

Site Energy Intensity Prediction

Description: According to a [report](#) issued by the International Energy Agency (IEA), the lifecycle of buildings from construction to demolition was responsible for 37% of global energy-related and process-related CO2 emissions in 2020. Yet it is possible to drastically reduce the energy consumption of buildings by a combination of easy-to-implement fixes and state-of-the-art strategies.

The dataset consists of building characteristics, weather data for the location of the building, as well as the energy usage for the building, and the given year, measured as Site Energy Usage Intensity (Site EUI). Each row in the data corresponds to a single building observed in a given year.

Source of dataset & data dictionary - [Click Here](#)

Problem Statement: You are provided with two datasets: (1) the `train_dataset` where the observed values of the Site EUI for each row are provided and (2) the `x_test` dataset the observed values of the Site EUI for each row are removed and provided separately in `y_test`. Your task is to predict the Site EUI for each row (using the complete training dataset), given the characteristics of the building and the weather data for the location of the building. Use the test sets for validation and testing.

The target variable is `site_eui` and the evaluation metric is RMSE score.