## PREDICTION OF HOUSE PRICE USING MACHINE LEARNING

**PHASE 1: PROBLEM DEFINITION AND DESIGN THINKING**

**INTRODUCTION**

Predicting house prices using machine learning involves creating a model that can estimate **t**he selling price of a house based on various features or attributes on various features or attributes. Here’s a breakdown of the problem definition and the machine learning for this task.

**PROBLEM DEFINITION**

The goal is to predict the deal price of the houses using various machine learning algorithms. Housing sales prices are determined by numerous factors such as area of the property, location of the house, material used for construction, age of the property, number of bedrooms and garages and so on. This uses machine learning algorithm algorithms to build prediction models for houses. Here machine learning algorithms such as logistic regression and support vector regression, Lasso Regression technique and Decision Tree are employed to build a predictive model.

**DESIGN THINKING**:

**MACHINE LEARNING PROCESS:**

**1.Data Source:** Gather a dataset that includes both the target variable (house prices) and the features. This dataset should be comprehensive and clean. The given dataset link is <https://www.kaggle.com/datasets/vedavyasv/usa-housing>. The dataset includes the following features-Area income, Area house age, Area number of rooms, Area number of bedrooms, Area population, Price, Address.

**2.Data Preprocessing:** Cleaning and preprocess the data by handling missing values, incorrect values, encoding categorical variables and scaling or normalizing numerical features.

**3.Feature Selection:** Create new features or transform existing ones to improve the model’s performance. For example, you can calculate the price per square foot, create dummy variables for categorical features, or extract meaning information from text descriptions.

**4. Splitting the data:** Divide the data set into two parts: a training set and the testing test. The training set is used to train the machine learning model, while the testing set is used to evaluate its performance.

**5. Model Selection:** Choose an appropriate machine algorithm for regression task. Common choices include linear regression tasks. Common choices include linear regression, decision trees, random forest, gradient boosting and neural network. experiment with different algorithms Experiment with different algorithms to find one that performs best for your specific data set.

**6.Model Training:** Train the selected model on the training data. The Model learns the relationship between the input features and the target variable during this step.

**7. Model Evaluation:** Evaluate the model’s performance using metrics such as Mean Absolute Error, Mean Squared Error, and Root Squared Error on the testing data. This step helps you assess how well your model generalizes to unseen data.

**8.Hyperparameter Tuning:** Fine- tune the model by adjusting hyperparameters (e.g., learning rate, tree depth) to improve its performance.

**9.Model Deployment:** Once you are satisfied with the model’s performance. Deploy it in a real-world environment where it can make predictions on new, unseen data.

**10.Monitoring and Maintenance:** Continuously monitor the model’s performance and update it as needed. House prices can change over time, so your model may require periodic retraining.

**CONCLUSION:** Thus, the machine learning model using linear regression algorithm is very helpful in predicting the house prices for real estate customers. Here we used supervised machine learning algorithm to predict the prices of the houses.