

Business Problem - Predict the Price of Bangalore House

Using K Nearest Neighbor Regression - Supervised Machine Learning Algorithm

Load Libraries

```
In [ ]: import pandas as pd
```

Load Data

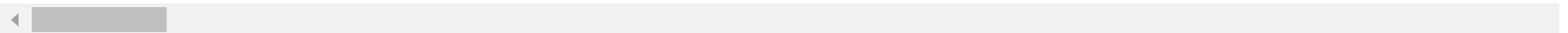
```
In [ ]: path = r"https://drive.google.com/uc?export=download&id=1xxDtrZKfuWQf1-6KA9XEd_eatitNPnkB"
df = pd.read_csv(path)
```

```
In [ ]: df.head()
```

Out[3]:

	bath	balcony	price	total_sqft_int	bhk	price_per_sqft	area_typeSuper built-up Area	area_typeBuilt- up Area	area_typePlot Area	avail
0	3.0	2.0	150.0	1672.0	3	8971.291866	1	0	0	
1	3.0	3.0	149.0	1750.0	3	8514.285714	0	1	0	
2	3.0	2.0	150.0	1750.0	3	8571.428571	1	0	0	
3	2.0	2.0	40.0	1250.0	2	3200.000000	1	0	0	
4	2.0	2.0	83.0	1200.0	2	6916.666667	0	0	1	

5 rows × 108 columns



Split Data

```
In [ ]: X = df.drop('price', axis=1)
        y = df['price']

        print('Shape of X = ', X.shape)
        print('Shape of y = ', y.shape)
```

```
Shape of X = (7120, 107)
Shape of y = (7120,)
```

```
In [ ]: from sklearn.model_selection import train_test_split
        X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=51)

        print('Shape of X_train = ', X_train.shape)
        print('Shape of y_train = ', y_train.shape)
        print('Shape of X_test = ', X_test.shape)
        print('Shape of y_test = ', y_test.shape)
```

```
Shape of X_train = (5696, 107)
Shape of y_train = (5696,)
Shape of X_test = (1424, 107)
Shape of y_test = (1424,)
```

##K Nearest Neighbor Regression - ML Model Training

```
In [ ]: from sklearn.neighbors import KNeighborsRegressor
```

```
In [ ]: regressor = KNeighborsRegressor(n_neighbors=5)
        regressor.fit(X_train, y_train)
```

```
Out[7]: KNeighborsRegressor(algorithm='auto', leaf_size=30, metric='minkowski',
                             metric_params=None, n_jobs=None, n_neighbors=5, p=2,
                             weights='uniform')
```

```
In [ ]: regressor.score(X_test, y_test)
```

```
Out[8]: 0.8858213777326956
```

Predict the value of Home

```
In [ ]: X_test.iloc[-1, :]
```

```
Out[10]: bath                2.000000  
         balcony            0.000000  
         total_sqft_int      1566.000000  
         bhk                2.000000  
         price_per_sqft      11494.252874  
         ...  
         location_Hosur Road  0.000000  
         location_Horamavu Banaswadi 0.000000  
         location_Domlur     0.000000  
         location_Mahadevpura 0.000000  
         location_Tumkur Road 0.000000  
         Name: 43, Length: 107, dtype: float64
```

```
In [ ]: regressor.predict([X_test.iloc[-1, :]])
```

```
Out[11]: array([184.4])
```

```
In [ ]: y_test.iloc[-1]
```

```
Out[12]: 180.0
```

```
In [ ]: y_pred = regressor.predict(X_test)  
         y_pred
```

```
Out[13]: array([ 77.378,  39.808, 120.2 , ...,  33.405,  63.   , 184.4  ])
```

```
In [ ]: y_test
```

```
Out[14]: 2435      80.00
          3113      40.00
          426     120.00
          1124      79.00
          1161      45.00
          ...
          2078      28.34
          6855      84.00
          4381      32.00
          3862      63.00
           43     180.00
          Name: price, Length: 1424, dtype: float64
```

Ab milenge next tutorial me,Tab tak ke liye SIKHATE SIKHATE kuch IMPLEMENT karte raho, Thank You.....:-)