



**DIGITAL[®]
EMPOWERMENT
NETWORK**

Digital Empowerment Network

Data Science

Week 02

House Price Analysis and Prediction

Mentor : Ali Mohiuddin Khan

Task 2 – House Price Analysis and Prediction

Project Submission Guidelines

1. Time Management

- Ensure you adhere strictly to the deadlines. Timeliness is crucial.

2. Proper Documentation

- Along with your project submission, prepare and submit proper documentation detailing your project.

This documentation should include:

- o **Project Overview:** A brief description of your project.
- o **Objectives:** The main goals and objectives of your project.
- o **Methodology:** The approach and methods you used to complete the project.
- o **Challenges:** Any challenges or hurdles you faced and how you overcame them.
- o **Conclusion:** The final outcome and any recommendations or future steps. Tasks

Objective: Analyze a dataset of house prices to understand pricing factors, identify outliers, and develop a model for future price prediction.

Tasks:

1. Data Cleaning and Exploration

Clean the data: Identify and handle missing values, inconsistencies, and outliers.

Explore the data: Analyze the distribution of house prices and other features. Identify potential relationships between features and price using visualizations (scatter plots, box plots, etc.).

2. Feature Engineering

Create new features that might be relevant for price prediction (e.g., age of the house, number of bedrooms per floor).

- o Consider encoding categorical features (e.g., location) into numerical values suitable for modeling.

3. Outlier Analysis

Identify houses with significantly higher or lower prices compared to similar properties.

Investigate the reasons for these outliers. Are there any specific features or combinations of features that contribute to the outliers?

4. Predictive Modeling

- o Train a machine learning model to predict house prices based on the available features. Popular choices for this task include linear regression, random forest, or gradient boosting.
- o Evaluate the performance of the model using appropriate metrics (e.g., mean squared error, R-squared).

5. Future Price Prediction

- o Use the trained model to predict future house prices based on hypothetical scenarios (e.g., what would be the price of a house with specific characteristics in a particular location?).

6. Report and Recommendations

- o Prepare a report summarizing your findings, including:
 - Data exploration results (key insights from visualizations)
 - Feature engineering techniques used
 - Outlier analysis (identification and explanation)
 - Model selection and evaluation results
 - Future price prediction examples
 - Recommendations for further analysis or data collection (if applicable)

Deliverables

- A well-documented script containing your data cleaning, exploration, feature engineering, modeling, and prediction code.
- A clear and concise report summarizing your findings and recommendations.

Evaluation

- Your work will be evaluated based on the following criteria:
 - o Completeness of tasks
 - o Data analysis skills (cleaning, exploration, visualization)
 - o Understanding of feature engineering concepts
 - o Ability to build and evaluate a machine learning model
 - o Quality and clarity of reporting

Additional Documentation Requirements

- Objectives: The main goals and objectives of your project.
- Methodology: The approach and methods you used to complete the project.
- Challenges: Any challenges or hurdles you faced and how you overcame them.
- Conclusion: The final outcome and any recommendations or future steps.

3. Originality

- Your project must be original and not copied from any source. Plagiarism will result in the cancellation of your project submission.

Note

- Regularly update your LinkedIn profile with your progress and achievements, as it is crucial for your career.

Deadline 30-July-2025