COLLGE NAME : ST.JOSEPH COLLGE OF ENGINEERING-2129

NAME : GOPIKA S

212921106020

PROJECT : HOUSE PRICE PREDICTING USING MACHINE LEARNING

Problem Definition:

Clearly define the problem: Predict the sale price of houses based on various features.

Gather data: Collect a dataset with historical housing data, including features and actual sale prices.

Define success metrics: Choose evaluation metrics like Mean Absolute Error (MAE) or Root Mean Square Error (RMSE) to measure the model's performance.

Data Collection and Preprocessing:

Collect a comprehensive dataset: Gather data on houses, including features like size, location, number of bedrooms/bathrooms, year built, etc.

Data cleaning: Handle missing values, outliers, and format inconsistencies.

Feature engineering: Create new features or transform existing ones to improve the model's performance.

Data splitting: Divide the dataset into training, validation, and test sets.

Exploratory Data Analysis (EDA):

Visualize the data to gain insights into feature distributions, correlations, and potential outliers.

Identify patterns and relationships that may affect house prices.

Model Selection:

Choose appropriate machine learning algorithms for regression, such as Linear Regression, Decision Trees, Random Forest, or Gradient Boosting.

Consider ensemble methods for better performance.

Feature Selection:

Use techniques like feature importance analysis or recursive feature elimination to select the most relevant features for the model.

Model Training and Evaluation:

Train the chosen model(s) on the training dataset.

Evaluate the models on the validation set using the defined metrics.

Tune hyperparameters to optimize model performance.

Model Interpretation:

Interpret the model's results to understand which features have the most significant impact on house prices.

Deployment:

Once satisfied with model performance, deploy it as a service or application for predictions.

Feedback Loop:

Continuously monitor the model's performance and update it as new data becomes available.

Ethical Considerations:

Ensure fairness and transparency in the model's predictions, avoiding biases and discrimination.

User Interface:

Create a user-friendly interface for users to input property details and get price predictions.

Maintenance and Updates:

Regularly update the model to adapt to changing real estate market dynamics.P