```python
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler, OneHotEncoder
from sklearn.compose import ColumnTransformer
from sklearn.pipeline import Pipeline
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_absolute_error, r2_score
import joblib

# Load dataset
df = pd.read_csv("C:\\Users\\hp\\Downloads\\Housing (1).csv")

# Features and target
X = df.drop(columns="price", axis=1)
y = df["price"]

# Identify categorical & numerical columns
categorical_cols = ["mainroad", "guestroom", "basement",
                    "hotwaterheating", "airconditioning",
                    "prefarea", "furnishingstatus"]

numerical_cols = [col for col in X.columns if col not in
categorical_cols]

# Preprocessing
preprocessor = ColumnTransformer([
    ("num", StandardScaler(), numerical_cols),
    ("cat", OneHotEncoder(drop="first"), categorical_cols)
])

# Build pipeline
model = Pipeline([
    ("preprocessor", preprocessor),
    ("regressor", LinearRegression())
])

# Train-test split
X_train, X_test, y_train, y_test = train_test_split(
    X, y, test_size=0.2, random_state=42
)

# Train model
model.fit(X_train, y_train)

# Predictions
y_pred = model.predict(X_test)

# Evaluate
print("MAE:", mean_absolute_error(y_test, y_pred))
print("R2 Score:", r2_score(y_test, y_pred))
```

```python
# Save model
joblib.dump(model, "model.joblib")
print("✅ Model saved successfully as model.joblib")

MAE: 970043.4039201641
R2 Score: 0.6529242642153177
✅ Model saved successfully as model.joblib

import streamlit as st
import joblib
import numpy as np
import pandas as pd

# Load model
model = joblib.load("model.joblib")

st.title("🏠 House Price Prediction App")
st.write("Enter the house details to predict the price:")

# Input fields
area = st.number_input("Area (sq ft)", 500, 20000, 7420)
bedrooms = st.slider("Bedrooms", 1, 10, 3)
bathrooms = st.slider("Bathrooms", 1, 10, 2)
stories = st.slider("Stories", 1, 5, 2)
mainroad = st.selectbox("Mainroad", ["yes", "no"])
guestroom = st.selectbox("Guestroom", ["yes", "no"])
basement = st.selectbox("Basement", ["yes", "no"])
hotwaterheating = st.selectbox("Hot Water Heating", ["yes", "no"])
airconditioning = st.selectbox("Air Conditioning", ["yes", "no"])
parking = st.slider("Parking Spaces", 0, 5, 1)
prefarea = st.selectbox("Preferred Area", ["yes", "no"])
furnishingstatus = st.selectbox("Furnishing Status",
                                ["furnished", "semi-furnished",
"unfurnished"])

# Prepare input data
input_data = pd.DataFrame([{
    "area": area,
    "bedrooms": bedrooms,
    "bathrooms": bathrooms,
    "stories": stories,
    "mainroad": mainroad,
    "guestroom": guestroom,
    "basement": basement,
    "hotwaterheating": hotwaterheating,
    "airconditioning": airconditioning,
    "parking": parking,
    "prefarea": prefarea,
    "furnishingstatus": furnishingstatus
```

```
}])

# Predict
if st.button("Predict Price"):
    prediction = model.predict(input_data)[0]
    st.success(f"□ Estimated House Price: ₹{prediction:,.2f}")
```