Create ubuntu instance, ssh into it

ssh -i [path to .pem file] ubuntu@ec2-[ip address].ap-southeast-1.compute.amazonaws.com

1. **Install pip3**

sudo apt install python3-pip

1. **Run the following to download the necessary files and packages:**

pip3 install pandas

pip3 install scikit-learn

pip3 install xgboost

pip3 install imblearn

1. **Create folders in the ubuntu instance**

mkdir input

mkdir processed

mkdir output

mkdir py

mkdir model

1. **Download the input files**

cd input

wget https://github.com/burritotales/dsa4266proj2/raw/main/col\_list.txt

wget https://github.com/burritotales/dsa4266proj2/raw/main/dataset2.json.gz

cd ../model

wget https://github.com/burritotales/dsa4266proj2/raw/main/top\_features\_xgboost\_model.pkl

cd ../py

wget https://raw.githubusercontent.com/burritotales/dsa4266proj2/main/aws\_prediction.py

1. **Run the py script**

python3 aws\_prediction.py

1. **View the output file**

cd ../output

nano dataset2\_output.csv