2-2022

Predicting Energy Consumption during Expo 2020

Rashed Waleed Albishre

RIT

Predicting Energy Consumption during Expo 2020

by

Rashed Waleed Albishre

A Capstone Submitted in Partial Fulfilment of the Requirements for the Degree of Master of Science in Professional Studies:

Data Analytics

Department of Graduate Programs & Research

Rochester Institute of Technology

RIT Dubai

February 2022

RIT

Master of Science in Professional Studies:

Data Analytics

Graduate Capstone Approval

Student Name**: Rashed Waleed Albishre**

Graduate Capstone Title**: Predicting Energy Consumption during Expo 2020**

Graduate Capstone Committee:

Name: Dr. Sanjay Modak Date:

Chair of committee

Name: Dr. Ehsan Warriachs Date:

Member of committee

Acknowledgments

Firstly, I would like to thank Dr. Sanjay Modak and Dr. Ehsan Warriach, my mentor, for their great support, and I am grateful for the efforts they made to make this happen. Secondly, I would love to thank my family for the endless support they have shown throughout my studying years, and how they encouraged me to continue my higher education. I also want to acknowledge my classmates for all the help they have given during the program. Lastly, a special thanks to the university and everyone working in it for providing such an excellent education to their students.

Abstract

The art of the Expo event is the bloom in the economy of the country where it is hosted along with the collaboration among the various areas of the country. Dubai availed the privilege of hosting Expo 2020. Dubai being a center of attention for tourist attraction and the growth in its economy, took pride in hosting the event while working on themes and ways to make the event stand out. Expo is known to bring visitors from around the world to one country for a duration of six months where the visitors are provided with an opportunity to enjoy the prestigious setup and festivities. While the event is successful in its ability to bring together people and work towards the enhancement of the economy in terms of the investments made and profits earned, one drawback resonates in the ways the energy is consumed. When focusing on Dubai as the host of the event, several key factors that are highlighted are the usage of various Expos and infrastructures to create a connection among communities while being able to accomplish and abide by the set goal and themes. However, in doing so the energy consumption seems to rise where several natural resources are used as well as impact on the environment is made. Thus, this report focuses on incorporating an econometric model to gain an insight to the growing energy demand. The model allows the user to evaluate the individual energy consumption demands by each of the involved infrastructures to reach a conclusion about the collective energy requirement. The proposal is then able to make suggestions about the ways that can be incorporated to limit the energy usage or to manage it in a wiser way.

Keywords: Machine Learning, Data Analysis, Dubai Expo, Energy Consumption, Prediction.

Table of Contents

List of Figures

List of Tables

Chapter 1

1.1 Background

An Expo 2020 provides innovations that are conducted by the collaboration of people together to give solutions for global provocation. With creative minds, the inventors build a future that has a great impact around the world. These events allow pioneers to share their ideas on how to enhance the world that humanity is facing. They exhibit their models and technologies to the showcase. An expectation was made that there will be an increase of 50 percent in electricity by 2020 from 2012. It was anticipated that 25 million visitors will attend Expo 2020 and massive amounts of electricity would have to be generated in order to cater the requirements of the event and the visitors attending. Significantly, Dubai has been observed to demonstrate a growth in GDP and economy over the years.

The increase in productivity and boom in economy can be represented through an increase in electricity consumption from 12240 GWh in 2001 to 44570 GWh in 2018 (Dahan, 2020). The growth rate in electricity contributes to a rising factor of 8.4 percent a year (Dahan, 2020). Economical activities and population growth are the factors to control the consumption of electricity. Dubai requires an investment of 250000 Mwh of power for Expo 2020 and a designated amount of around 7 billion dollars for infrastructure. Due to the increase of the economy, there will be a relative increase of electrical demand between 4-6 percent annually. As a result, the electricity consumption of Dubai will reach 9.6 GW, showing a 50 percent elevation in comparison to the levels from 2012 (Saraf, 2014).

The increase of budget annually is utilized to implement infrastructures with clean coal, solar, and smart grids (Saraf, 2014). Moreover, to increase the capacity of power generation an additional investment of 2 billion dollars will be required. United Arab Emirates receives 99 percent of its fuel mainly through natural reservoirs of natural gas and the remaining fuel is derived from solar power through the set solar panels. Therefore, Expo 2020 implemented a sustainable solution to reduce the use of electricity to minimize the consumption of energy. For instance, the Expo 2020 will be incorporating energy efficient technologies to conserve energy by a thorough and careful designing and construction of the pavilions, Expos, and public realm (Saraf, 2014). Several components in the households require elements of energy consumption that comes from electricity, gas and water usage (Teba, 2018).

Moreover, the additional sustainable methods that will be taken into consideration will work towards carbon footprint, recycling, reuse of materials, and attain neutral balance for water (Saraf, 2014). Another initiative is to gain energy from renewable sources by using only 50 percent of the total energy (Saraf, 2014). To achieve that, developers need to manufacture photovoltaic (pv) solar panels to install for the top roof and into the walls of the Expos (Saraf, 2014). Solar energy is an efficient way to sustain development in the Middle East and generation powers. In Emirates, the main fuel that has been used by the year of 2019 is gas followed by solar energy and then coal (Ritchie & Roser, 2021).

1.2 Statement of problem

The choice of selecting Dubai as a host for expo 2020 means making it a centre of attraction for 182 countries for a six-months duration. The attraction of various countries also means a rise in the investments made. It is approximated that the expo 2020 would generate an investment of approximately US $100 billion US $150 billion. However, what is the downfall to all the positive aspects of hosting expo 2020 at Dubai? When considering the downfall, the key element that comes to mind is the idea of the electricity supply to cater to almost 25 million people that would attend expo and to make the Expo city ready. Thus, the concern for energy consumption hinders the path of hosting such an event. This study sheds light on using a predictive model to determine the energy consumption during Expo 2020.

When considering the growth of any economy, the two key elements that are constantly brought up are the resources used and the energy generated. While energy generation is an expensive process, its consumption is not controlled by means of efficiency. With different sectors working together to generate the revenue for the economy, it becomes challenging to monitor the distribution of energy consumption. Thus, emphasis on a prediction model is required that can monitor the energy consumption and limit it when required and possible.

1.3 Project goals

To explore the energy consumption during Expo 2020 held at Dubai by using a predictive model. The Expo 2020 is a rich event in terms of the utilization of infrastructure, the setup of activities and mini themes to cater to a diverse range of countries and visitors. Thus, the project aims towards using a model that allows the navigation to discover the energy consumption and ways in which it could be modified.

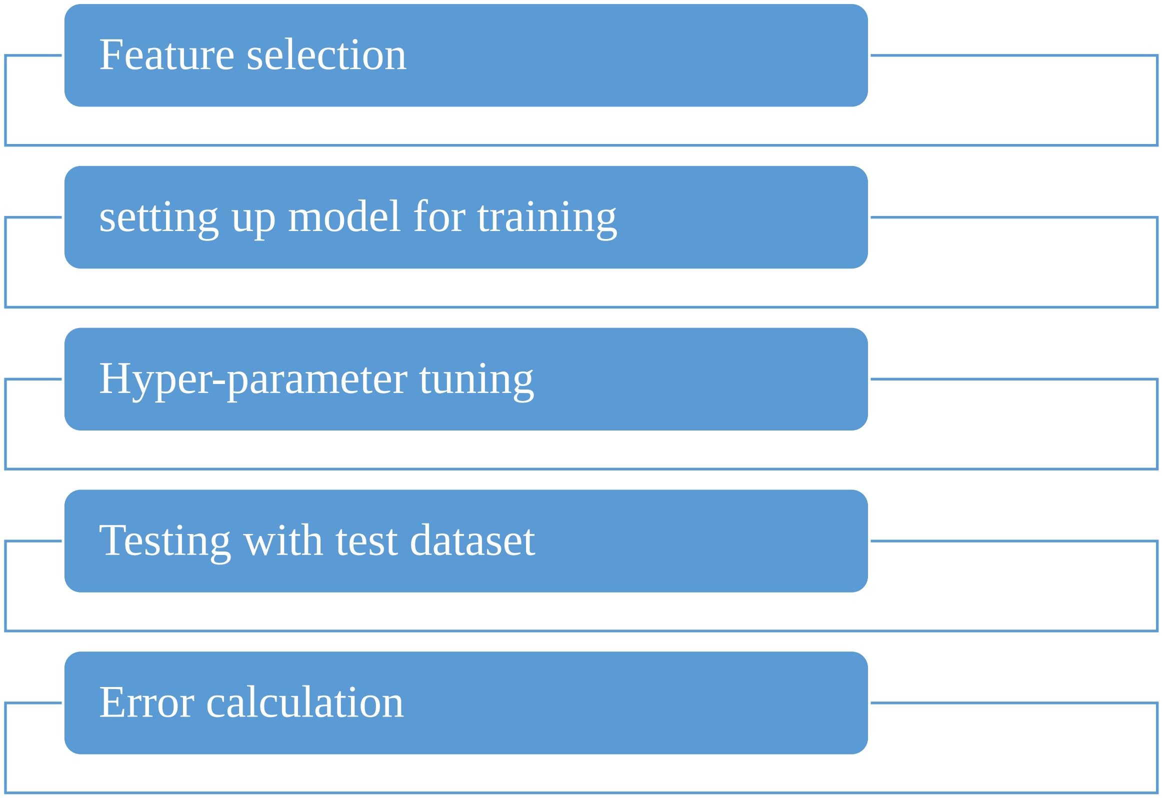
The key goals of the project are as follows:

- To determine the distribution of energy consumption in Expo 2020.
- To evaluate the usage of energy during the six month event period by using a mathematical model.
- To evaluate the impact of using sustainable alternatives to saving energy consumption.
- To recommend ways in which energy consumption can be monitored and managed.
- To evaluate the current energy consumption of UAE.
- Work towards incorporating a predictive model that can be used by any business or event to determine their energy consumption.
- Incorporate effective strategies that enable the business to work along with the model in determining ways of managing energy consumption.

1.4 Methodology

We went through these steps below, but we have summarized them in three sections which are initialization, visualization and simulation and model testing. I will use KNN algorithm beside LTSM. k-nearest neighbors (KNN) algorithm, to help solve classification and simple problem, as well as long-short-term memory (LTSM) which is helpful in the field of deep learning, classifying processing and making predictions. Both will be explained deeper later in the project analysis part.

Methodology Steps



Data Pre-processing

In-built python libraries are imported to the working notebook to pre-process the data. The main libraries used here are Pandas and NumPy. Pandas here assisted us to read the data stored on computer\sin CSV format. Pre-processing of data entails discovering any abnormalities or missing values in data. We have two CSV files two, one having genuine Dubai expo data for year 2020 and other having output of energy additional simulations for 2020. Following stages outline the pre-processing of data for this study.

- Load Required Libraries, Panda and Numpy
- Load CSV data file
- Clean data

- Remove duplicates and null values.
- Analyze the data create prediction model

Python Libraries

Pandas

- Pandas is a Python machine learning package that is essential to any work done in the field. Listed below are a few things that it can assist with and that are relevant to this work:
- It facilitates the reading and saving of data in CSV format.
- It converts the data from a CSV file that is kept in the system to a data Frame that is simple to deal with.
- It computes the most important statistics of data, such as the mean, median, maximum, and minimum.
- Provides the distribution of the whole set of data or the distribution of a specific column in the set of data.
- Resampling data and managing missing data
- Merging two datasets and taking a subset of a data collection
- Rearranging the dataset in pandas time-series format
- Creating a time-series format for the dataset

Numpy

NumPy is another essential Python library to have on hand. It is used in the data analysis process to execute scientific computing activities.

It has the following features:

• Practical linear algebra, Fourier transform, and random number capabilities

• An object that represents an N-dimensional array of data.

1.1 Limitations of the Study

Unfortunately, one of the biggest obstacles that I have faced is the lack of information. After contacting Dewa to provide the needed data sets, such as power and consumable energy in Dubai before and during expo 2020, to be able to compare, I have been told that it is not possible because it is considered as highly secured data, therefore my request was denied. As well as my computer having limitations, such as not having big server's performance to run huge amount of data.

Chapter 2 – Literature Review

Due to the very recent nature of the announcement, there is almost no literature dealing directly with the relationship between Dubai and the Expo 2020. However, there has been considerable commentary in the form of political and economic punditry, much of which is academic. Piers Schreiber, Vice President of Corporate Communications & Public Affairs at the Jumeirah Group, claimed that ‘the Expo will create up to 270,000 jobs in the region, bringing great economic and social benefits’ (cited in Wilson, 2013).

Among these are an injection of roughly €17.7 billion into the economy and a migration of talent from abroad (Wilson, 2013). These estimations are supported by the forecasting group Oxford Economics, which claims that the event will contribute nearly $40 billion to Dubai’s GDP and create 277,000 new jobs over the next seven years (Big News Network, 2013). A similar argument is put forward by Rose and Spiegel (2009), whose work suggests that ‘mega events’ lead to a substantial increase in trade (approx. 30%); however, they also show that ‘unsuccessful bids to host the Olympics have a similar positive impact on exports…trade is attributable to the signal a country sends when bidding to host the games, rather than the act of actually holding a mega-event’ (p1). While all of this boils down to the positive impacts on country’s economy and boost in tourism, it can also result in rapid environmental degradation by way of increased energy consumption and arrival of tourists as shown in a study conducted in turkey which showed positive relationship among energy consumption, increased tourism and environmental degradation (Katircioglu, S. T. 2014). This warrants the need for sustainable energy source and beforehand planning of energy consumption. A study concluded that during

Milan expo 2015, emission of one ton CO_2 per square meter of exhibition was recorded (Gallo, et al., 2020).

Based on comparisons with other ‘mega events’, it is often argued that the effects of the Expo 2020 are likely to be negative for Dubai. Rose and Spiegel (2009). The same is argued by Nitsch and Wendland (2013), who also point to the large initial investment in facilities and infrastructure associated with mega events, which can place a considerable burden on the local or national economy; there tends also to be a dramatic and unpredictable effect on property prices. The conclusion of Nitsch and Wendland (2013) is that mega events tend to have an overwhelmingly negative effect on population growth (i.e., a population decline), as measured relative to a control group.

However, Nitsch and Wendland (2013), and Rose and Spiegel (2009), point to the difficulty of estimating the impact of major events. Problems quantifying the effects, especially on phenomena such as labor markets, are often exacerbated by the fact that many studies are commissioned ex ante by biased groups. Moreover, in conducting analyses such as this, it is difficult to find an adequate sample size due to the infrequency of mega events. Then there are problems with the intangibility of cultural and social spillover effects, as well as the economic multiplier (Rose and Spiegel, 2009).

To prepare for Expo 2020, an important fact about electricity is that the intensity of power displays a rapid growth rate economically in relation to growth rate of electricity. That in return reveals that the GDP is expected to reach Dh410 billion in 2019 (Dahan, 2020). With this increase in GDP, more tourists will arrive, and the population will increase. This will lead to an overall surge in energy use which is other than the expo. It also means that Dubai will need to increase its energy reservoirs in order to meet with the increased demand. The money and income will be invested in increasing energy resource which will also result in increased use of fuel to meet the demand. Dubai's main source of energy is natural gas. Although steps are being taken to move towards sustainable energy (Taleb, H. M., 2015). In Dubai 100% of the population have access to electricity, in 2019, electricity consumption of the country was 130.03 TWh- (Hannah Ritchie & Max Roser, 2020) which is said to be increased in 2020 due to Expo.

The two essential factors that are preferred for economic growth are energy and materials. United Arab Emirates is a country of attraction where around 20 million visitors will be drawn to Expo 2020 resulting in the doubling of the record standing of 11 million visitors that attended in the year of 2012 (Dahan, 2020). To accommodate the visitors an infrastructure expansion will be required. The elasticities are in high demand for electricity depending on the income and price as follows in respectively; the inelasticity of (0.80 and for the elasticities of -0.96) (Dahan, 2020).

The goal of Dubai's economy resonates in its ability to provide continuous evolution for the future and sufficient growth economically. Based on the research conducted the total demand of energy consumption will gain a rise to 44816 GWh in 2019 and to 44696 GWh in 2020 prior to EXPO 2020, and then will eventually slows down to reach 43709 GWh in 2021 and 40606 GWh in 2022 (Dahan, 2020). In addition, a positive correlation between wealth and energy consumption is observed. It has been demonstrated that the wealthier the country, the higher the consumption in energy there is (Arouri et al. 2012). On the other hand, some studies have shown results that contradicts this statement. The studies presented, when a country becomes more developed and wealthy, their energy consumption decreases by virtue of using efficient machinery and use of sustainable sources (Ouyang & Li, 2018). Hence, electricity serves as one of the most important sources of living in Dubai and attention should be paid to enhance this resource in a way that the strategy can be applied to developing countries.

The sector of manufacturing uses about 65 percent of energy distribution for industrial energy consumption (Kant & Sangwan, 2015). Also, machine appliances are very common in the manufacturing industry.

Consumption of electrical energy that is used from machine appliances causes approximately 99% environmental impacts (Kant & Sangwan, 2015). Thus, it has been observed that the implementation of machining parameters can lead to the usage of lower rated motors, drives and auxiliary equipment (Kant & Sangwan, 2015). Such a convention will then result in a decline in the energy consumption (Kant & Sangwan, 2015). In addition, when analyzing energy consumption, other variables like parameters, performance, the reliability, and cost of materials play a key role in understanding ways to reduce the energy consumption (Kant & Sangwan, 2015).

When referring to energy efficiency, models are incorporated to investigate means of technologies and standard operating procedures that can be effectively implemented to reduce the volume of energy per unit of industrial production (Kant & Sangwan, 2015). The United Arab Emirates is the 7th largest country that produces the largest petroleum in the world (Kant & Sangwan, 2015). The consumption pattern of electrical energy in UAE during the last period has been analyzed as a function of population and weather sensitive parameters. Linear multiple regression models of energy consumption for different seasons have been developed.

The models account for consumption variations during the winter, summer and post monsoon seasons (Ranjan M, 1999). The country's economy is dependent on their fuel such as natural gas and petroleum (Kant & Sangwan, 2015). Dubai is eventually being the vital financial and trading center in the Middle East. The long-term Investments in non-energy sectors, such as infrastructure and technology work towards providing UAE with insurance against declines in oil price declines and any contributing factors to global economic stagnation (Eia, 2020).

For internal demand growing (Eia, 2020). The growth will resonate from the country's high-sulfur gas deposits (Eia, 2020). This investigation work is carried out throughout the design of an Expo's structural plan. Accurate projections may be obtained by thoroughly analysing each influencing aspect. The area of the Expo structure is one of the most energy-intensive components in European nations, accounting for around a quarter of total energy consumption.

Approximately 40% to 50% of the available energy is used. According to the paper, the authors propose three options for predicting the energy use. First method is the core technique which is subject to physical models, which are restricted in their application to three category groups of similar items Next, in the second technique, the computational model is used. The assumption of energy use is based on machine intelligence models, which are utilized.

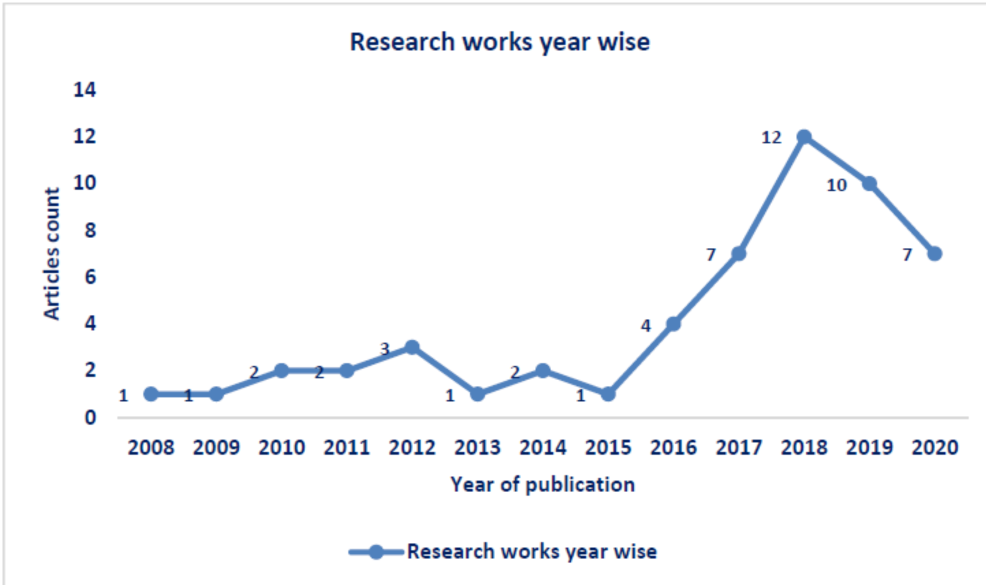


Figure 1 Research Over the years

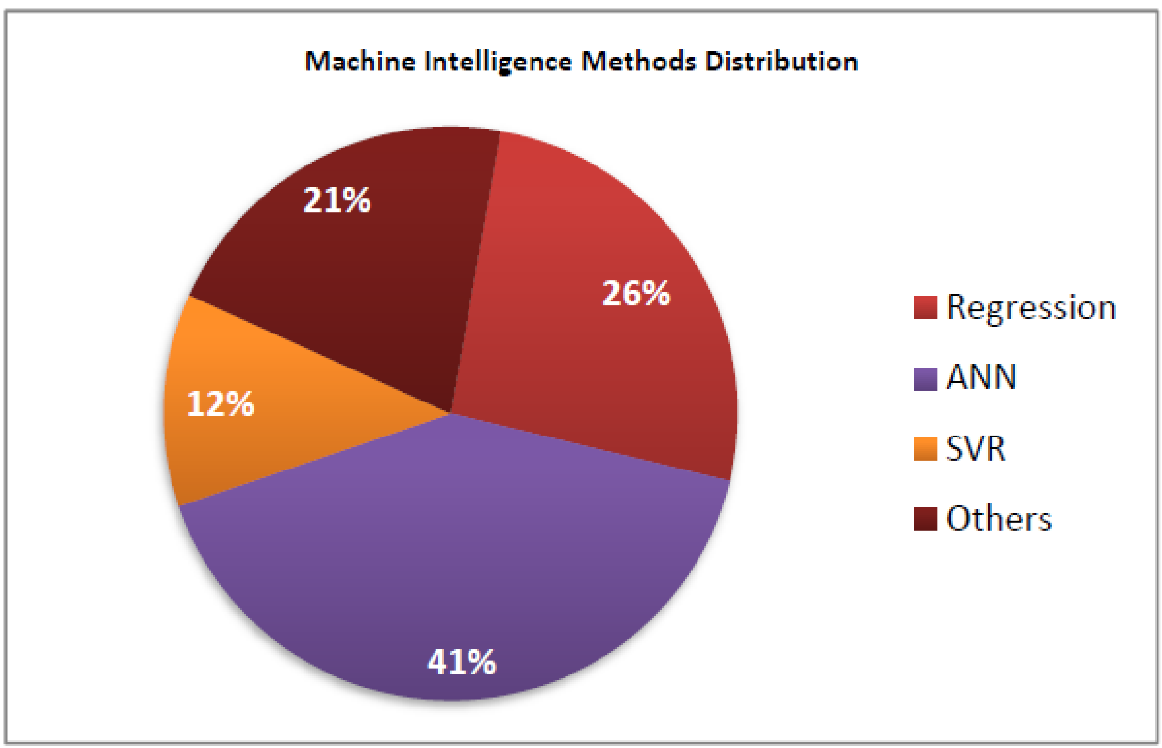


Figure 2 Machine Learning Algorithms

The latest of them is Hybrid systems are a kind of approach that employs both physical and computational intelligence. A set of methods that propose sensor-based models that are based on machine learning processes. Create correlations between consumption and effect variables (such as temperature and season) to test hypotheses (during the day, previous usage). While sensor-based prediction models have been utilized, there is a need for more.

There hasn't been much research done on the structural integrity of public sector Expo structures. This information-driven technique is used to manage private Expo constructions with several labels. In a nutshell the authors of a research reported on the energy consumption in residential and commercial metropolitan areas. The previous year of expo is base for the study, with a full review of previous, existing, and predicted trends in energy consumption for the per capita energy consumption needs of a country. This study also offers policymakers with a quick concept on how to monitor and handle the situation. To make energy more efficient and viable, we need to make new plan for power consumption.

From both techniques i.e., computational machine intelligence technique and the machine learning method, the most effective one the computational machine intelligence approach which combines the previous two methods and is the most advanced technique.

Artificial intelligence-enabled approaches for forecasting in urban area Expos are being explored. Creating a hybrid model by combining the two approaches that have been provided, this model is a combination of two different types of models.

In order to evaluate its performance, it was also compared to a few other models that were already in use. When compared to the other, the research indicated that it had greater predicting ability.

Models that are being considered Table 6 provides a quick comparison of the various forecasting techniques. Approaches that take specific factors into consideration. Engineering approaches that are based on statistics are often used. Techniques, neural network-based models, and SVM models were employed to accomplish the goals of the study. The research and analysis for all of the models, the same data collection and instances were utilized to create them.

Basically, it's the comparison of different kinds of computational techniques based on various criteria which shows the amount of complexity of the data, the utility of the model, and the speed with which it can be executed here. A good data complexity level means that the data set samples are either fewer or that the data set samples are more diverse. These models make use of a restricted number of feature sets. After that, an effort was made to depict the greatest possible computational technique determined by the precision of the forecast. In accordance with the calculation time, it is indicated in this research that the execution speed is essential for each model. It had been noticed that when it comes to complexity of data sets, execution rate, and precision accuracy,

Neural network models are used for outperforming their counterparts in terms of performance. Even though it is understood that neural models are difficult to utilize, they still serve as one of the most important models as compared to the competition for many applications especially in the analysis of energy usage.

The error rate analysis showed variance on case-to-case basis which indicated that it was less accurate as compared to the regression methodology. So, it can be concluded that the ANN can outperforms other nonlinear models in terms of accuracy and precision. A unique technique for the computation of the energy consumption of a commercial Expo structure in an urban region may be developed from a specified count of distinct Expo design characteristics utilizing public data obtained from various residential or commercial sectors.

In a few further experiments, the researchers used various technologies to examine the conveyance of energy and the factors of energy consumption in large Expo structures located across a metropolis. According to the findings of the investigation, these arrangements are only applicable to a limited fraction of the Expo structures in a particular city, and thus are only capable of regulating a tiny amount of energy consumption at the metropolitan level. For this reason, it proposed a proactive model of energy consumption that was driven by an information-driven technique for combining city level infrastructures to overcome this restriction.

It was decided to use computational methods such as linear regression and neural networks to the city's infrastructure energy management to anticipate the city's future power consumption. The algorithms utilized were support vector machines, neural networks, and random forests. The correctness of the model prototype was evaluated throughout the Expo and postal division stages, with the real usage of information being used to determine results. In a few research papers, the authors described an optimization-driven strategy to integrating several information mining approaches, including random forests, hierarchical clustering, and k-medoids clustering.

A deeper dive into the subject resulted in the development of a deep learning predictive methodology combined with a long-term transient memory intelligence model for regulating energy consumption in urban Expos in order to reduce the amount of information being transmitted and to better coordinate the inter-Expo impact with the information-driven energy model.

In this paper, we looked at a variety of machine learning methods and artificial intelligence techniques for forecasting Expo-level energy retention and optimization. The most important takeaway from the study is that machine-learning processes such as neural networks, support vector machines, decision trees, and K-means models are widely employed because they provide better results in less time than other models. From 2008 to the present, the deployment of machine-learning processes to handle Expo energy consumption projects has grown in popularity.

In the foreseeable future, it is projected to be a continuing source of curiosity and investigation. Apart from machine learning approaches, our study finds that deep learning models and optimization techniques have also played an important role in the research of energy consumption analysis in urban sector Expos, which is encouraging. As shown by the data-driven expectation techniques, they are more appropriate than engineering procedures in many situations. They provide a superb approach for open data (energy consumption, climate, transitory, and inhabitant), which is not difficult to secure from structures using the Internet of Things and communication propels (IoT).

Conclusion:

After going through various papers, it was apparent that this deep dive into the methods of prediction led us to the decision of using two main models for prediction purposes:

- RNN Model: This model is used for electricity forecasting over a long- or medium-term horizon of time. This model can also be used to calculate missing data. It also performs on a higher level as compared to 3 level perceptron neural network.
- LSTM Model: it is a new approach which applies effective function of random time which enhances its accuracy in forecasting. It has also been proven to be superior to other models.

On the other side it was very clear that the source of the dataset is very important in such a prediction, which assign us trying to bring a very accurate dataset from a much verified source such as Dubai electricity water authority (DEWA), or Dubai Static Center (DSC). Which are listed as governmental sector.

Chapter 3- Project Description

3.1 Project Overview

The project will go through several phases in which it will prepare the selected dataset to be utilized to test and model the proposed predictions. The project will start by the pre-processing steps of data cleaning, and normalization which consist of dealing with redundant, and missing data and making sure that the data the is present within the dataset to integrity. The dataset will be later explored by visualizing the data using numerous appliances to generate a diverse selection of visuals to gain insight from the different features and their relationships. Testing and modeling will focus on verifying our prediction by various means such as model comparison and several estimates regarding several dimensions through the power usage and years.

3.2 **Data Overview**:

System Energy Requirement (Gwh)
Years
Residential Energy Consumed Units (GWh)
Residential Number of consumers
Commercial Number of consumers
Commercial Energy Consumed Units (GWh)
Duration
Industrial Number of consumers
Industrial Energy Consumed Units (GWh)
Area Capacity
Other Number of consumers
Other Energy Consumed Units (GWh)
Total Number of consumers
Total Energy Consumed Units (GWh)
Power Maintenance

Chapter 4- Project Analysis

4.1 Initialization

```
import warnings
warnings.filterwarnings('ignore')
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
plt.style.use('fivethirtyeight')
```

The figure above shows the Loading of all required libraries and packages.

```
df = pd.read_csv('Electric_consumption.csv')
```

Upload dataset and safe in the data frame called df. Data file is in csv format

```
df.columns=['Date', 'Consumption']
df=df.dropna()
df['Date'] = pd.to_datetime(df['Date'])
df.set_index('Date', inplace=True) #set date as index
df.head()
```

	Consumption
Date	
2012-10-12	7.098
2012-10-13	11.087
2012-10-14	13.223
2012-10-15	10.257
2012-10-16	9.769

4.2 Consumption Visualization:

The below section will encompass the visualization for the power consumption between 2013 and 2014.

```
plt.xlabel("Date")
plt.ylabel("Consumption")
plt.title("production graph")
plt.plot(df)
```

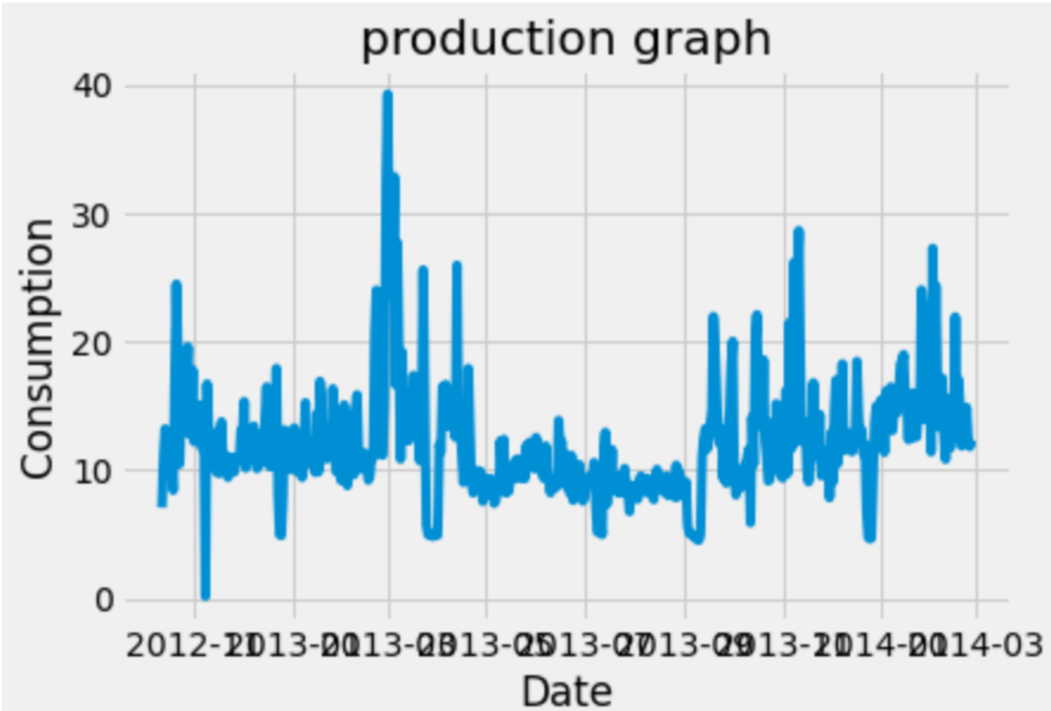
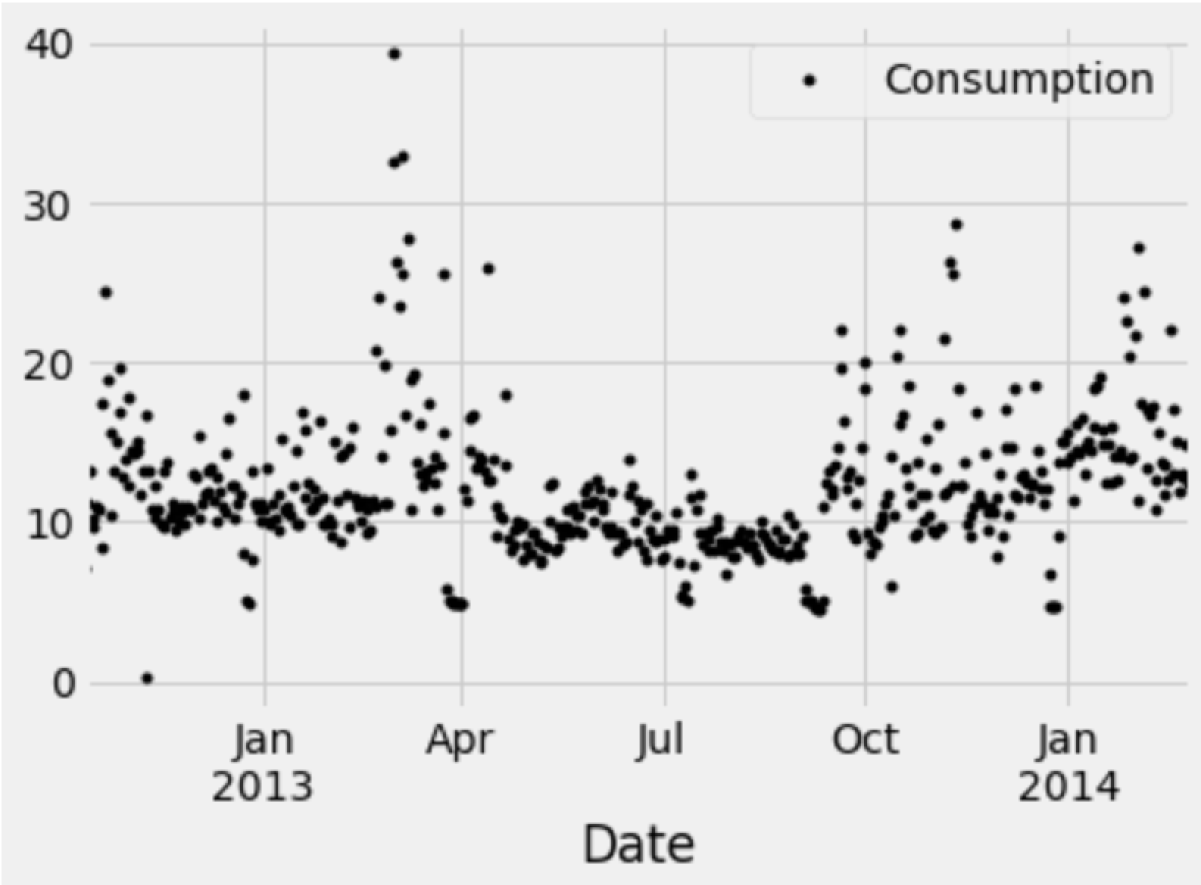


Figure 3 Energy Consumption through years.

```
df.plot(style='k.')
plt.show()
```



Both graphs above shows the distribution of the consumption ratio between January 2013 and January 2014.

Figure 4 Energy Consumption Actual

4.3 Simulations, and Model Testing:

The next section has been broken into two parts. The first portion is concerned with the simulation results obtained by SPSS software, while the second section is concerned with the performance of the ML model.

First, we compare dataset provided the actual power consumption and predicted power consumption. Where we can see the data fluctuation over the graph by using Simple RNN Model.

Figure 5 Prediction made by simple RNN model

Secondly, we compare dataset provided the actual power consumption and predicted power consumption. Where we can see the data fluctuation over the graph by using Simple LSTM Model. Each model gives different power consumption prediction over the same time and power.

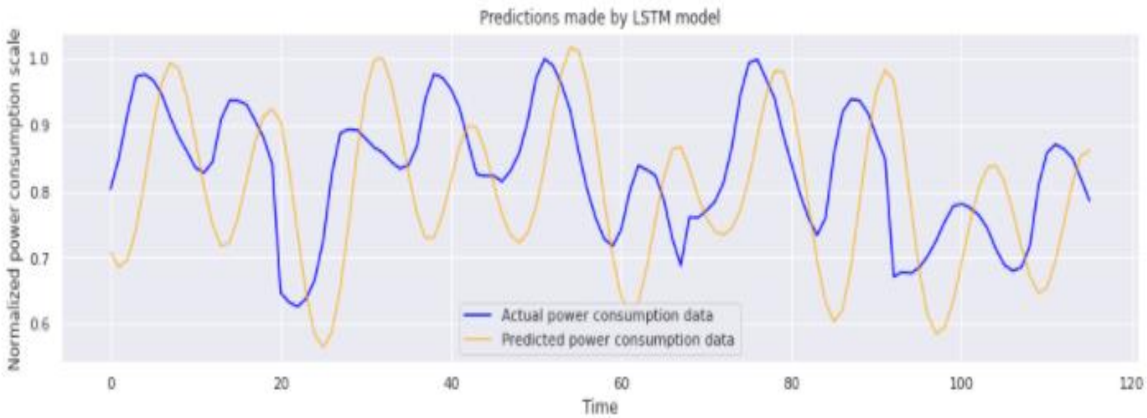


Figure 6 Predictions made by LTSM model

	Time Frame	Power Consumption Scale		Accuracy
		Actual	Predicted	
Simple RNN model	0	0.80	0.80	98-99%
	20	0.65	0.80	
	40	0.95	0.93	
	60	0.75	0.75	
	80	0.85	0.85	
	100	0.77	0.80	
	120	0.79	0.79	
LSTM model	0	0.81	0.7	75-80%
	20	0.65	0.93	
	40	0.95	0.82	
	60	0.62	0.71	
	80	0.88	0.95	
	100	0.78	0.7	
	120	0.79	0.87	

The above table shows the comparison of two model Simple RNN and LSTM with time frame, power consumption scale and accuracy. Here the prediction has been made on different time scale range from 0 to 120 with interval of 20 each. Power consumption scale is divided into two parts, first actual power consumption and second predicted power consumption by two models (Simple RNN and LSTM).

Prediction made by simple RNN shows similarity between actual power consumption and prediction power consumption, which is highest around 0.95 and 0.93 respectively at 40 time frame while lowest is 0.65 and 0.85 respectively at 20 time frame. From the above table it is shows that RNN model shows perfect prediction of power consumption. On the other hand LSTM model shows dissimilarity between actual power consumption and predicted power consumption. In LSTM at some time frame the predicted value shows higher power consumption than actual and at some time frame predicted value shows lower power consumption than actual.

When we compare both the model Simple RNN shows higher accuracy as actual power consumption and predicted power consumption are same which is around 98-99 % while LSTM model accuracy shows unpredicted power consumption between actual and predicted power consumption which is around 75-80%.

Analysis

- The original data was in numerous excel files, which I turned into one clear csv file after some data wrangling and cleaning.
- Importing the necessary python modules was followed by loading the dataset and saving it as a csv file (pandas data frame) for easy analysis.
- Cleaning up the data (Duplicates, null values, data types)
- Displaying the data as a data frame allows you to view it and begin studying the data and gaining insights based on trends.

4.4 Setting up the Model

The first step in creating the model is to divide the data into two sets: training and testing. The relationship between the two may be any ratio, but more training data equals greater learning, and that is exactly what the job requires. In the Scikit learn package for Python, there is an inherent function called train test split. The instructions that were used to divide the data set are shown in the following lines.

X_train, X_test, Y_train, Y_test= train_test_split (X, Y, shuffle=True, test_size=0.2)

The characteristics array is specified by the letter X, and the target value is represented by the letter Y. If the default settings are used within the brackets, this command will automatically divide the data into two groups: 75 percent for training and 25 percent for testing. If the default parameters are not used, this command will split the data manually. The following are some of the most significant criteria to consider:

Test_size = the percentage of test data in input data

Shuffle= whether to shuffle or not the data while splitting. True means yes.

After we have divided the data, we will present the model that will be employed, which will be a random forest in this case. We use the Random Forest Regressor from the ensemble techniques in the Scikit package to predict forest fires. There are just a few settings that need to be specified for a random forest. However, the default values of the parameters are utilized for the initial execution. Following are some example command lines that demonstrate how the model is fitted using the fit command:

```
rf = RandomForestRegressor (n_estimators=1000, random_state=42)
rf.fit(X_train, y_train)
```

4.5 Model for actual and predicted model

Starting with KNN model, k-nearest neighbours (KNN) algorithm is a simple, supervised machine learning algorithm that is mainly used to solve different issues in both and regression problems. It's easy to implement it and understand it but the hard part you may face in it is that it has a major drawback of becoming significantly slows as the size of that data in use grows**.**

Second comes the LTSM algorithm, long short-term memory (LSTM) is basically an artificial recurrent neural network which is used in the field of deep leaning. LSTM networks are well-suited to classifying, processing and making predictions based on time series data, since there can be lags of unknown duration between important events in a time series.

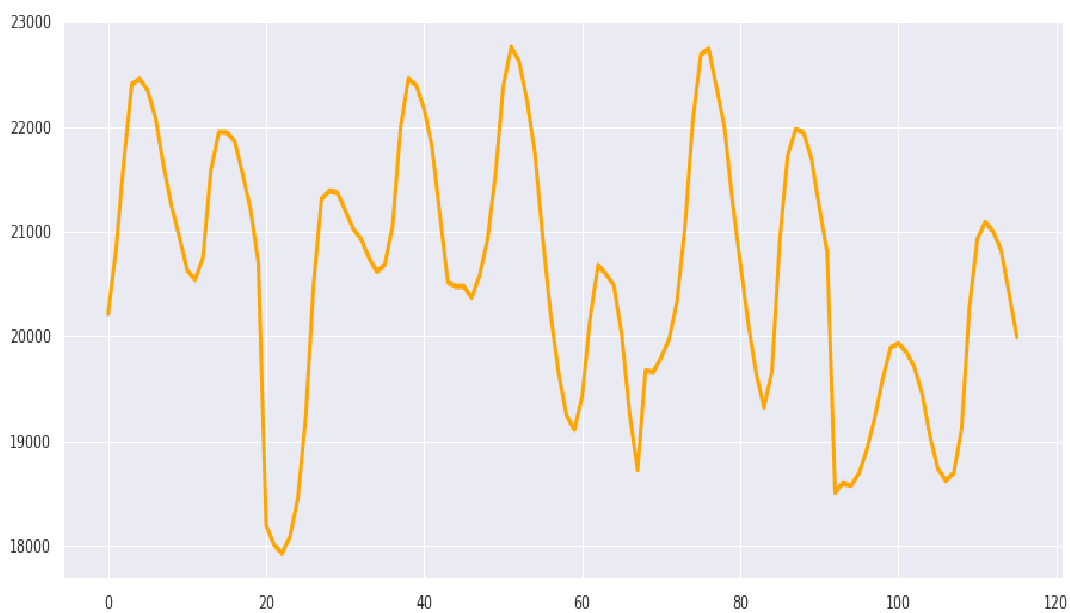


Figure 7 Actual and Predicted model Graphs

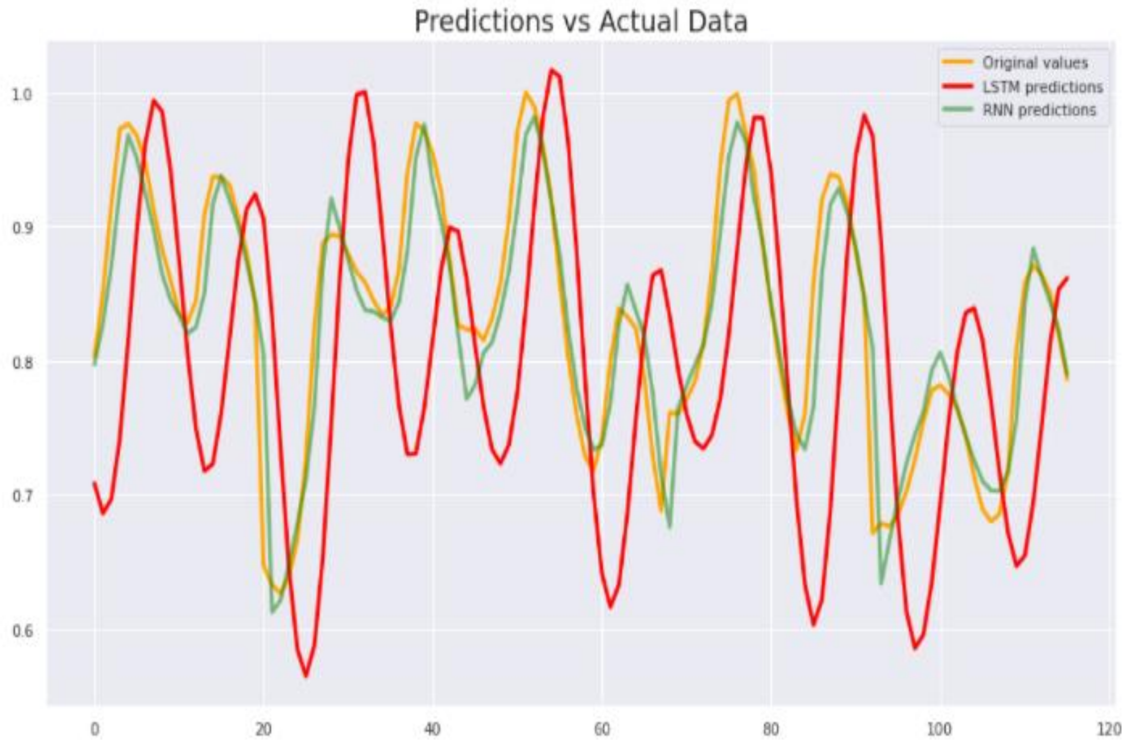


Figure 8 Prediction Vs Previous and actual data

Figure 8 above shows the actual and predicted data from our model. From the figure 7 we can see that there is a high correlation between the actual and the predicted data showing that the model if fit.

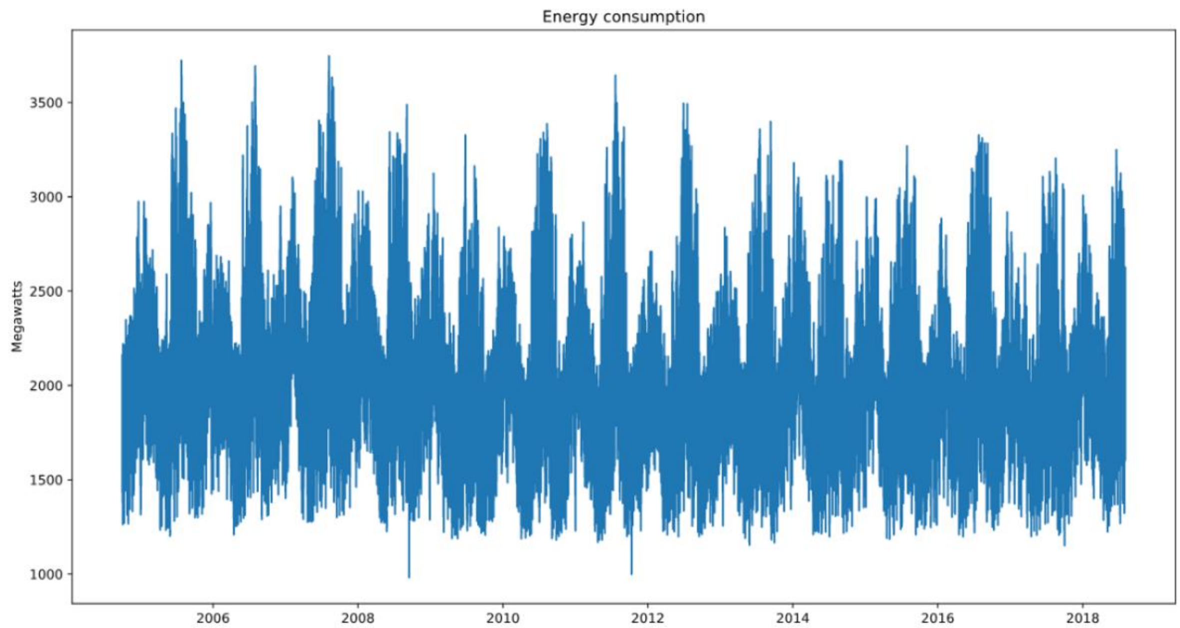
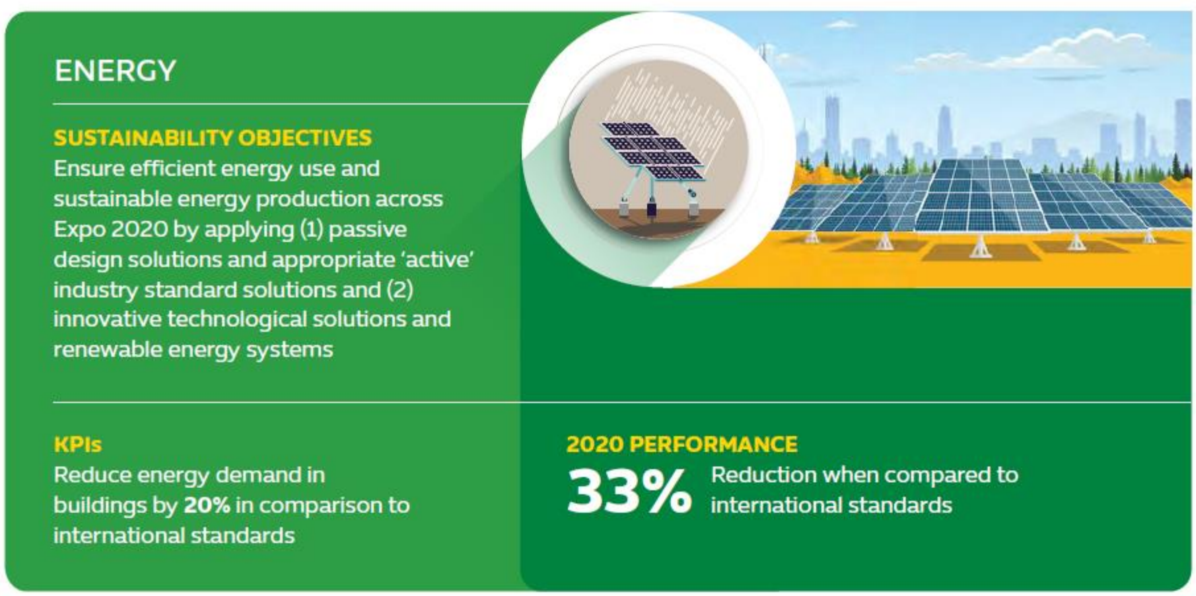


Figure 3 Energy Consumption in megawatts

The expo 2020 is one of the most prominent and workable power consumption plan over the decade where smart sustainable plan worked by using solar system and active industry standards.



The result analysis in python by using discussed libraries and algorithm we can see the less power consumption been predicted in the results of this research work. Over all 33% reduction is recorded in expo 2020 when we compared our result with the given data.

Chapter 5 Conclusion

5.1 Conclusion

To predict energy services in Expo, this research integrated numerical and predictive machine learning supervised methodologies. It was discovered that trying to acquire and prepare real-world power consumption data for use in EP simulations resulted in higher prediction outcomes. For the first and third quarters of the year, the forecast consumption based on actual consumption was closer to the predicted power consumption.

The Random Forest Regressor ML model was used to anticipate the energy services based on EP simulated data, and the results were promising. The model also included certain characteristics based on historical and engineering information, as well as data from the first year's metrology and some historical and engineering knowledge. The findings revealed that historical data, such as consumption in the previous hour, had a strong relationship with the output of the system. In this case, the Feature Selection stage resulted in more accurate findings while also saving computing time.

Specifically, the obsolete EP models were shown to have limitations since they underestimated the overall consumption while overestimating the amount of energy used by the HVAC system. EP models were most likely created a long time ago, and as a result, they do not take into consideration any changes that have occurred in the equipment or the energy use policy of the Expo.

Furthermore, the models were created using an outdated version of EP software (v8.2) that is no longer accessible on the company's website. The models were upgraded to the most recent version of the program (v8.8) by using a specific command provided by subsequent versions of the software to cope with similar issues. These conversions required a significant amount of time, and it is possible that they had an impact on performance.

The meteorological data was not readily accessible for the duration of the trials due to a lack of resources. It is estimated that over 15% of the power data was extrapolated or generated based on patterns in the data. Some meteorological entities were collected from another metrological station that had an impact on the performance and were used in this study.

5.2 Recommendations

Testing with more datasets may increase the accuracy of the information and prediction which will help in the diffrentiality and comparison between one places to another, especially if it was tested in other countries in the future, which will lead to more steady plans to overcome the usage of electricity. As well as you can use the same dataset but with different subsets with this data. As well as testing random combinations or specified sets is recommended due to the environment of predictions, the more the available data and tests, the more it’s easier to predict the right accurate thing.

5.3 Future Work

Further research should aim for trying new advances techniques in different ways of summarizing different subsets either from the same dataset or another. It is also recommended to use other dataset which can be more balanced and accurate such as datasets from the main supplier of electricity in the country that is being studied for example DEWA in Dubai. This will lead into calculating feature across different dataset which may or may not have similar predictions which by it turn will affect the approach and focus of the selection that will get you much more accurate models and create a greater prediction process. As well as for the future researches, computer limitation should be considered in terms of giving accurate predictions and avoid forcing the model to use all the available processors.

References

References

Arouri, M. E. H., Youssef, A. B., M'henni, H., & Rault, C. (2012). Energy consumption, economic growth and CO2 emissions in Middle East and North African countries. *Energy policy*, *45*, 342-349.

Azadeh, M. A., & Sohrabkhani, S. (2006, December). Annual electricity consumption forecasting with Neural Network in high energy consuming industrial sectors of Iran. In *2006 IEEE International Conference on Industrial Technology* (pp. 2166-2171). IEEE.

Badal, F. R., Das, P., Sarker, S. K., & Das, S. K. (2019). A survey on control issues in renewable energy integration and microgrid. *Protection and Control of Modern Power Systems*, *4*(1), 1-27.

Benedetti, M., Cesarotti, V., Introna, V., & Serranti, J. (2016). Energy consumption control automation using Artificial Neural Networks and adaptive algorithms: Proposal of a new methodology and case study. *Applied Energy*, *165*, 60-71.

Bilil, H., Aniba, G., & Maaroufi, M. (2014). Multiobjective optimization of renewable energy penetration rate in power systems. *Energy Procedia*, *50*, 368-375.

Dahan, A. (2020). The Economic Effect of Expo 2020 On The Demand For Electricity In Dubai Economics, Estimation, & Forecast (2001 – 2022). *Journal Of Critical Reviews, 7*(5). doi: 10.31838/jcr.07.05.148

Divina, F., Garcia Torres, M., Goméz Vela, F. A., & Vazquez Noguera, J. L. (2019). A comparative study of time series forecasting methods for short term electric energy consumption prediction in smart buildings. *Energies*, *12*(10), 1934.

Gómez, M., & Rodríguez, J. C. (2019). Energy consumption and financial development in NAFTA countries, 1971–2015. *Applied Sciences*, *9*(2), 302.

Kant, G., & Sangwan, K. S. (2015). Predictive modelling for energy consumption in machining using artificial neural network. *Procedia Cirp*, *37*, 205-210.

Muhammad, B. (2019). Energy consumption, CO2 emissions and economic growth in developed, emerging and Middle East and North Africa countries. *Energy*, *179*, 232-245.

Ouyang, Y., & Li, P. (2018). On the nexus of financial development, economic growth, and energy consumption in China: New perspective from a GMM panel VAR approach. *Energy Economics*, *71*, 238-252.

Ranjan, M., & Jain, V. K. (1999). Modelling of electrical energy consumption in Delhi. *Energy*, *24*(4), 351-361.

Ritchie, H., & Roser, M. (2022). *Energy*. Our World in Data. Retrieved 23 February 2022, from Energy consumption modelling using deep learning technique—a case study of EAF. *Procedia CIRP*, *72*, 1063-1068.

Sadorsky, P. (2011). Financial development and energy consumption in Central and Eastern European frontier economies. *Energy policy*, *39*(2), 999-1006.

Taleb, H. M. (2015). Natural ventilation as energy efficient solution for achieving low-energy houses in Dubai. *Energy and buildings*, *99*, 284-291.