Name: Muaaz Hassan Ansari

Id: DHC-107

**Task 3: Multimodal ML – Housing Price Prediction Using Images + Tabular Data**

**Objective**

The objective of this task is to **predict housing prices** using both structured tabular data and house images. By combining **image features extracted using Convolutional Neural Networks (CNNs)** with traditional tabular data (e.g., number of rooms, location, area), we build a **multimodal machine learning pipeline** capable of robust regression modeling.

**Dataset**

The dataset consists of two modalities:

1. **Tabular Data**
   * Housing attributes such as size, number of rooms, location, etc.
2. **Image Data**
   * Pictures of houses corresponding to the tabular entries.

Both datasets are aligned such that each house has structured features and an image.

**Methodology**

1. **Preprocessing Tabular Data**
   * Handle missing values
   * Normalize/scale numerical features
   * Encode categorical features
2. **Image Feature Extraction**
   * Resize images to a standard size (e.g., 224×224)
   * Use a pretrained CNN (e.g., ResNet50) to extract deep features
3. **Feature Fusion**
   * Concatenate CNN-based image embeddings with tabular features
   * Pass the fused features to a regression model (e.g., Dense layers)
4. **Model Training & Evaluation**
   * **Loss function**: Mean Squared Error (MSE)
   * **Metrics**: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE)

**Implementation Steps**

1. Load and preprocess the tabular dataset.
2. Load and preprocess the image dataset.
3. Build a CNN model for image feature extraction.
4. Combine image features with tabular features.
5. Train the combined model for regression.
6. Evaluate performance using MAE and RMSE.
7. Save and export the trained model.

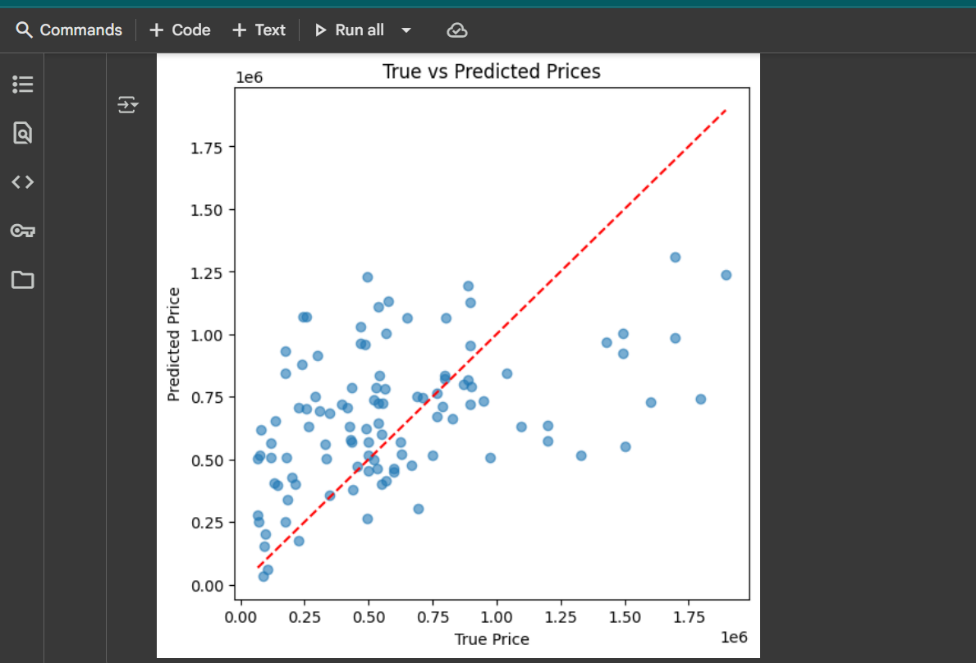
**Evaluation Metrics**

* **Mean Absolute Error (MAE):** Measures the average magnitude of errors in predictions.
* **Root Mean Squared Error (RMSE):** Penalizes larger errors and provides a more sensitive measure.

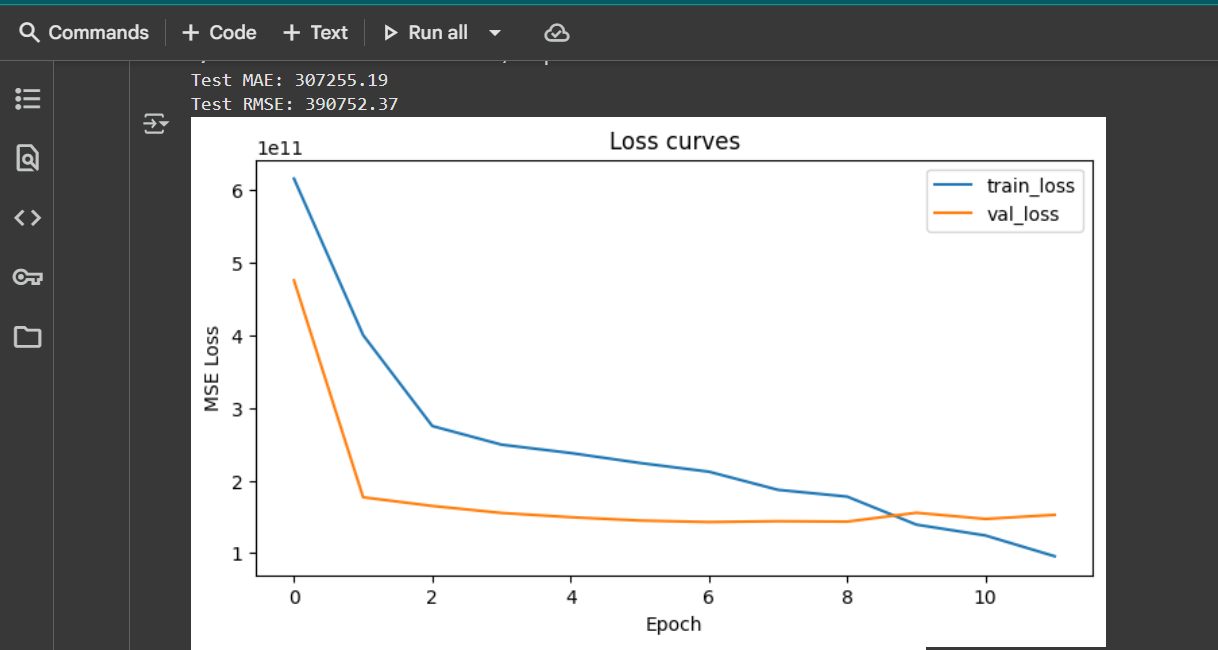
**Results**

The results of the multimodal model can be visualized as follows:

* 📊 **Predicted vs Actual Housing Prices**



* 📈 **Training vs Validation Loss Curve**



**Example Results:**

* Test MAE: 307255.19
* Test RMSE: 390752.37

**Conclusion**

This project demonstrates the effectiveness of **multimodal machine learning**, where CNN-based image features are combined with tabular features to predict housing prices. The fusion of modalities improves prediction accuracy compared to using tabular or image data alone.