data. From the validation matrices result we can conclude that Random Forest Regressor performed better but, in some cases, GaussianNB performed close or better than Random Forest Regressor. For future work, we plan on developing a model that can predict the power demand for the future using deep learning approaches that contains one or more hidden layers e.g. Convolutional Neural Networks (CNNs).

10. REFERENCES

[1] T. Kim and S. Cho, *Intelligent Data Engineering and Automated Learning – IDEAL 2013*, vol. 8206. Springer International Publishing, 2013.

[2] Independent Electricity System Operator, “Hourly Demand Report.” [Online]. Available: http://reports.ieso.ca/public/Demand/. [Accessed: 23-Apr-2019].

[3] “Station Results - Historical Data - Climate - Environment and Climate Change Canada.” [Online]. Available: http://climate.weather.gc.ca/historical_data/search_historic_data_stations_e.html?searchType=stnProv&timeframe=1&lstProvince=ON&optLimit=yearRange&StartYear=2002&EndYear=2019&Year=2019&Month=4&Day=22&selRowPerPage=25&txtCentralLatMin=0&txtCentralLatSec=0&. [Accessed: 23-Apr-2019].

[4] Martin Hansen, “Prediction of Electricity Usage Using Convolutional Neural Networks,” 2017.

[5] K. T. Williams and J. D. Gomez, “Predicting future monthly residential energy consumption using building characteristics and climate data: A statistical learning approach,” *Energy Build.*, vol. 128, pp. 1–11, Sep. 2016.

[6] R. F. Berriel, A. T. Lopes, A. Rodrigues, F. M. Varejao, and T. Oliveira-Santos, “Monthly energy consumption forecast: A deep learning approach,” *Proc. Int. Jt. Conf. Neural Networks*, vol. 2017–May, pp. 4283–4290, 2017.

[7] F. Rodrigues, C. Cardeira, and J. M. F. Calado, “The Daily and Hourly Energy Consumption and Load Forecasting Using Artificial Neural Network Method: A Case Study Using a Set of 93 Households in Portugal,” *Energy Procedia*, vol. 62, pp. 220–229, Jan. 2014.

[8] “Data-Driven Insights with MATLAB Analytics: An Energy Load Forecasting Case Study - MATLAB & Simulink.” [Online]. Available: https://www.mathworks.com/company/newsletters/articles/data-driven-insights-with-matlab-analytics-an-energy-load-forecasting-case-study.html. [Accessed: 23-Apr-2019].

[9] J. Van Der Merwe, “Carrots Monthly Market Report,” *Amt*.