import pandas as pd

from sklearn.linear\_model import LinearRegression

import matplotlib.pyplot as plt

import pickle

csv\_file\_path = '\data\_train\Low-fidelity output + High-fidelity prediction residuals (Active learning)\_train.csv'

data = pd.read\_csv(csv\_file\_path, encoding='GBK')

print(data.head())

X = data[['Final value']] # Independent variable, sum of low-fidelity and high-fidelity predicted values

y = data['label'] # Dependent variable, true value (in fact, this is a probability value, not a label value)

model = LinearRegression()

model.fit(X, y)

print(f"Regression coefficient: {model.coef\_[0]}")

print(f"Intercept: {model.intercept\_}")

model\_save\_path = '\Linear model\linear\_model.pkl'

with open(model\_save\_path, 'wb') as file:

pickle.dump(model, file)

y\_pred = model.predict(X)