Phase1: Problem Definition and Design Thinking

The scope of this document is to identify the problem and develop a project on predicting house prices using machine learning

Problem Definition:

Through research it has been found out that there are the following problems with the chatbot. The problems are listed below:

1. Data Collection:

Gather a dataset containing information about houses, including features like size, location, number of bedrooms, etc and ensure the dataset is clean, free of missing values, and properly labeled.

2. Data Preprocessing:

Perform data cleaning and handle missing values and encode categorical variables into numerical format.

Normalize or standardize numerical features to ensure they're on the same scale.

3.Feature selection: Choose the most relevant features to include in the model.

4. Model Selection

Choose a machine learning algorithm suitable for regression tasks. Common choices include Linear Regression, Random Forest, or Gradient Boosting and split your dataset into training and testing sets.

5. Model Training:

Train the selected model using the training dataset. Fine-tune hyperparameters to optimize performance.

6. Model Evaluation:

Evaluate the model's performance using metrics like Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), or R-squared.

Design Thinking:

By having understood the above problem. We would be designing a solution to solve it.

1. Machine learning algorithms for predicting house prices could involve using advanced deep learning models, reinforcement learning, or other innovations to better capture complex patterns in housing data.

2. Efforts to make machine learning models more interpretable and explainable will likely increase, helping users understand the factors driving house price predictions.

3. Personalized predictions based on individual preferences and requirements could become a standard offering, allowing users to input their specific criteria for a home and receive tailored price estimates.

4. Machine learning can be applied to predict not only the price of homes but also the potential return on investment for real estate investors, guiding them in making informed decisions.

5. Lenders and insurers can use machine learning to assess the risk associated with property transactions, leading to more accurate underwriting and pricing.