**Name: Shaik Mahaboob**

**Roll no: 20211CEI0142**

**Mini Project Report: House Price Prediction using Applied Machine Learning**

**1. Introduction**

**1.1 Background**

The objective of this mini project is to develop a machine learning model for predicting house prices based on various features. Accurate house price prediction is crucial for real estate professionals, homeowners, and potential buyers to make informed decisions.

**1.2 Scope**

The project focuses on exploring and implementing machine learning techniques to analyze a dataset containing relevant features such as square footage, number of bedrooms, location, and other pertinent factors affecting house prices.

**2. Dataset Overview**

**2.1 Data Source**

The dataset used for this project was obtained from **kaggle.com**. It comprises a collection of real estate data, including both numerical and categorical features.

**2.2 Data Exploration**

Features: The dataset includes features such as square footage, number of bedrooms, bathrooms, location details, and other relevant attributes.

Target Variable: The target variable is the house price, which we aim to predict.

**3. Data Preprocessing**

**3.1 Handling Missing Data**

Any missing or null values in the dataset were addressed through appropriate imputation techniques or removal of incomplete records.

**3.2 Feature Engineering**

New features were created, and existing features were transformed to enhance the model's predictive performance. Feature scaling and normalization were applied as needed.

**4. Model Selection**

**4.1 Algorithms Considered**

Several machine learning algorithms were considered for the house price prediction task, including:

Linear Regression

Decision Trees

Random Forest

Gradient Boosting

Neural Networks

**4.2 Model Evaluation**

Models were evaluated using metrics such as Mean Squared Error (MSE), Mean Absolute Error (MAE), and R-squared to determine their predictive performance. Cross-validation techniques were employed to ensure robustness.

**5. Model Training and Tuning**

The selected model was trained on the dataset, and hyperparameter tuning was performed to optimize the model's performance. Grid search or random search techniques were used for this purpose.

**6. Results**

The performance of the machine learning model was assessed using the evaluation metrics mentioned earlier. The final model achieved satisfactory results in terms of accuracy and precision in predicting house prices.

**7. Discussion**

**7.1 Challenges Faced**

Discuss any challenges encountered during the project, such as data quality issues, feature selection difficulties, or model performance challenges.

**7.2 Limitations**

Highlight any limitations of the model and potential areas for improvement.

**8. Conclusion**

This mini project successfully developed a machine learning model for predicting house prices based on a given set of features. The model demonstrates promising results, and future work may involve further refinement, incorporating additional features, and exploring advanced modeling techniques.