

Apartment Price Predictor using Location Data

Data Section

For the machine learning model, I will be using two open source datasets available on Kaggle. The two sets contain the following data:

- Apartment names
- Latitude and Longitudes
- Prices
- Area
- Unit Type

Apart from these, I will be using the Foursquare API to leverage location data. The data set contains property for buying, and would be targeted mostly by people with an intention to settle such as a family. The data leveraged from the API would focus more on their requirements. The following data is planned to be extracted:

- Transport facilities nearby and proximity to each such as Railway stations, Metro stations, Bus stops, Airport. For this parameter, I am considering the distance to the closest of each.
- Daily requirements for running the house: Frequency of grocery stores, salons, hospitals and other medical services within a particular radius
- Presence of Educational institutions nearby for students in the house.
- Recreational Areas nearby for relaxing: Frequency of Restaurants, Cinemas, Parks.

The data sets used can be found here: <https://www.kaggle.com/sangarshanan/bangalore-apartments-dataset>

Assumption:

I am treating each apartment as an independent entity and ignoring amenities provided by the builder within the houses. The model will be generalized based on number of BHK and location data.