HOUSING PRICE PREDICTION ML PROJECT

1.1 About Project

I have created a Machine Learning Regression Model Using Python. I will be analysing a house price predication dataset for finding out the price of a house on different parameters. I will do Exploratory data Analysis, split the training and testing data, Model Evaluation and Predictions.

1.2 Requirement of Project

1.2.1 Hardware Requirement

1.Laptop

1.2.2 Software Requirement

Operating System : Windows 10 an application

Language   : Python 3.8

1.3 Introduction

Machine learning is an application of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed. **Machine learning** focuses on the development of computer programs that can access data and use it learn for themselves. Machine learning is of three types:

1. **Supervised learning:**
    It is the **machine learning** task of **learning** a function that maps an input to an output based on example input-output pairs. A **supervised learning** algorithm analyzes the training data and produces an inferred function, which can be used for mapping new examples.

2. **Unsupervised learning:**
    It is a type of **machine learning** that looks for previously undetected patterns in a data set with no pre-existing labels and with a minimum of human supervision.

3. **Reinforcement learning** (RL):

It is an area of **machine learning** concerned with how software agents ought to take actions in an environment in order to maximize the notion of cumulative reward. **Reinforcement learning** is one of three **basic machine learning** paradigms, alongside supervised **learning** and unsupervised **learning**.

Linear regression: It is a type of regression analysis where the number of independent variables is one and there is a linear relationship between the independent(x) and dependent(y) variable. The red line in the above graph is referred to as the best fit straight line. It comes under the supervised learning.
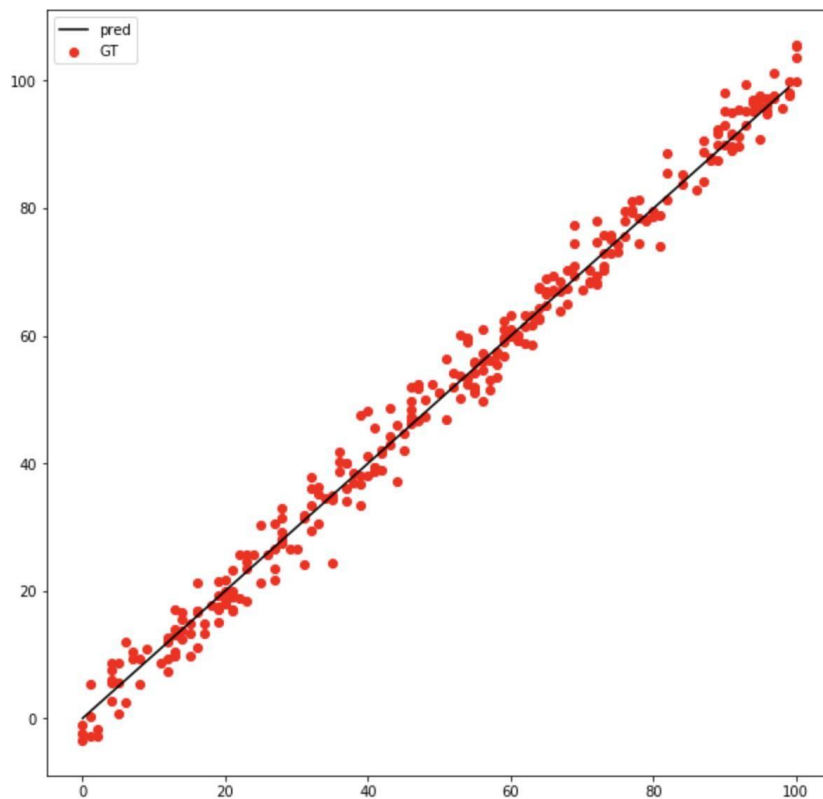
1.3.1 Working of the Project

We will be importing Scikit Learn, Pandas, Seaborn, Matplotlib and Numpy.

```
import pandas as pd import

numpy as np import seaborn as
```

```
sns import matplotlib.pyplot as

plt
```

Importing Data and Checking out

```
HouseDF = pd.read_csv(r' C:\Users\FARATH\Downloads\Linear-
Regression-Model-for-House-Price-Prediction-master\
USA_Housing.csv')

HouseDF.head()
```

X and y list

```
X = HouseDF[['Avg. Area Income', 'Avg. Area House Age', 'Avg. Area
Number of Rooms', 'Avg. Area Number of Bedrooms', 'Area
Population']]

 y =
HouseDF['Price']
```

Split Data into Train, Test

```
from sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test = train_test_split(X, y,
test_size=0.4, random_state=101)
```

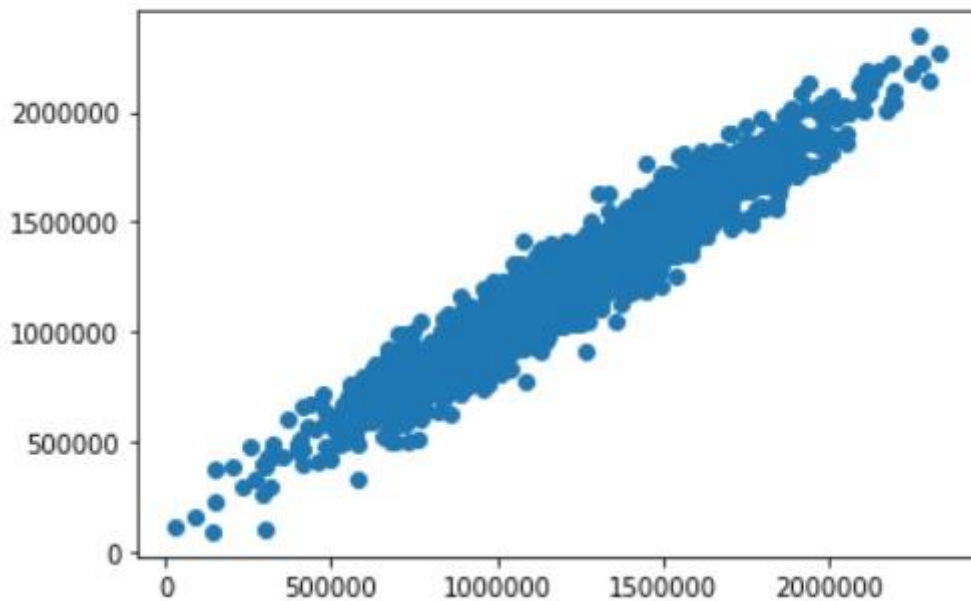Creating and Training the LinearRegression Model

```
from sklearn.linear_model import LinearRegression

lm = LinearRegression()  lm.fit(X_train,y_train)
```

Predictions from our Linear Regression Model

```
predictions = lm.predict(X_test)   plt.scatter(y_test,predictions)
```
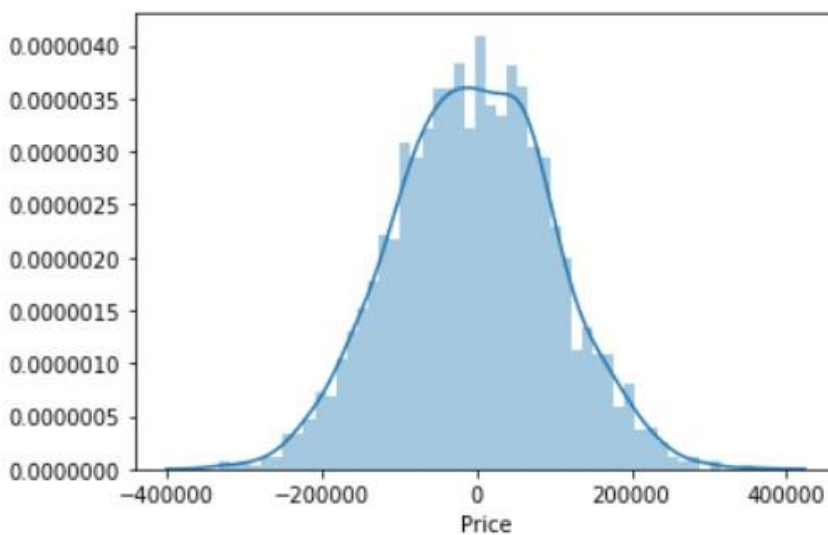
1.3.2 Graph Diagram

`<matplotlib.collections.PathCollection at 0x67b87ccc88>`

In the above scatter plot, we see data is in a line form, which means our model has done good predictions.

```
sns.distplot((y_test-predictions),bins=50);
```


In the above histogram plot, we see data is in bell shape (Normally Distributed), which means our model has done good predictions.

From the above graph(scatter plot) itself we would say that our model has done a good predictions.

REFERENCE

- Machine Learning, Tom Mitchell, McGraw Hill, 1997
- Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow