Real Estate Price Prediction Project

Given By: Mr. Ajay

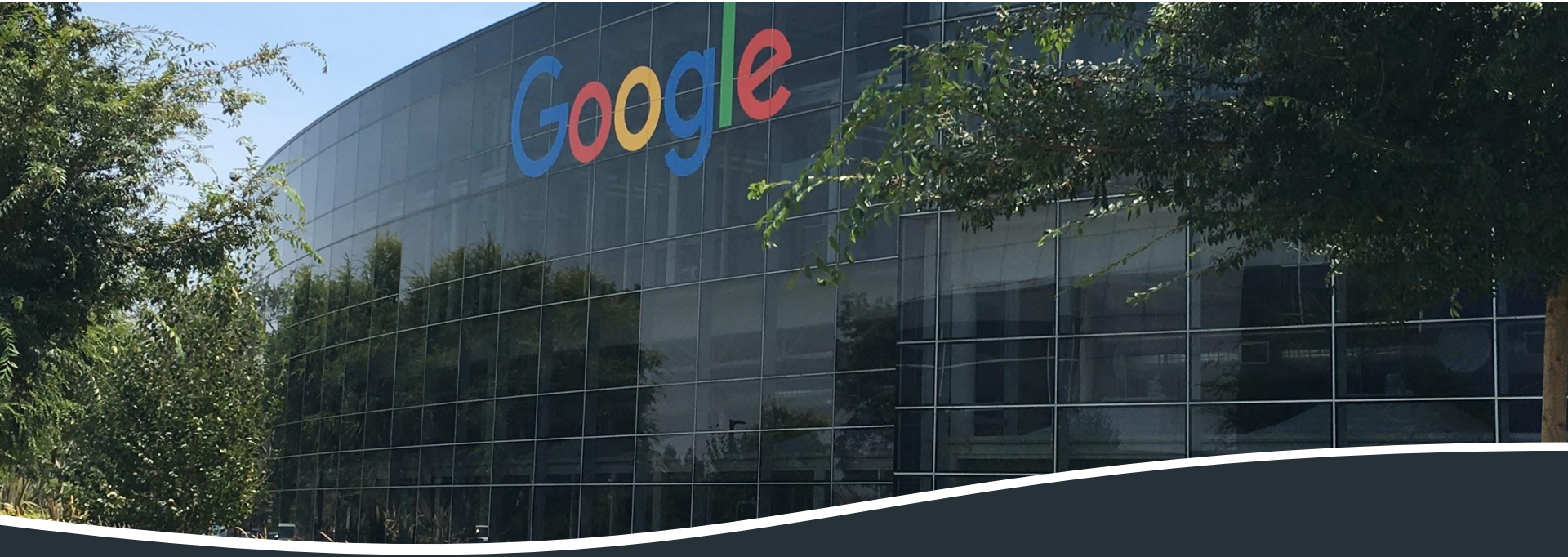
Bengaluru_House_Data

Abstract idea

*In this project ,I am experiment with a house prediction datasets , and to explore how machine learning algorithm can be used to find the pattern in data . We were expected to gain experience using a machine learning library and were expected to submit a report about datasets and the algorithm used . After performing the required tasks on a datasets of our choice , herein lies our final report .

Abstract idea

What are the things that a potential home buyer considers before purchasing a house? The location, the size of the property, vicinity to offices, schools, parks, restaurants, hospitals or the stereotypical white picket fence? What about the most important factor — the price?



Offices

Schools

Restaurants

Parks

Hospital


10
20
EURO

Abstract idea

Potential homeowner, over 9,000 apartment projects and flats for sale are available in the range of 42-52 lakh, followed by over 7,100 apartments that are in the 52-62 lakh budget segment, says a report by property website Makaan. According to the study, there are over 5,000 projects in the 15-25 lakh budget segment followed by those in the 34-43 lakh budget category.

Abstract idea

Buying a home, especially in a city like Bangalore, is a tricky choice. While the major factors are usually the same for all metros, there are others to be considered for the Silicon Valley of India. With its help millennial crowd, vibrant culture, great climate and a slew of job opportunities, it is difficult to ascertain the price of a house in Bengaluru.

DATASETS

	area_type	availability	location	size	society	total_sqft	bath	balcony	price
0	Super built-up Area	19-Dec	Electronic City Phase II	2 BHK	Coomee	1056	2.0	1.0	39.07
1	Plot Area	Ready To Move	Chikka Tirupathi	4 Bedroom	Theanmp	2600	5.0	3.0	120.00
2	Built-up Area	Ready To Move	Uttarahalli	3 BHK	NaN	1440	2.0	3.0	62.00
3	Super built-up Area	Ready To Move	Lingadheeranahalli	3 BHK	Soiewre	1521	3.0	1.0	95.00
4	Super built-up Area	Ready To Move	Kothanur	2 BHK	NaN	1200	2.0	1.0	51.00

```
[ ]  House_data.shape
```

```
(13320, 9)
```

Data Statistical Analysis

```
House_data.describe()
```

	bath	balcony	price
count	13247.000000	12711.000000	13320.000000
mean	2.692610	1.584376	112.565627
std	1.341458	0.817263	148.971674
min	1.000000	0.000000	8.000000
25%	2.000000	1.000000	50.000000
50%	2.000000	2.000000	72.000000
75%	3.000000	2.000000	120.000000
max	40.000000	3.000000	3600.000000

Null values

```
area_type          0
availability       0
location           1
size              16
society         5502
total_sqft         0
bath              73
balcony          609
price              0
dtype: int64
```

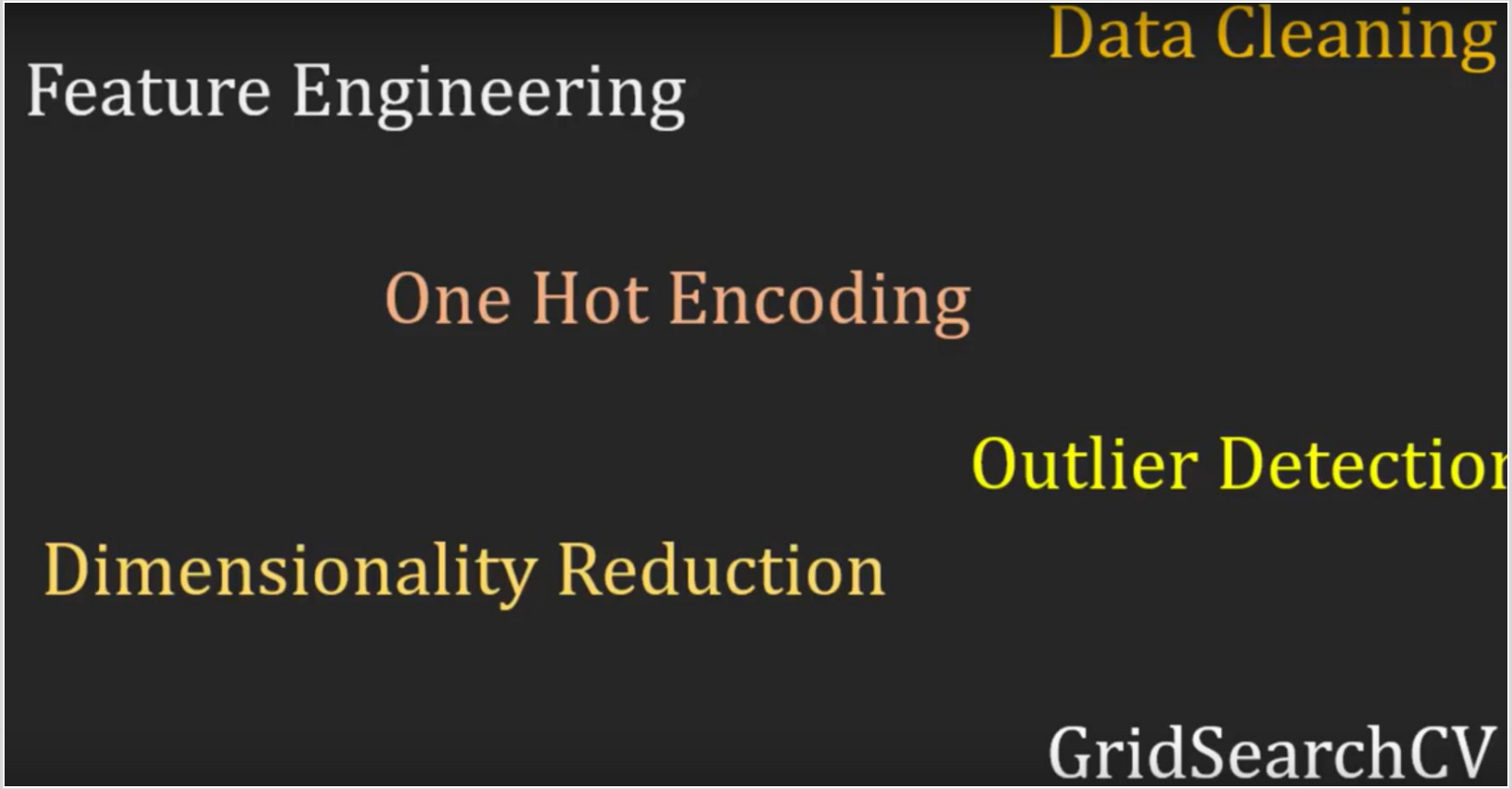
After Handling Null values

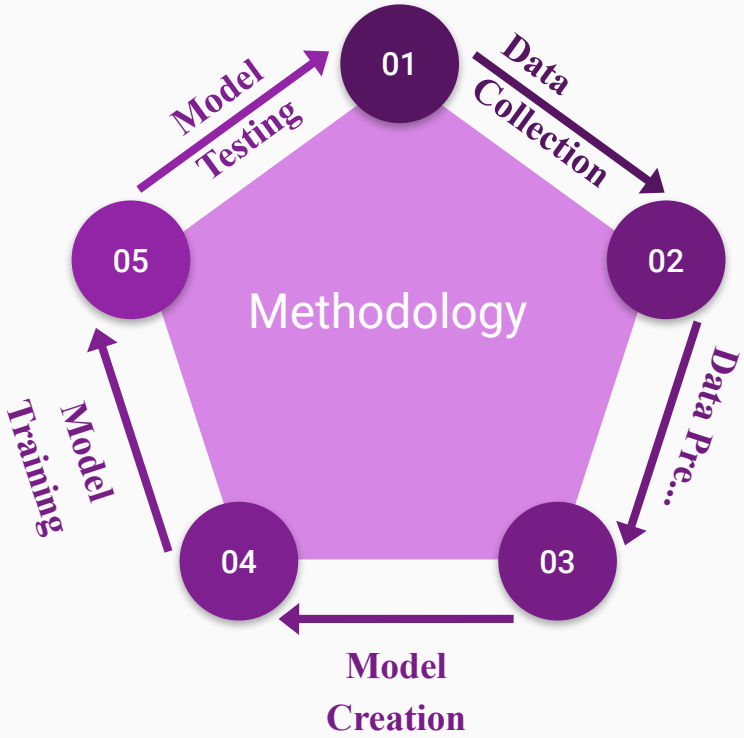
```
[ ]  X.isnull().sum()
```

```
location      0
size          0
total_sqft    0
bath          0
balcony       0
price         0
dtype: int64
```

Methodology

(I am using cool data science technique)

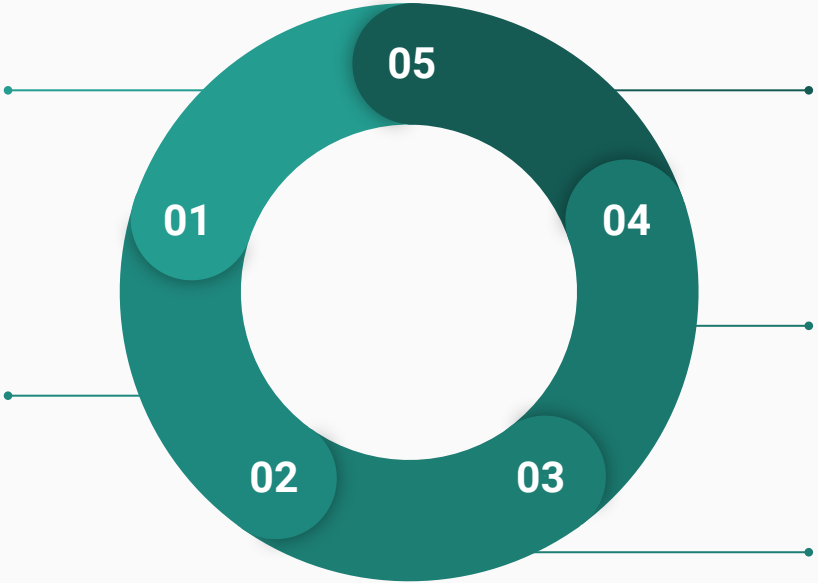

Data Cleaning
Feature Engineering
One Hot Encoding
Outlier Detection
Dimensionality Reduction
GridSearchCV


Methodology
01
Data Collection
02
Data Pre...
03
Model Creation
04
Model Training
05
Model Testing

Bengaluru_House_Data

Data Preprocessing

- Null value handling
- Categorical Data Handling
- Feature Selection(PCA)
- Tuning the hyper-parameters



Model Testing

20% DATA

MSE

Model Training

80% DATA

Model Creation

- Linear Regression
- Lasso Regression
- Decision_Tree Regression


THANK
YOU