Prediction of Energy consumption in a building using Machine Learning

79.Prashik Patil, 44.Narendra Golla, 62.Ramyasri Karukuri, 009.Nitin Sai Kumar

DEPARTMENT OF CHEMICAL ENGINEERING

VNIT NAGPUR

Abstract

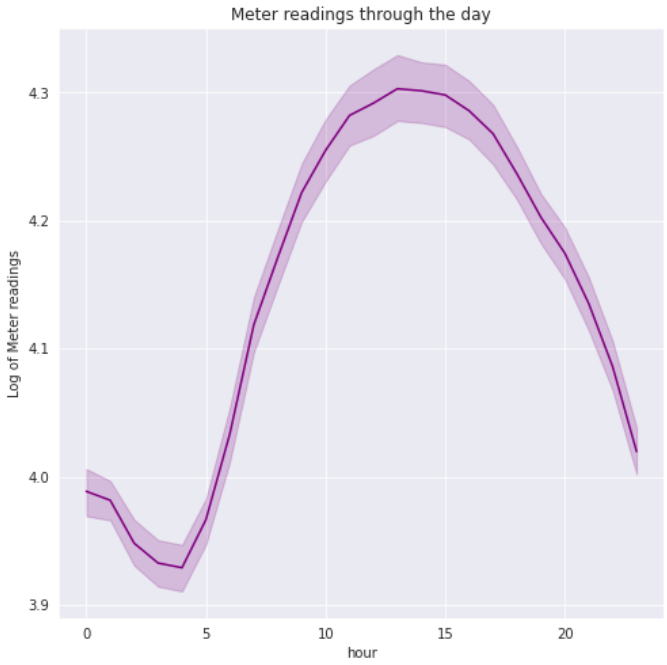
This paper presents the prediction of Energy consumption in a non-residential building using Machine Learning. For this we required various data that include the energy consumption and other factors required, we have collected the data for two years. We will talk about the collection of the data, screening of the data, weather data and meta- data of the buildings. We use this data for data mining tasks like prediction bench marking, prototyping, anomaly detection, energy analysis and building type classification.

Introduction

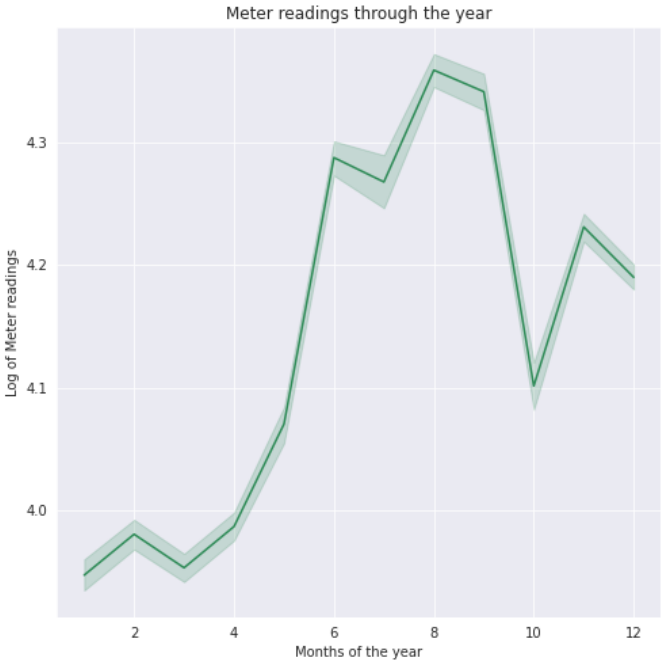
Reducing overall energy consumption and associated greenhouse gas emissions in the building sector is critical to meeting our long-term sustainability objectives. One bright spot Adoption of technology is a metering infrastructure for energy. which is widely used throughout the world The most recent From interval to continuous data collection, there has been a paradigm shift in the collection of building energy data. metres, also known as data loggers, have produced unprecedented results. a lot of information Data-driven analysis has been used in this study. provided useful insights into the built environment Among various applications, prediction of building energy use, and The use of machine learning models for forecasting has been a success.

Data Exploration and Pre-processing

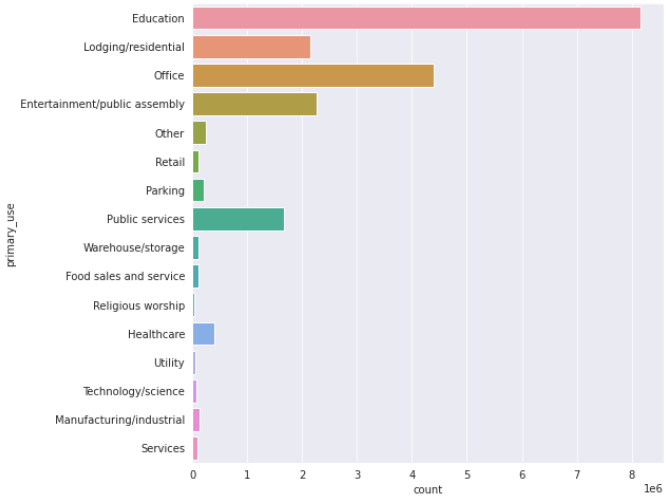
The dataset had over 20 features. The target value was meter reading.



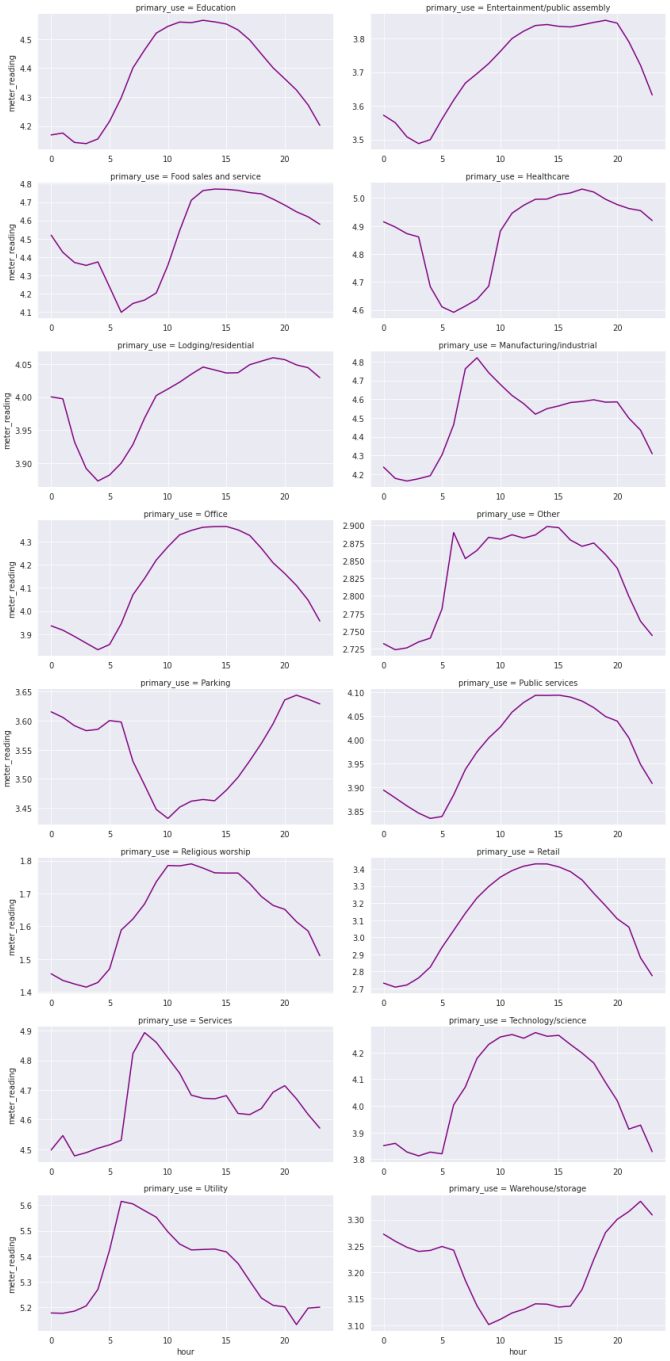
The above plot shows that the metre readings drop around 4 a.m. and then rise again after 6 a.m. The hourly usage peaks in the afternoon around 3 p.m. This behaviour is obvious because most institutes, industries, and other buildings are fully operational during this time.



The average monthly consumption is low in the early months and then rises after April with the arrival of spring. It peaks in August and then begins to fall again in the autumn and winter months.



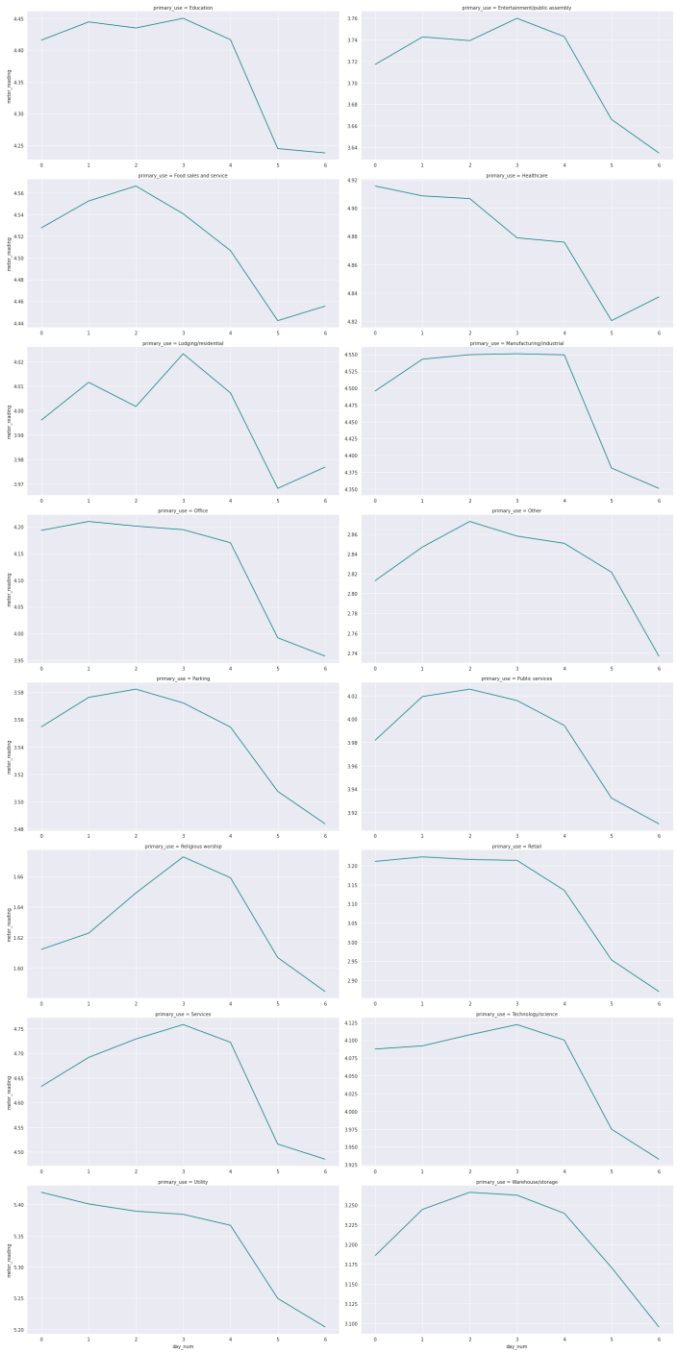
The educational institutes have the most readings, followed by the office, entertainment, and lodging/residential buildings.



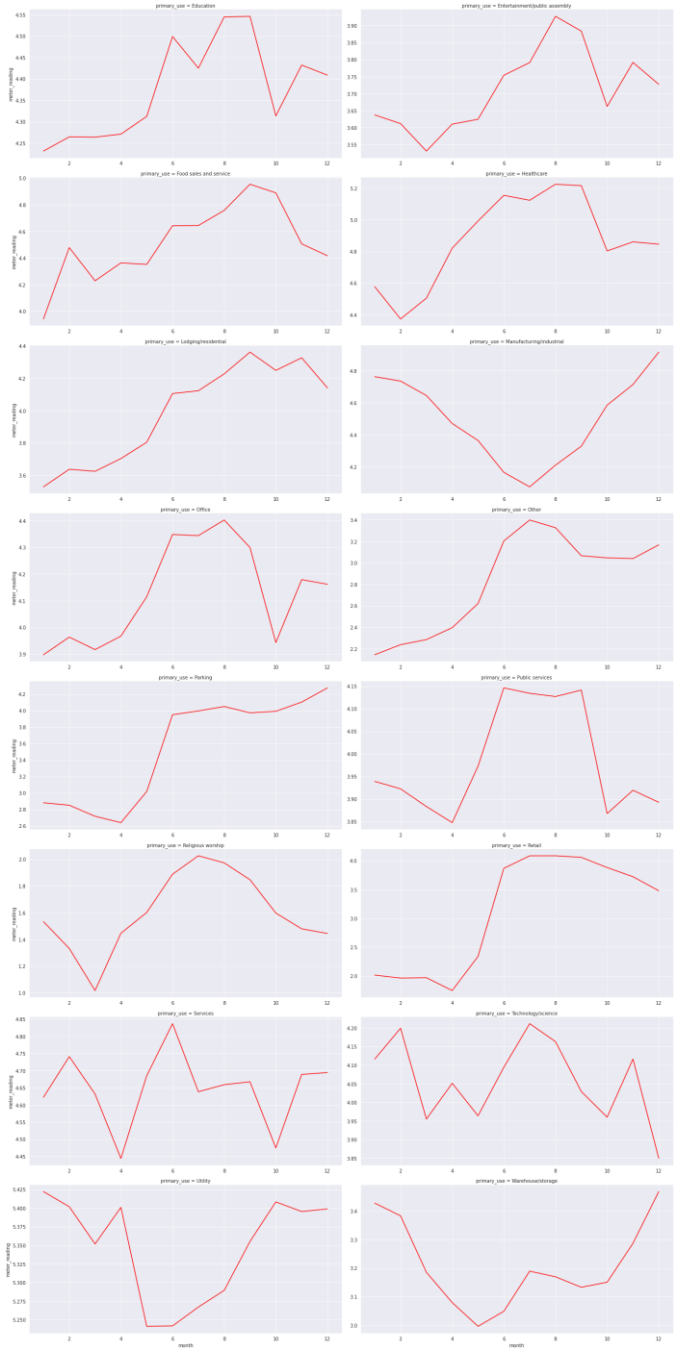
- • The peak usage for educational institutes is from 9 a.m. to 3 p.m., after which the usage drops. This is explained by the fact that most educational institutions are only open during the day.
- • Consumption of entertainment and food sales buildings increases after 10 a.m. and remains stable until 8 p.m.
- • Usage in healthcare and residential buildings is similar in that it peaks around 4 p.m. and then drops in the evening.
- • Office buildings, retail, technology, and public service buildings also consume the

majority of energy during the day, primarily from 10 a.m. to 4 p.m.

- • The parking lots are used the least in the morning and the most after 3 p.m.



We have plotted the average daily usage for each of the primary usage categories in the figure above. According to the plots above, average consumption is lowest on Sundays for all types of buildings except food sales and entertainment. On Saturdays, they have the lowest consumption.

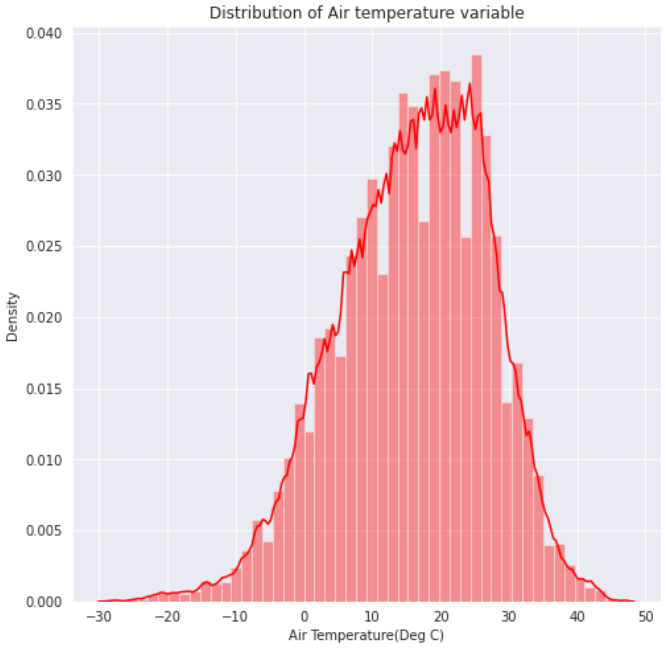


We've shown the average hourly consumption for each of the key usage categories in the graph above.

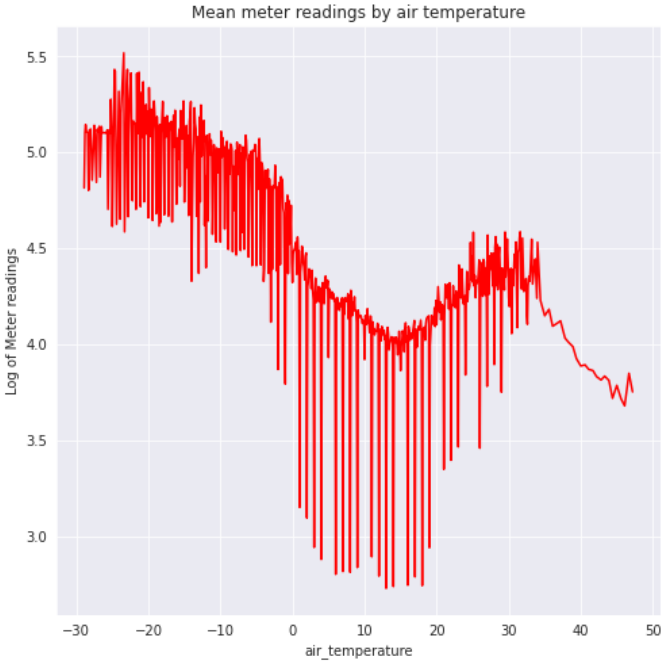
• August and September are the busiest months for educational establishments. During the first few months of the year, utilisation is minimal.

• The consumption of entertainment and food sales buildings peaks in August, with a minimum in March and a high in August.

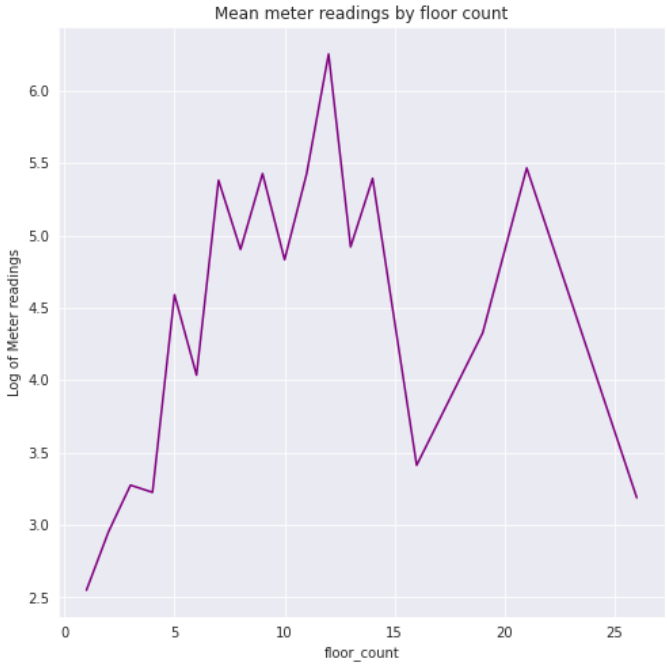
• There is a distinct downward trend in manufacturing, which reaches a low point in July before increasing again.



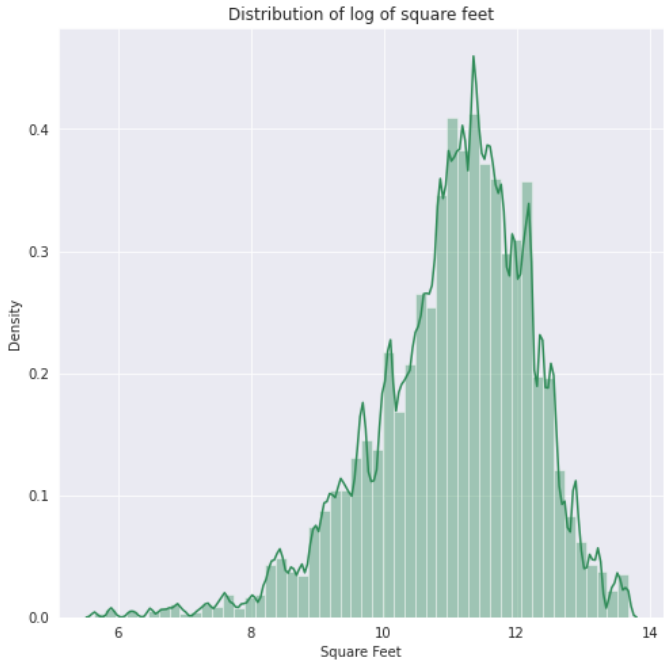
The temperature in the air roughly follows a normal distribution. The maximum observations occur between 0 and 30 degrees Celsius.



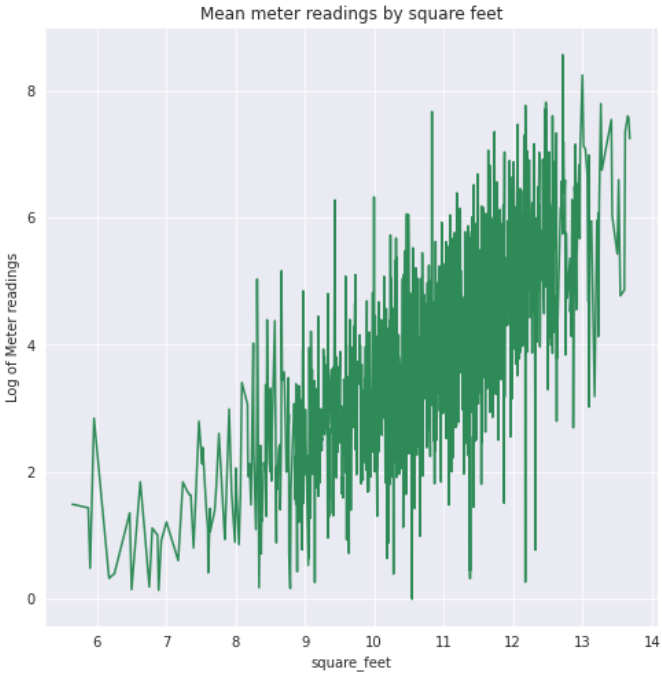
When the temperature is negative, the average metre readings are high; once the temperature begins to rise, the metre consumption decreases. It rises once more when the temperature exceeds 15 degrees Celsius.



When the floor count exceeds 10, the metre reading rises and then falls. This could be due to a variety of factors, such as not using all of the floors at the same time.



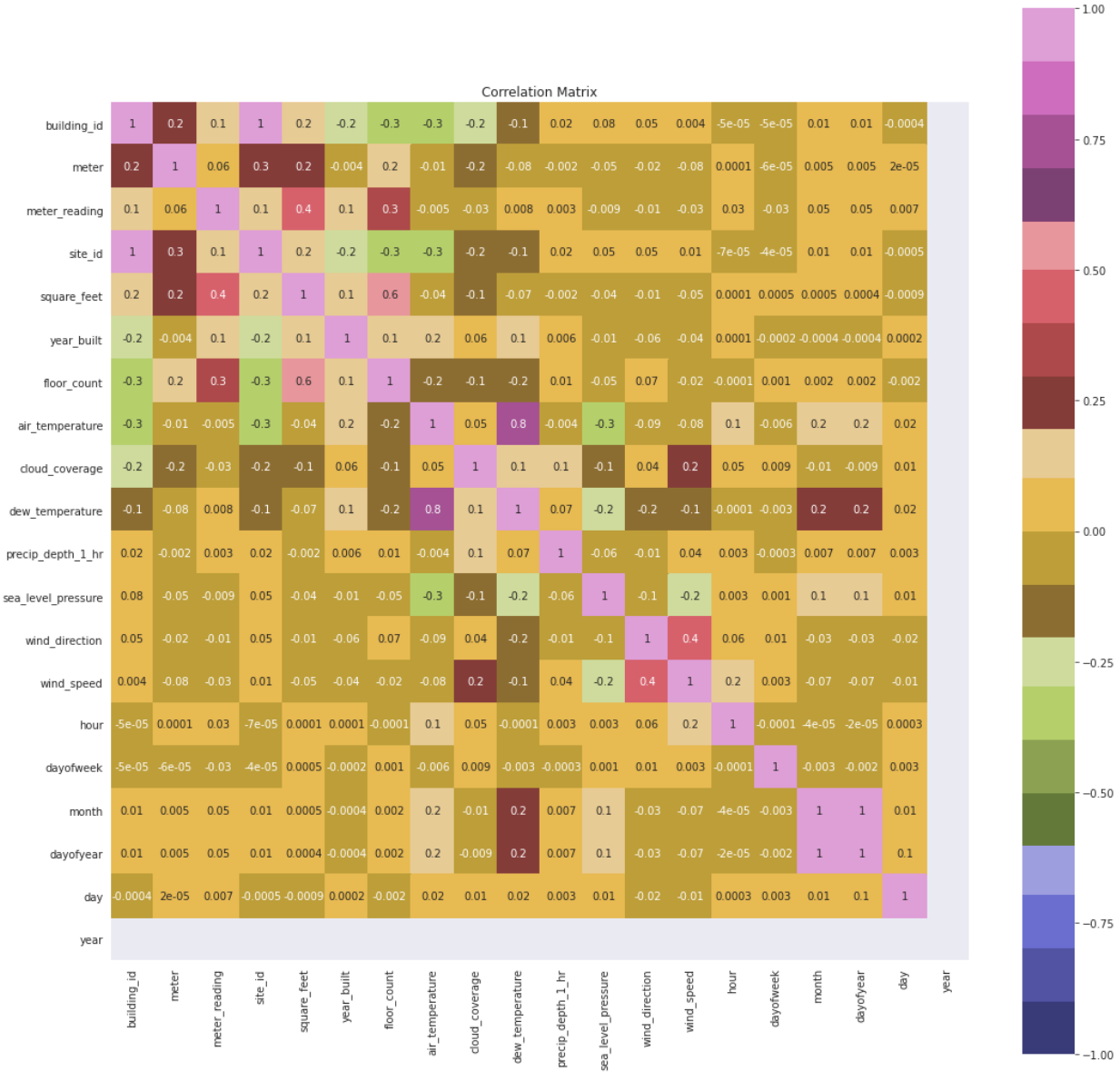
After log transformations, it looks better. We'll plot the mean energy consumption plot with square feet and look for trends.



grows.

The correlation matrix of the features available has been shown. The target variable has a substantial positive association with the square feet and floor count features. As we all know, the size of a structure has a significant impact on how much energy it consumes. The year a building was built also has a minor positive link, as older buildings use more energy than newer ones.

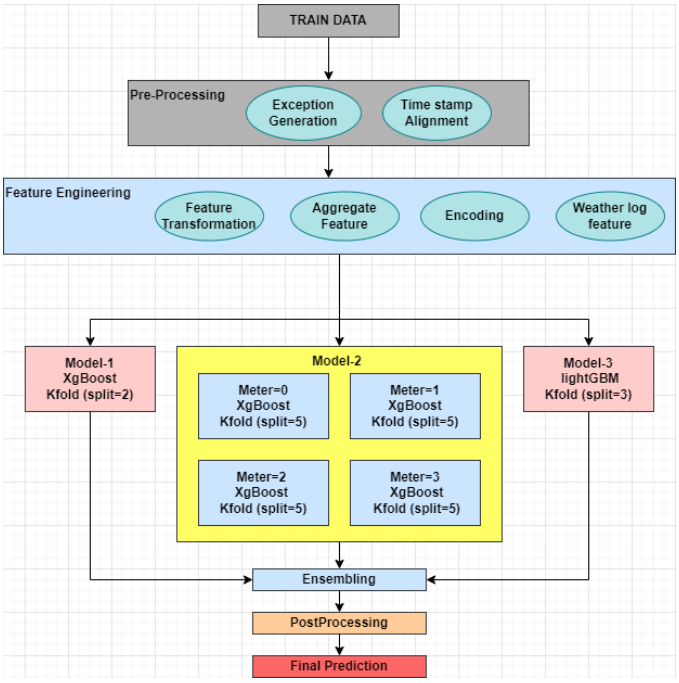
The two have a clear good relationship. The metre consumption increases as the size of the structure



Feature Engineering

- We handle the missing values by filling the feature mean value in the null except for the cloud coverage, sea level pressure and depth feature, for them we used median.
- Added new features- holidays. This feature will help us understand sudden drop in levels of energy usage.
- We split the date time feature in day, month, year and time features. This helps us to study the data on basis of groups.

Machine Learning Architecture



We used a custom architecture to find the final predictions.

Final Predictions = Model1 X (0.3) + Model2 X (0.4) +Model3 X (0.3)

Model 1 -> XgBoost

Model 2-> XgBoost with 4 boosting models

Model 3-> LightGBM

Custom architecture gave us a **RMSLE** score of 0.98.

Conclusion

From Data set of energy meters from the non-residential buildings with the range of two full years, we could able to extract some of the below insights from different plots using machine learning

As the building ages older, it tends to consume more energy than newer ones due to lack of energy efficient equipment or technology used in the building. Similarly, size of the building is directly proportional to meter consumption in the building. As the size of the building increases, the meter consumption also increases and vice-versa. In a year, August has peak average monthly consumption. We have minimum consumption in the initial months and then rises after April in the onset of spring. Relating Average meter reading to the temperature, It is high when the temperature is negative once it starts increasing, the meter consumption reduces. It again increases when the temperature rises above 15 deg C.

Having a glance at the type of building sector, Educational institutes have the peak usage in August and September. Food sale buildings has peak usage in August and minimal usage in March on consumption for entertainment basis. On a normal day consumption drops to a minimum in the early hours and then rises after 6am. It hits the peak in the afternoon and then drops after 7pm.

We can reduce the energy consumption in a building using above predictions by establishing and maintaining effective energy management systems for monitoring and controlling energy use in large public and commercial buildings