Title of the Project

Team ID:24 **Supervisor: Prof.Bavrabi Ghosh**
Student Name: Ankan Ghatak **Roll No.: 83**
Student Name: Anubhab Chaudhuri **Roll No.: 103**

Introduction

- The real estate industry is influenced by various factors such as economic conditions, political stability, demographics, and technology, among others.
- The ability to predict real estate prices accurately can be useful for various stakeholders such as buyers, sellers, and investors.
- In the real estate industry, machine learning algorithms can be used to develop models for real estate price prediction.

Objectives and Challenges

- This research paper aims to develop and compare various machine learning models for real estate price prediction in three Indian cities - Chennai, Bengaluru, and Delhi.
- Some challenges are:

- Managing 3 different datasets with uncommon columns
- Removing Outleirs
- Dealing with Nan Values

Our Approach

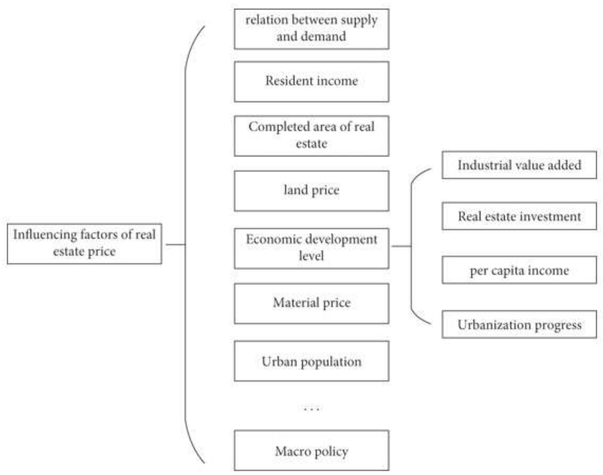


Fig. 1 Index Factors affecting real estate prizes

- Linear Regression: It assumes a linear relationship between the independent and dependent variables and aims to find the best fit line to predict the target variable.
- Decision Tree: It works by recursively partitioning the data into smaller subsets based on the features to predict the target variable.
- Lasso Regression: This model uses shrinkage. Shrinkage is where data values are shrunk towards a central point as the mean. The lasso procedure encourages simple, sparse models (i.e. models with fewer parameters).

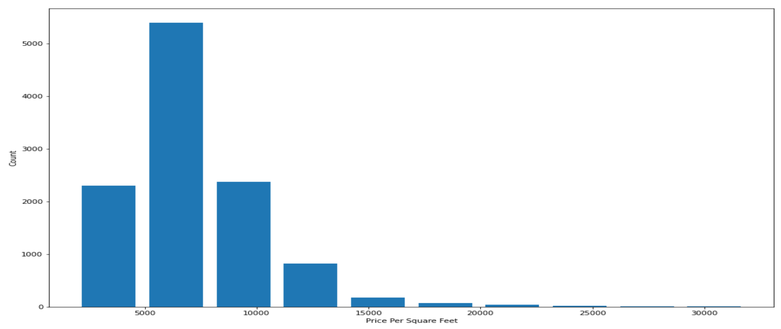
Results

```
[194]: predict_price('1st Phase JP Nagar',1000, 2, 2)
       c:\users\anubhab\appdata\local\programs\python\python39\lib\site-packages\
       nearRegression was fitted with feature names
         warnings.warn(
[194]: 81.80288347852803

[170]: predict_price('Indira Nagar',1000, 3, 3)
       c:\users\anubhab\appdata\local\programs\python\python39\lib\site-packages\
       nearRegression was fitted with feature names
         warnings.warn(
[170]: 178.2198177230117

[174]: predict_price('Karapakkam',1000, 3, 3)
       c:\users\anubhab\appdata\local\programs\python\python39\lib\site-packages\
       nearRegression was fitted with feature names
         warnings.warn(
[174]: 71.2640470677168

[177]: predict_price('Shivaji Nagar',1000, 3, 3)
       c:\users\anubhab\appdata\local\programs\python\python39\lib\site-packages\
       nearRegression was fitted with feature names
         warnings.warn(
[177]: 75.71646050647018
```



ALGORITHM	CHENNAI (X)	BENGALURU (Y)	DELHI (Z)	ALL TOGETHER (X+Y+Z)
LINEAR REGRESSION	0.845574	0.818354	0.897554	0.849693
DECISION TREE	0.741746	0.725728	0.799807	0.696853
LASSO REGRESSION	0.799029	0.687439	0.870960	0.623923

Discussion

- There are many factors affecting real estate prices, such as market supply and demand, wage income, and economic development level.
- There is a complex relationship between these influencing factors..
- The research uses the information gain method to extract the principal components.

Conclusion and Future Scope

- This research paper developed and compared various machine learning models for real estate price prediction in three Indian cities - Chennai, Bengaluru, and Delhi.
- The models were trained using various algorithms such as Linear Regression, Decision Tree and Lasso Regression.
- Future work can include incorporating more features such as the age of the property, amenities, and nearby infrastructure to improve the accuracy of the models.

References

1. Liu, G. (2022). Research on Prediction and Analysis of Real Estate Market Based on the Multiple Linear Regression Model. Scientific Programming, 2022, 1–8. https://doi.org/10.1155/2022/5750354
2. _Real Estate Price Prediction using Supervised Learning. (2022b, December 15). IEEE Conference Publication | IEEE Xplore. https://ieeexplore.ieee.org/abstract/document/10014818
3. _Zhao, Y. (2022, August 29). PATE: Property, Amenities, Traffic and Emotions Coming Together for Real Estate Price Prediction. arXiv.org. https://arxiv.org/abs/2209.05471
4. Tchuente, D., & Nyawa, S. (2021). Real estate price estimation in French cities using geocoding and machine learning. Annals of Operations Research, 308(1–2), 571–608. https://doi.org/10.1007/s10479-021-03932-5