Capstone Project 1

HOUSING PRICE PREDICTION

N BHARATH | Mentor – SIDDHARTH DIXIT

HOUSING PRICE PREDICTION

OVERVIEW

This project is to predict the housing prices in the given area considering various elements like whether the house contains car garage, swimming pool and how many bedrooms it contains and what is the dimensions of the building etc.

CLIENT

This project does not have a definitive client. But the analysis performed could be of use to anyone in the Real Estate Business (House Owners, Buyers, Tenets, etc).

If we ca make use of the data to find some insights from the past data and could possibly predict the future, many House Owners and Tenets could use this recommendation system on their own. Giving good recommendations directly entails one or many of the following:

1. Customers use the platform more frequently due to the quality and relevance of content shown to them.
2. Better User Experience. Customers do not relay on brokers and will search on their own according to their need and deed.

DATA

The data used in this project has been obtained from Kaggle and it is available in csv (Coma Separated File) format, the data set consists of 1460 instances of training data and 1460 of test data. Total number of attributes equals 81, of which 36 is quantitative, 43 categorical + Id and Sale Price.

Techniques Used to Solve the Problem

- **Data Wrangling:** This section describes the various data cleaning and data wrangling methods applied on the Movie datasets to make it more suitable for further analysis. The following sections are divided based on the procedures followed.
- **EDA (Exploratory Data Analysis):** This step helps us to find out important features of the data, and relation between dependent and independent variables of the data. It is an approach for summarizing, visualizing, and becoming intimately familiar with the important characteristics of a data set. EDA is performed in order to define and refine the selection of feature variables that will be used for machine learning.

- **Inferential Statistics:** This step includes checking the Mean, Median, Mode along the various tastings like T-Test, Z-test etc. Along with this part it includes various steps like drawing inferences using plots such as pair plots, bar plots, histograms.
- **Machine Learning:** After these steps I need to predict the house prices using various regression models.