Assignment 3

Name : Smitesh Nitin Patil Class : 1CSD1

Student Id : 22223696

For this machine learning task, a regression problem was given, here I have documented my efforts of creating two regression models for predicted the target variable electricity consumption.

Task 1: Select an appropriate ML package to use for this regression task. Your report should briefly introduce your chosen package and your reasons for selecting it. In the ML package, select two different regression algorithms that you will apply to the dataset to learn two different regression models. Here are some possible choices, though other appropriate algorithms are fine too: linear regression, multi-layer perceptron, decision trees, k-nearest neighbours, support vector machines, etc. In your report, include a brief, clear description of both algorithms. Ensure that you acknowledge all of your sources of information. [2 marks max.]

Answer: There are many open-source python packages readily available for developing the complete machine learning pipeline. From data pre-processing to model evaluation metrics. For the machine learning task, sci-kit learn package was used. Sci-kit learn has sub packages for many regression ML models like MLPRegressor, Linear Regression, Gradient-boosting regressor, etc. Also, it is relatively easy to evaluate your models as it has many evaluation metrics for regression models predefined in the library like RMSE(root mean squared error) and R2 error.

From the sklearn library, two regression models were used

1. K-nearest neighbours Regressor
2. Gradient Boosting Regressor

1. K-nearest neighbours : KNN is a supervised learning algorithm, KNN utilises distance measuring metrics to make predictions on samples.