FINAL PROJECT PROPOSAL: HOUSE PRICE PREDICTION

1. Problem Statement: Our group is committed to finding a solution to the difficult prediction of the housing price problem. The goal is to create a reliable and accurate machine learning model that can forecast residential property sales prices based on a variety of factors, including square footage, the number of bedrooms and bathrooms, location, and other important factors. Potential buyers, sellers, and real estate professionals will all find this predictive model to be quite helpful in assisting them in making well-informed decisions regarding real estate transactions.

2. Description of Data Set: For data analysts, the real estate markets, such as those in America, give an intriguing chance to evaluate and forecast where property prices are headed. Property price forecasting is becoming more crucial and advantageous. Property values are a reliable predictor of a nation's economic health and general market state. We are organizing a sizable collection of real estate sales records that are kept in an unknown format and have unknown data quality issues based on the information supplied.

3. Implementation Plan:
Milestone 1: Data Collection and Preprocessing
Duration: Weeks 1-2
   Tasks:
- Collect and download the dataset from Kaggle.
- Clean the dataset, handling missing values and outliers.
- Perform data exploration and visualization to gain insights into the dataset.

Milestone 2: Feature Engineering and Selection
Duration: Weeks 3-4
   Tasks:
- Identify relevant features that influence house prices.
- Engineer new features if needed.
- Select the most significant features to include in the model.

Milestone 3: Model Development
Duration: Weeks 5-8
   Tasks:
- Choose appropriate machine learning algorithms for regression.
- Split the dataset into training and testing sets.
- Train and fine-tune models using cross-validation.
- Evaluate model performance using appropriate metrics.

Milestone 4: Documentation and Presentation
Duration: Week 8-9
   Tasks:
- Prepare comprehensive documentation for the project, including data sources, model details, and user instructions.
- Create a presentation summarizing the project's key findings and the model's performance.

4. Team members & task allocation:
1. Shivanshu Singh: - Task: Data Collection and Preprocessing
 Responsibilities: Gather house price data from various sources.
   • Ensure data quality and consistency.
   • Perform initial data preprocessing, including handling missing values and data cleaning.

2. Visharad Ravi: - Task: ML Algorithm Design and Development
 Responsibilities: Design machine learning algorithms and models for house price prediction.
   • Implement the chosen machine learning algorithms.
   • Conduct initial testing and debugging of the models.

3. Uday Kurella: Task: Model Training, Evaluation, and Reporting
 Responsibilities: Train machine learning models using data.
   • Perform model evaluation and validation.
   • Generate comprehensive reports and documents.