ME-781 Course Project: Interim Report

Housing Price Prediction using Metropolitan Cities' Real Estate data

Compiled by Group17-

Kaustabh Patel	190100063
Lakshya Chaplot	19D100010
Niharika Maheshwary	190100141
Siri Verma	180050052

Problem Objective and Introduction

HOUSING MAGICIAN

- We are a Mumbai based startup aspiring to lead in providing reliable and cheap housing and real estate prices across different metropolitan cities of India.
- Our speedy, efficient and cross-platform app will help deliver the most appropriate housing price to the user based on their customized filters.
- In the post-pandemic era of risky brokers, distrustful real estate offices and economic uncertainty, we plan to enter the B2C market for boosting the residential housing sales in India.
- Our current objective is to provide housing price predictions in major metropolitan cities of India, but plan on expanding our reach to other cities in the near future.

Problem Definition

Customer Requirement: Quick, easy, hassle-free way to deal with purchasing real estate in metropolitan cities which helps in determining housing prices for potential buyers and sellers.

Market Survey: 99acres (price estimator), Land Value Calculator (estimating price of the land), MagicBricks PropWorth (estimate the value of a property).

USP: Providing the user with a heat map to help them understand the prices and status of the property according to their needs and demands. It is easy to understand as well as user friendly to help citizens make the best decision with regards to property prices.

Protection of USP: Copywriting of entire code and novel solution developed along with patenting. Encryption of the code is another way to protect the USP

Barrier to Entry:

- Copyright and patent issues by other organizations and companies which lead to some roadblock
- Confidentiality clauses when trying to determine price points for various houses and properties

Operational Structure:

- Various criteria and demands uploaded by the user
- Prediction model uses different variables to get the best response
- From this model lets us know the price and status for different houses and properties in the selected vicinity with the demands mentioned

Technology Landscape Management

Patents:

- https://patents.google.com/patent/WO2010128924A1/en
- https://patents.justia.com/patent/7509261

Literature:

- https://arxiv.org/abs/1707.04868
- https://arxiv.org/abs/1808.02547
- https://www.sciencedirect.com/science/article/pii/S1877050920316318
- https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3565512
- https://www.emerald.com/insight/content/doi/10.1108/IJHMA-11-2018-0095/full/html
- https://ieeexplore.ieee.org/abstract/document/7603227

Open Libraries:

- Numpy, pandas, keras, tensorflow, scikit-learn, matplotlib, glob

Project Plan-GANNT Chart

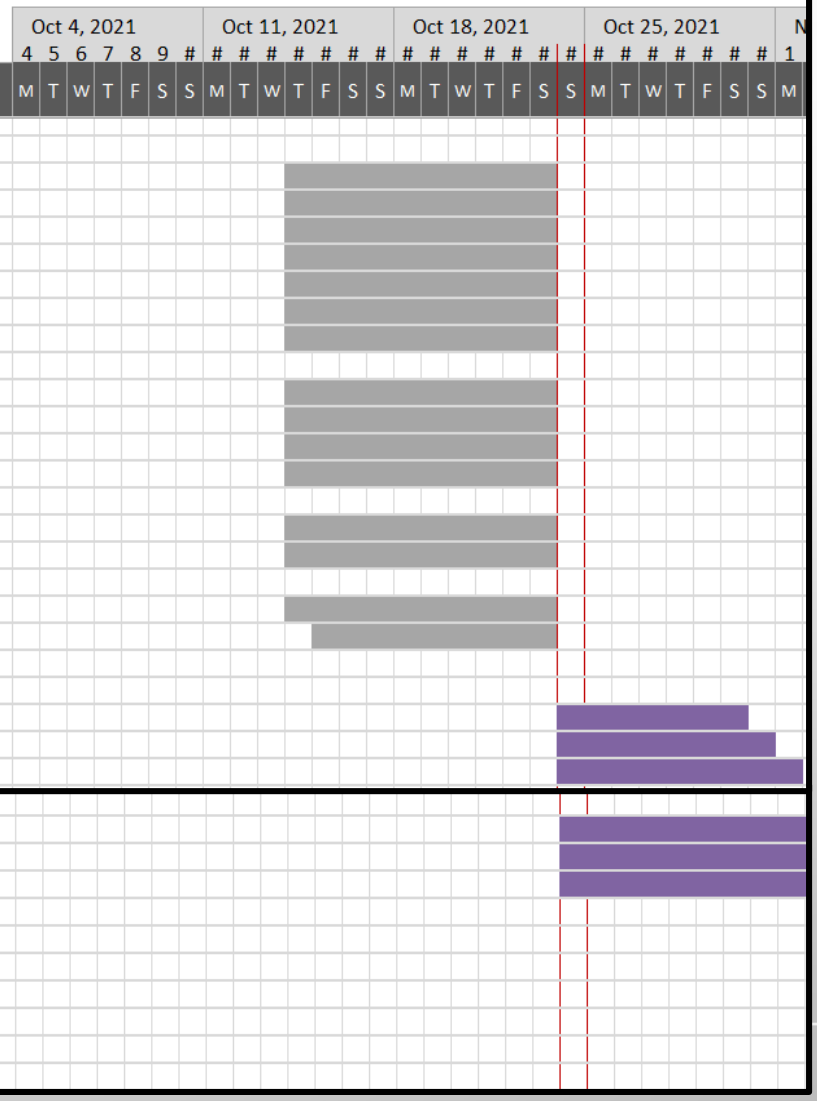
Housing Magician

Group-**17**, ME781 Course Project, IIT-Bombay

Housing price predictor app

Project Start:	Wed, 13-10-2021
Today:	Sun, 24-10-2021
Display Week:	0

Sr.No	TASK	ASSIGNED TO	PROGRESS	START	END
1	**Problem Definition & Objective**				
1.1	Customer (end-user) requirement	Niharika	100%	14-Oct-21	23-Oct-21
1.2	Market survey (for exisitng solns)	Niharika	100%	14-Oct-21	23-Oct-21
1.3	Key Differentiator	Niharika	100%	14-Oct-21	23-Oct-21
1.4	Unique Selling Point (USP)	Niharika	100%	14-Oct-21	23-Oct-21
1.5	Protection of USP	Niharika	100%	14-Oct-21	23-Oct-21
1.6	Barrier to Market Entry	Niharika	100%	14-Oct-21	23-Oct-21
1.7	Business Case (setup cost and capital)	Niharika	100%	14-Oct-21	23-Oct-21
2	**Technology Lanscape Management**				
2.1	Patents	Kaustabh	100%	14-Oct-21	23-Oct-21
2.2	Published literature	Kaustabh	100%	14-Oct-21	23-Oct-21
2.3	Open libraries	Kaustabh	100%	14-Oct-21	23-Oct-21
2.4	Proprietary libraries	Kaustabh	100%	14-Oct-21	23-Oct-21
3	**Project Planning**				
3.1	GANNT chart	Lakshya	100%	14-Oct-21	23-Oct-21
3.2	RASIC chart	Lakshya	100%	14-Oct-21	23-Oct-21
4	**Conceptual Design**				
4.1	High level activity object diagrams	Siri Verma	100%	14-Oct-21	23-Oct-21
4.2	Choice of language, OS, Data set	Siri Verma	100%	15-Oct-21	23-Oct-21
4.3	ML model selection	Lakshya	0%		
5	**Code Development-1**				
5.1	Coding convention practices	Siri Verma	0%	24-Oct-21	30-Oct-21
5.2	Basic user interface	Siri Verma	0%	24-Oct-21	31-Oct-21
5.3	Output visualization	Lakshya	0%	24-Oct-21	1-Nov-21
6	**Code Development-2**				
6.1	Error handling	Lakshya	0%	24-Oct-21	17-Nov-21
6.2	Auto document generation	Siri Verma	0%	24-Oct-21	18-Nov-21
6.3	Code testing	Lakshya	0%	24-Oct-21	19-Nov-21
7	**Product Marketing**				
7.1	Brochure development	Niharika	0%	20-Nov-21	25-Nov-21
7.2	Marketing presentation	Niharika	0%	21-Nov-21	26-Nov-21
7.3	Marketing audio-video of 1min	Siri Verma	0%	22-Nov-21	27-Nov-21
8	**Product Demonstration**				
8.1	Demonstration presentation	Kaustabh	0%	21-Nov-21	27-Nov-21
8.2	User manual	Lakshya	0%	22-Nov-21	28-Nov-21



https://tinyurl.com/GANNT-17

The 0% completion indicates the work hasn't been started yet.

Project Plan-RASIC Chart

- **Responsible:** The person who is ultimately responsible for delivering the project and/or task successfully.
- **Accountable:** The person who has ultimate accountability and approval authority; they review and assure quality and are the person to whom "R" is accountable.
- **Supporting:** The team or person(s) supporting the "real" work with resources, time or other material benefit. They are committed to its completion.
- **Informed:** Those who provide input and must be informed of results or actions taken but are not involved in final decision-making .
- **Consulted:** Those who provide valuable input into product design or establish quality review criteria. Their buy-in is important for successful implementation.

https://tinyurl.com/RASICgroup17

	Project Deliverable (or Activity) \ Role	Project Members			
		Kaustabh	Lakshya	Niharika	Siri Verma
1.1	Customer (end-user) requirement	A	I	R	I
1.2	Market survey (for exisitng solns)	A	I	R	I
1.3	Key Differentiator	A	C	R	I
1.4	Unique Selling Point (USP)	I	I	R	A
1.5	Protection of USP	I	I	R	A
1.6	Barrier to Market Entry	I	I	R	A
1.7	Business Case (setup cost and capital)	C	A	R	I
2.1	Patents	R	I	A	I
2.2	Published literature	R	I	A	I
2.3	Open libraries	R	I	A	I
2.4	Proprietary libraries	R	I	I	I
3.1	GANNT chart	I	R	A	C
3.2	RASIC chart	C	R	A	I
4.1	High level activity object diagrams	A	I	I	R
4.2	Choice of language, OS, Data set	A	I	C	R
4.3	ML model selection	C	R	I	A
5.1	Coding convention practices	I	C	A	R
5.2	Basic user interface	I	A	I	R
5.3	Output visualization	I	R	I	A
6.1	Error handling	I	R	I	A
6.2	Auto document generation	I	A	I	R
6.3	Code testing	I	R	A	I
7.1	Brochure development	A	I	R	I
7.2	Marketing presentation	I	I	R	A
7.3	Marketing audio-video of 1min	I	A	I	R
8.1	Demonstration presentation	R	A	C	I
8.2	User manual	I	R	I	A

Conceptual Design

ML Model Selection:

- We have modelled "Predicting Housing Prices" as a regression problem as it allows us to give more insights to the relation between the attributes and the house price in a better way.
- After applying the linear regression model and interpreting the data, we plan to use regularization models like Ridge and Lasso Regression, and other ML models like polynomial regression and KNN algorithms, to predict the output. We will compare and analyse the models based on their accuracy, and choose the best model to predict prices in our start-up project.

ML Dataset Selection:

- https://www.kaggle.com/anmolkumar/house-price-prediction-challenge
- https://www.kaggle.com/ruchi798/housing-prices-in-metropolitan-areas-of-india
- The data attributes considered to predict the price will include features of both the house/property as well as the features of the neighbourhood. Additional features like crime rates, water and electricity issues, and availability of schools, grocery marts and hospitals in the vicinity will also be added.
- Our project will work on data collected for the metropolitan cities - Bangalore, Chennai, Delhi, Hyderabad, Kolkata and Mumbai.

Choice of Language: Python

Operating System: Windows 10

Conceptual Design

High Level Activity Object Diagram: