DESIGN INTO INNOVATION TO SOLVE THE PROBLEM

PHASE-2 DOCUMENT SUBMISSION

PROJECT: innovation to solve the problem

TEAM MEMBERS

110521106002-R.Abirami
110521106003-K.Arularasi
110521106014-C.Monisha
110521106325-K.Ponnarasi

Title: House Price Prediction Using Python

SOME OF THE STEP ARE INCLUDE IN MY PROJECT :

- **Problem Definition**:
  - Understand the problem of house price prediction and its objectives.

- **Data Collection**:
  - Gather a comprehensive dataset of historical house prices and relevant features (e.g., square footage, number of bedrooms, location, etc.).

- **Data Preprocessing**:
  - Handle missing data and outliers. ○ Perform data normalization or standardization.
  - Split the data into training and testing sets.

- **Feature Engineering**:
  - Select relevant features.
  - Create new features if necessary (e.g., feature extraction from text or images).

- **Model Selection**:
  - Choose the appropriate machine learning algorithm(s) for regression (e.g., Linear Regression, Random Forest, XGBoost).

- **Model Training**:
    - Train the selected model(s) on the training data.

- **Model Evaluation**:
    - Evaluate the model's performance on the testing dataset using appropriate metrics (e.g., Mean Absolute Error, Root Mean Squared Error).

- **Hyperparameter Tuning**:
    - Fine-tune the model's hyperparameters to improve performance.

- **Final Model**:
    - Select the best-performing model as the final model.

- **Deployment**:
    - Deploy the final model to a production environment.

- **Inference**:
    - Allow users to input information about a house they want to predict the price for.

- **User Interface**:
    - Create a user-friendly interface for users to input data.

- **House Price Prediction**:
    - Use the deployed model to predict the house price based on user input.

- **Display Result**:
    - Show the predicted house price to the user.

Incorporating the feedback loop:

- **Model Evaluation:**

After evaluating the model's performance, check if it meets the desired criteria. If it does, proceed as usual. If not, continue to the feedback loop.

- **Feedback Gathering:**

Collect feedback from users or domain experts about the model's performance and any issues they encountered.

- **Data Update:**

Update the dataset with new data if available, potentially including recent house prices and additional features.

- **Model Retraining:**

Retrain the model using the updated dataset and incorporate any feedback-driven improvements or adjustments to the model.

- **Loop Back to Model Training:**
- Return to the model training stage with the updated data and model improvements.

In practice, each step may involve multiple sub-steps and iterations to fine-tune the model and optimize its performance. Additionally, data privacy, model explainability, and ethical considerations should be taken into account throughout the entire process.