# Team NoCodeDevelopers

Varun Srinivas
Rishiraj Datta
Anirudh K

# What is the issue?

- House Price Prediction has been a topic of interest for every household for a very long time
- Many conversations end up with remorse, as people feel like they overpaid for a particular property
- Often times there are brokers and other middle-men who muddle our valuation
- It is thus essential to work on an unbiased estimator of prices based on various economic factors.

# Our Model

- After researching through various property broker sites and talking to adults involved in the real estate industry, we came up with a list of factors on which the price of a property depends upon:
  - Location
  - Sq. Ft
  - BHK
  - No. of Hospitals Nearby
  - No. of Parks Nearby
  - No. of Schools Nearby
  - No. of houses nearby(to determine whether it is a residential area or not)
  - No. of Stores Nearby
  - No. of Malls Nearby
  - No. of Metro Stations Nearby
  - No. of restaurants nearby

# Alternate Models and Reason for Rejection

- Inflation and Interest rate time series data - Flow variable vs stock variable issue
  - Not a viable combination for learning algorithms
- Unsupervised Learning algorithms like K-means clustering or AdaBoost
  - need to have a look at the current prices, only then, we can get close estimates
- Older data for geolocation analysis
  - API usage issue  + issue of redundancy

# Model Significance

- When a person wants to buy a house, they want to look for the availability of hospitals for the wellbeing of their parents/old people, and schools for their children
- More number of restaurants, malls and metro stations in an area means that it is a developing area with a lot of commercial spots.
- As a proxy for future development, we assume places close to maps and restaurants will see greater economic growth and hence use these parameters for predicting prices.
- We define a radius for different parameters
- For example, a person would need a hospital or school closer to their house, than a restaurant.

# Tech Stack and APIs used

- Backend:
  - Google Colab, Python, ScikitLearn, Pandas, Matplotlib, HERE Search - Forward Geocoder API, HERE Search - Browse API, HERE Maps API,HERE Autosuggest API.
- Frontend:
  - HTML, CSS, Javascript

# Working

- Using data from Magicbricks, we get the Location, Sq.Ft and BHK data. The location is in the form of text address, and we aim to get the coordinates of that location using HERE Search - Forward Geocoder API.
- With the help of HERE Search - Browse API, we get the different attributes like Hospitals, Parks etc.
- We define a radius for different parameters.
- For example, a person would need a hospital or school closer to their house, than a restaurant.

# Future Plans

- Use sentimental analysis of different areas, to detect people's sentiment about posh areas, high crime areas from the news or any other data source. We will be leveraging NLP to get these parameters and add it to our model to improve the performance.
- We will use temporal data to work with inflation data, interest rates and time series prediction of data. For this we can use LSTMs.
- We could also work with highlighting a particular area of interest
- Status inclusion: Ready to move, Possession by date ⇒ to include the effect of time value of money and opportunity cost