**1. A brief on the approach used to solve the problem:**

The problem statement provide was a time-series problem. First I want to know the data. For this I start to find out trends in energy consumption by plotting it on graph. Then I go for feature creation that will be used for further analysis. After analyzing all created features and their relations I select most appropriate ones. Then I started building ML Models for prediction. To cover the basics and test the model. I initially started with a Machine Learning approach. I experimented with the most widely used algorithms and found XGBoost Regression and RandomForest Regression to give the highest accuracy. However, when tested against the ground truth (test data), it performed poorly. Correspondingly, I moved with XGBoost model to tune hyper parameter. Therefore, I went ahead with the said model to test which of the hyperparameters gave the highest score as per the ground truth. This feedback system allowed me to me to obtain better ‘future covariates’ for my model and optimize other hyperparameters to create a better model with each trail. However this model is not giving accurate results but I continued with this for now (We can use DL model for further improvement). Finally, I scripted the entire process in a modular fashion to create a pipeline that could be deployed and automated for future use.

**2. Which Data-preprocessing / Feature Engineering ideas really worked? How did you discover them?**

ML Approach: In this ML approach, I initially set index as datetime converted from ‘object’ datatype to ‘datetime’ datatype. Then I look for missing data and I found that only 2% of data has null values in energy column, so I decided to drop them. Then I generated feature columns namely ‘year’, ‘month’, ‘day’, ‘quarter’, ‘day of week’, ‘day of month’ and ‘hour’ from the new datetime index. Following this, I created new feature named as ‘Holidays’ which will include whether that date is holiday or not as per Indian calendar. Weekend also marked as holidays because most of industry has weekly off on these days. Then I started visualizing data, energy demand at various time frame. Then I get to know about seasonal trend with linear growth is followed by energy demand. So that I decided to which ML model we can use for prediction. Then I found correlation between all features and deleted unrelated columns. After that, I went ahead with test, train, and validation splitting closely followed by standard scaling to bring all the feature ranges to fit the common scale.