**Energy Forecasting: BigDEAL**

**Challenge**

Francisco Flores - 0178269 – frflores@cougarnet.uh.edu

Shashank Bhushan – 2093536 – sbhusha3@cougarnet.uh.edu

Anubhav Dixit - 2100232 – adixit4@cougarnet.uh.edu

**Submission Date: December 1, 2022**

**Course Name: INDE6360 Engineering Analytics**

**ABSTRACT:** The power systems around us are rapidly evolving and this rapid evolution with itself brings lots of benefits to humankind in terms of energy demand fulfillment at the very end. But these advancements also bring lots of challenges with them and as a result of these many upgrades, one such challenge is that the security and stability of the power supply grid is always at a probability of risk because of peak load demand spikes. These sharp increase in demand are characterized by stronger randomness and less predictability. To this end, this research paper offers a comparative machine learning based model building for application to peak load demand forecasting techniques. In this study, we first provide a clear and comprehensive problem statement for forecasting peak load demand. Second, we will give some overview of Dataset used for working on BigDEAL Challenge which will further going to get continued by identification of proposed methods and justification for the application of these methods on the data provided by Dr. Tao Hong. Third, we will move further with the insights into the data analysis part of our project using proposed methods of Time Series Forecasting Using Regression, Boxplots, Decision Tree, and Random Forest machine learning methods. We will provide a detailed explanation of how we tried to fit provided data to get best forecasting results and how we are going to minimize the mean square error and other errors in our trained models. Finally, we will conclude with key takeaways from our results and possibilities of future work that can happen on this data.

**INTRODUCTION**

The power systems around us are rapidly evolving due to the implementation of high performance generators, computerized relays, advanced metering infrastructure on the supply side as well as it is also getting upgraded vastly at demand side through advent of local energy utilization applications such as electric vehicles and other modern domestic and commercial load utilization applications, and also it is enhanced by an increase in the penetration of intermittent renewable energy at both transmission and distribution level. This fast-paced evolution with itself brings a lot of challenges in terms of peak load demand spikes in power requirements. These spikes are at many times highly random as well as they are not easily predictable by using modern sensing equipment. This is what makes it one of the biggest threats to stability of power supply grids and very often leads to failure of its security which ultimately results in big events like power grid failures and blackouts.

As per the study, it is seen that among different power generation units, peak load demand units are the worst performing units with lowest efficiency and highest cost. These units have so much bad efficiency that even a 5 to 15 percent reduction in peak load demand spikes can make a humongous difference in terms of cost saving on the investment of safety equipment, energy saving in the power supply lines without any need to worry about losses in power supply systems and protection from grid failures due to de-synchronization of supply systems from different sources of power supply. Controlling these factors are surely going to help with effective Peak Load demand management. This makes Peak Load Demand forecast a major area of focus for industry leaders and researchers. Dr. Tao Hong is one such researcher from UNC Charlotte whose major focus is on Energy Forecasting, and he is also the director of Big Data Energy Analytics Laboratory (BigDEAL).

The real need or motivation for Peak Load Demand Forecast and its benefitted parties and stakeholders are listed as follows:

1. Grid Operators: Obviously, Grid Operators are the party which are going to get benefitted most from Peak Load Demand Forecast. They will be able to increase power generation equipment utilization rate. They can also reduce cost of power generation and can avoid pressure on grid supply during peak load hours.
2. Electricity Retailers: With less losses on the supply side these retailer companies can fix electricity tariffs so as to increase the power supply quality and can also maximize their profits. These retailers can also offer different schemes for bigger and commercial consumers of power supply to help them avoid consuming more power during peak load demand hours and help them with keeping their internal unit power utilization in an efficient manner.
3. Commercial End Users: They will plan and manage their electrical power utilization accordingly and can save huge number of sums on their electricity bills.
4. Domestic End Users: They can go for usage of smart electrical devices which are not only cost efficient but also energy efficient.
5. Government: Government as a big player can frame policies which are widely effective for regional economic growth of a nation or a state in terms of efficient power utilization and can utilize the saved money on social welfare and other works.

Coming to the challenge part, it was posted on October 2nd, 2022, by Dr. Tao Hong and the cleaned data containing exogenous variables like Weather Variables, Calendar Variables from the year 2002 to year 2006 was provided on October 29th, 2022. Load data was provided for every hour from the year 2002 to year 2006. Prediction was to be done for the year 2007. Next phase of steps included to identify and apply machine learning techniques like Regression Trees, Pruned Trees, Boosted Trees, Bagged and Random Forest techniques to our refined data. The main focus was to predict the load data for year 2007 and minimize errors in different techniques. After application of different models Root Mean Square Error for each prediction got minimized and the best results of around 5-6% got achieved by our Random Forest model. Considering our expertise in the concepts and the tight schedule of competition our model capacities worked at satisfactory levels in which Random Forest Model was the best performing. Overall, it was a nice experience to participate in a time bound energy forecasting challenge and we gained an impressive experience of applying different machine learning techniques to real-world data.

**PROBLEM STATEMENT**

The theme of BigDEAL Challenge 2022 is peak load forecasting. The competition includes a qualifying match and a final match.

The qualifying match is ex post one-year ahead forecasting. The match includes three tracks: 1) H (hourly loads of one year); 2) M (magnitude of 365 daily peaks); and 3) T (timing of 365 daily peaks).

The final match is ex ante day ahead peak load forecasting.

**DATA SOURCE**

Qualifying Match Data provided by Duke Energy for the BigDEAL Challenge 2022. The Data set contains,

* Single excel file containing load data for past years
* 5 Columns containing data for time, ex. Year, Month
* 4 Columns containing data for temperature
* 1 Column containing load data from past years

Calendar

Description automatically generated with low confidence

**BIBLIOGRAPHY**

1. Electrical peak demand forecasting – A review: Shuang Dai, Fanlin Meng, Hongsheng Dai, Qian Wang, Xizhong Chen (2021)
2. Forecasting electricity load: Interview with Prof. Tao Hong:

<https://www.youtube.com/watch?v=ypD6f4K0LTU>

1. Lecture 12 Peak Load Forecasting: Kern

https://www.youtube.com/watch?v=A6C51j9gDBU&t=1265s

1. Dr. Tao Hong BigDEAL Laboratory Page:

<http://www.drhongtao.com/bigdeal>

1. Peak Load Forecasting - Load Forecasting - Power System Planning and Reliability: Ekeeda

<https://www.youtube.com/watch?v=cmv1BV5UO9Y>

1. New York ISO Independent System Operators:

<https://www.nyiso.com/real-time-dashboard>