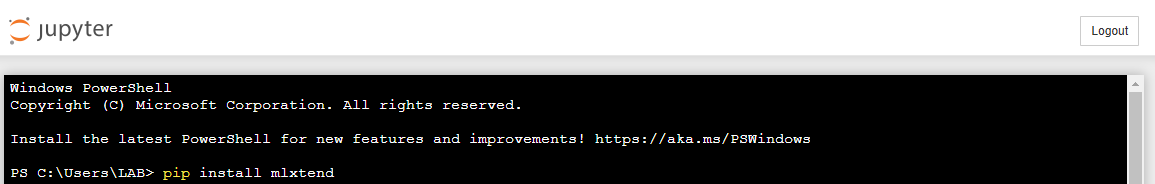
**Experiment-5**

**AIM: Develop a program for Bias, Variance, remove duplicates, Cross Validation:**

**Install module from terminal:**

**>pip install mlxtend**

****

**import csv file housing.csv from path or user and enter the path where it exists:**

**for example: 'D:/housing.csv'**

**program:**

from pandas import read\_csv

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

from sklearn.linear\_model import LinearRegression

from mlxtend.evaluate import bias\_variance\_decomp

import numpy as np

# Load dataset

url = 'D:/housing.csv'

dataframe = read\_csv(url, header=None)

# Separate into inputs and outputs

data = dataframe.values

X, y = data[:, :-1], data[:, -1]

# Ensure that X and y are numpy arrays

X = np.array(X)

y = np.array(y)

# Split the data

X\_train, X\_test, y\_train, y\_test = train\_test\_split(X, y, test\_size=0.33, random\_state=1)

# Define the model

model = LinearRegression()

# Estimate bias and variance

mse, bias, var = bias\_variance\_decomp(model, X\_train, y\_train, X\_test, y\_test, loss='mse',

num\_rounds=200, random\_seed=1)

# Summarize results

print(f'MSE: {mse:.3f}')

print(f'Bias: {bias:.3f}')

print(f'Variance: {var:.3f}')

Output:

MSE: 135.372

Bias: 113.311

Variance: 22.061