

Report

Household Electricity Consumption

Theory :-

With the recent increase in smart meters across the residential sectors, we have large publicly available datasets. With such data, the power consumption of individual households can be tracked in almost real-time.

Such prediction can help power companies regulate their supply; also, the consumer can use this information to make better decisions both financially and environment-consciously.

In this project, we address this challenge by trying 5 different Machine learning Algorithms to do a comparative analysis to see which approach works best.

Steps:-

We first collect the data.

The type of our data is secondary.

➤ Data Description:-

- 1)The data is collected from UCI website, provided by senior researchers from France
- 2)Data involves more than 2 million records
- 3)Data consists of 47 months ranging from December 2006 to November 2010

As we see in the study that there are total 9 parameters

➤ Further Analysis:-

- 1)After collection of data we import library which are required for model fitting.
- 2)Then we loaded the collected data in the system.
- 3) We drew sample from loaded data set as the data is too large
- 4)The most important task of data cleaning was done after the sample from dataset has drawn.
- 5)After cleaning of data we checked normality of data by using different statistical models.
- 6)By checking normality of data we came to know that outliers are present in data,so our next task was handling of outliers.
- 7)Outliers are handled using IQR(Inter Quartile Range) Method.
- 8)After handling outliers our data was ready for model fitting.
- 9)We have fitted total five models :-

- Linear Regression
- Ridge Regression
- Lasso Regression
- Support vector Regression
- Decision Tree Regressor
- Random Forest Regressor

Report

Household Electricity Consumption

10) From the analysis we come to know that support vector regression model fits best to the given data than rest of the models.
