import pandas as pd

import pickle

from sklearn.pipeline import Pipeline

from sklearn.compose import ColumnTransformer

from sklearn.preprocessing import OneHotEncoder

from xgboost import XGBRegressor

from sklearn.model\_selection import train\_test\_split

# Load your data

df = pd.read\_csv("Tourism\_Master.csv", encoding='ISO-8859-1')

df.dropna(inplace=True)

# Features and target

X = df[["Country", "Attraction", "VisitSeason", "AttractionPopularity", "UserAvgRating"]]

y = df["Rating"]

# Preprocessing

cat\_cols = ["Country", "Attraction", "VisitSeason"]

preprocessor = ColumnTransformer([

    ("cat", OneHotEncoder(handle\_unknown='ignore'), cat\_cols)

], remainder='passthrough')

# Pipeline

pipeline = Pipeline([

    ("preprocessor", preprocessor),

    ("model", XGBRegressor(n\_estimators=100, learning\_rate=0.1, max\_depth=6, random\_state=42))

])

# Train

X\_train, X\_test, y\_train, y\_test = train\_test\_split(X, y, test\_size=0.2, random\_state=42)

pipeline.fit(X\_train, y\_train)

# Save locally

with open("regression\_model.pkl", "wb") as f:

    pickle.dump(pipeline, f)

print("✅ Model retrained and saved locally.")