**Ideation Phase**

**Defining the Problem Statements**

| **Date** | **26-09-2023** |
| --- | --- |
| **Team ID** | **676** |
| **Project Name** | **House Price Prediction using ML** |

**House Price Prediction using Machine Learning**

**Problem Definition and Design Thinking**

**Introduction**

The task at hand is to develop a machine learning model that can accurately predict house prices based on a set of relevant features. House price prediction is a common problem in the real estate industry and has a wide range of applications, from helping buyers make informed decisions to assisting real estate professionals in setting competitive prices for listings.

In this document, we will outline the problem statement, the steps involved in solving it, and the design thinking approach that will guide our project.

**Problem Statement**

Objective: Develop a machine learning model that can predict house prices with a high level of accuracy.

Data: We have a dataset containing various features of houses (e.g., size, location, number of bedrooms, number of bathrooms, etc.) along with their corresponding sale prices. This data will be used to train and evaluate our machine learning model.

**Key Challenges:**

1. Data Quality: Ensuring the dataset is clean, complete, and free of errors.

2. Feature Selection: Identifying the most relevant features for accurate price predictions.

3. Model Selection: Choosing the appropriate machine learning algorithm(s) for the task.

4. Model Evaluation: Evaluating the model's performance using appropriate metrics.

5. Deployment: Creating a user-friendly interface or API for end-users to make predictions.

**Design Thinking Approach**

**Empathize:**

Before diving into solving the problem, it's crucial to empathize with the users and understand their needs. In this case, our primary users are potential homebuyers and real estate professionals. We need to gather insights into what factors are most important to them when considering house prices and how accurate predictions can benefit them.

**Actions:**

- Conduct surveys or interviews with potential users to gather their perspectives.

- Analyse historical real estate market trends to identify critical pricing factors.

- Seek feedback from domain experts in the real estate industry.

**Define:**

Based on our understanding of the problem and the users' needs, we will define clear objectives and success criteria for our project.

**Objectives:**

- Develop a machine learning model that achieves a Mean Absolute Error (MAE) of less than $X on the test data.

- Create a user-friendly web application for users to input house details and receive price predictions.

**Ideate:**

Brainstorm potential solutions and approaches to address the problem. This phase involves thinking creatively and considering various algorithms and techniques for house price prediction.

**Actions:**

- Explore different machine learning algorithms such as linear regression, decision trees, random forests, and neural networks.

- Experiment with feature engineering techniques to enhance model performance.

- Consider incorporating external data sources (e.g., neighbourhood crime rates, school quality) to improve predictions.

**Prototype**

Create a prototype of the machine learning model and the user interface for price prediction.

**Actions:**

- Develop a Jupyter Notebook or Python script for data pre-processing, model training, and evaluation.

- Create a simple web interface using tools like Flask or Django to allow users to input house details.

- Test the prototype with a subset of the dataset to ensure it meets performance objectives.

**Test**

Evaluate the model's performance using appropriate metrics and gather feedback from users.

**Actions:**

- Split the dataset into training and testing sets.

- Train the model on the training set and evaluate it on the testing set.

- Use metrics such as MAE, Root Mean Square Error (RMSE), and R-squared to assess model performance.

- Collect user feedback on the web interface for usability and accuracy.

**Implement:**

Once the prototype meets the defined objectives and receives positive feedback, proceed with full implementation.

**Actions:**

- Train the final machine learning model on the entire dataset.

- Deploy the model as part of a production-ready web application.

- Conduct thorough testing to ensure the application is robust and user-friendly.

**Iterate**

Continuous improvement is essential. Gather user feedback and iterate on the model and interface to enhance accuracy and usability.

**Actions:**

- Monitor the model's performance and retrain it periodically with updated data.

- Address user feedback and make necessary improvements to the web interface.

- Stay informed about advancements in machine learning and real estate pricing models for potential enhancements.

**Conclusion**

In this document, we've outlined our approach to solving the problem of house price prediction using machine learning. We've defined the problem, identified key challenges, and laid out a design thinking approach that involves empathizing with users, defining objectives, ideating potential solutions, prototyping, testing, implementing, and iterating.

Our ultimate goal is to develop an accurate and user-friendly solution that provides valuable insights for both homebuyers and real estate professionals in the housing market. By following this structured approach, we aim to create a reliable tool that contributes positively to the real estate industry.