**PROJECT: PREDICTING HOUSE PRICES USING MACHINE LEARNING**

**TEAM MEMBER: Yamini T**

**Phase 1 Submission**

**Abstract:**

Predicting house prices is a crucial task in the real estate industry, as it helps buyers, sellers, and investors make informed decisions. Machine learning offers a powerful approach to tackle this challenge by leveraging historical data and various features of properties. This abstract outlines a design thinking approach to predict house prices using machine learning, encompassing problem definition, data collection, feature engineering, model selection, and evaluation.

**Design thinking process**:

**Empathize:**

Understand the stakeholders' needs, such as homebuyers, real estate agents, and investors.

Gather insights into the factors that influence house prices, like location, size, amenities, and market trends.

**Define:**

Clearly define the problem: Develop a well-defined problem statement, e.g., "Predict house prices accurately based on property features."

Set measurable objectives: Determine the metrics for evaluating the model's performance, such as Mean Absolute Error (MAE) or Root Mean Square Error (RMSE).

**Ideate:**

Identify potential data sources: Explore publicly available real estate datasets, local property listings, or APIs.

Brainstorm features: Consider factors like location (geographical coordinates, neighborhood characteristics), property characteristics (size, number of bedrooms, bathrooms, amenities), historical sales data, and economic indicators.

**Prototype:**

Data Collection: Gather a comprehensive dataset containing historical sales records, property details, and relevant economic data.

Data Preprocessing: Clean and preprocess the data, handling missing values, outliers, and encoding categorical variables.

Feature Engineering: Create new features or transform existing ones to extract valuable information.

Model Selection: Choose appropriate machine learning algorithms, such as linear regression, decision trees, or ensemble methods.

Model Development: Build and train the selected models using a portion of the dataset.

**Test and Iterate:**

Split the data into training and testing sets to evaluate model performance.

Use evaluation metrics (e.g., MAE, RMSE) to assess the model's accuracy.

Iterate on the model by fine-tuning hyperparameters, trying different algorithms, or incorporating feedback from stakeholders.

**Implement:**

Deploy the trained model as a web application, mobile app, or API for stakeholders to access and utilize.

Ensure scalability and real-time updates for market changes.

**Evaluate and Improve**:

Continuously monitor the model's performance and gather user feedback.

Implement improvements based on new data and emerging trends.

Consider retraining the model periodically to maintain accuracy.

In summary, this design thinking approach provides a systematic framework for predicting house prices using machine learning. By understanding the problem, defining objectives, collecting and preprocessing data, selecting models, and iterating based on evaluation, we can create a valuable tool for various stakeholders in the real estate industry.