Supervised ML (Assignment1)

Assignment Project

Problem Statement

A real estate company **HomeVista Properties** operates across multiple cities and handles thousands of residential property sales every year. The company wants to automate its house pricing process.

They wants to use **Machine Learning** to build an intelligent system that can **predict the market price of a house automatically** based on its physical features, location, and condition.

You are hired as a **Machine Learning Engineer** to build a regression model that can accurately predict the house price using historical property data.

Your task is to analyse the dataset, perform data preprocessing, train a **Linear Regression model**, and evaluate its performance.

Dataset Description

Each row represents one residential house and its physical, location, and construction details.

Feature	Description
Id	Unique identification number for each house
MSSubClass	Type of dwelling involved in the sale (numerical code representing building class such as 20 = 1-Story, 60 = 2-Story, etc.)
MSZoning	General zoning classification of the sale (Residential Low Density, Medium Density, etc.)
LotArea	Lot size in square feet
LotConfig	Lot configuration (Inside, Corner, Cul-de-sac, etc.)
BldgType	Type of dwelling (1Fam, 2Fam, Duplex, Townhouse, etc.)
OverallCond	Overall condition rating of the house (scale 1–10)
YearBuilt	Original construction year
YearRemodAdd	Year the house was remodeled or additions were made
Exterior1st	Exterior covering on house (VinylSd, MetalSd, HdBoard, etc.)

Feature	Description
BsmtFinSF2	Type 2 finished square feet of basement
TotalBsmtSF	Total square feet of basement area
SalePrice (Target)	Final selling price of the house