# Importing libraries

import pandas as pd

import numpy as np

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

from sklearn.ensemble import RandomForestRegressor

from sklearn.metrics import mean\_squared\_error, r2\_score

import matplotlib.pyplot as plt

import seaborn as sns

# Load your dataset

# Example: UCI Air Quality dataset or your own CSV

data = pd.read\_csv('AirQuality.csv') # replace with your path

# Display first few rows

print(data.head())

# Preprocess the data

data = data.dropna() # remove rows with missing values

# Features and target

X = data.drop(['AirQualityIndex'], axis=1) # replace with actual target column

y = data['AirQualityIndex']

# Train-test split

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

# Initialize and train model

model = RandomForestRegressor(n\_estimators=100, random\_state=42)

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

# Predict

y\_pred = model.predict(X\_test)

# Evaluation

print("RMSE:", np.sqrt(mean\_squared\_error(y\_test, y\_pred)))

print("R^2 Score:", r2\_score(y\_test, y\_pred))

# Feature importance plot

plt.figure(figsize=(10,6))

sns.barplot(x=model.feature\_importances\_, y=X.columns)

plt.title('Feature Importance in Air Quality Prediction')

plt.show()