### Predicting House Price Using Machine learning.

**Problem Definition: Predicting House Prices using Machine Learning**

**Problem Statement:**

The goal of this project is to develop a machine learning model that can accurately predict the selling prices of residential houses based on various features and attributes associated with the properties.

**Key Components:**

**1. Dataset:** Gather a comprehensive dataset of historical housing information, including details such as square footage, number of bedrooms and bathrooms, location, neighbourhood characteristics, year built, lot size, and other relevant features. This dataset will serve as the foundation for model training and testing.

**2. Data Pre-processing:** Clean and pre-process the dataset to handle missing values, outliers, and categorical variables. This may involve techniques such as data imputation, feature scaling, one-hot encoding, and data normalization to ensure the data is suitable for machine learning.

**3. Feature Selection/Engineering:** Identify the most relevant features that have the most impact on house prices. Additionally, consider creating new features or transforming existing ones if they can provide valuable information to the model.

**4. Splitting the Data**: Divide the dataset into training, validation, and test sets. This separation ensures that the model's performance can be evaluated on unseen data and prevents overfitting.

**5. Model Selection:** Choose an appropriate machine learning algorithm or ensemble of algorithms for regression, as this is a regression problem. Common choices include linear regression, decision trees, random forests, support vector machines, or gradient boosting methods.

**6. Model Training:** Train the selected model(s) on the training data using appropriate training techniques and hyper parameter tuning to optimize performance.

**7. Model Evaluation:** Assess the model's performance on the validation set using relevant evaluation metrics such as Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), or R-squared (R²). Fine-tune the model if necessary.

**8. Model Testing:** Validate the model's performance on the test dataset to ensure its ability to generalize to unseen data.

**9. Interpretability:** If required, analyse the model to understand which features are most influential in predicting house prices. This information can be valuable for both buyers and sellers.

**10. Deployment:** Once satisfied with the model's performance, deploy it in a real-world application, such as a website or mobile app, where users can input property details and receive price predictions.

**Success Criteria:**

The success of the project will be measured based on the model's ability to make accurate predictions of house prices on the test dataset. The chosen evaluation metric (e.g., RMSE or MAE) should meet predefined criteria for accuracy and generalization.

**Ethical Considerations**:

Ensure that the model development process adheres to ethical guidelines, avoids bias and discrimination, and respects privacy laws and regulations. Transparency in how the model makes predictions is also important.

**Next Steps:**

Once the initial model is deployed, continuous monitoring and potential retraining may be necessary to adapt to changing market conditions and maintain prediction accuracy.

This problem definition outlines the steps and considerations involved in predicting house prices using machine learning. It's essential to follow best practices in data science and machine learning throughout the entire process to develop a robust and accurate predictive model.