1. In order to complete this assignment, I utilized Amazon sagemaker. First I set up AWS EMR environment. This helped me create the ec2 instances and security groups needed to communicate with each other. So in this case, I utilized amazon sagemaker to run a parallel model.

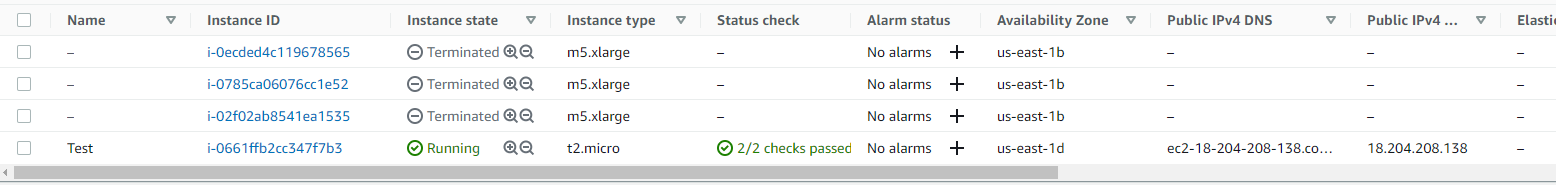
Graphical user interface, text, application

Description automatically generated

EC2 Instances & Security Groups:

Graphical user interface, text, application

Description automatically generated



1. Next, I created a sagemaker notebook which enabled me to write my python code in Jupyter notebook to train the model. I also created an S3 bucket to store the training and validation set.

Graphical user interface, application

Description automatically generated

Graphical user interface, text, application

Description automatically generated

1. Once this was set up, I was able to write out my code to train my ML model. For this assignment, I decided to go with a random forest model.

Text

Description automatically generated

Graphical user interface, text, application, email

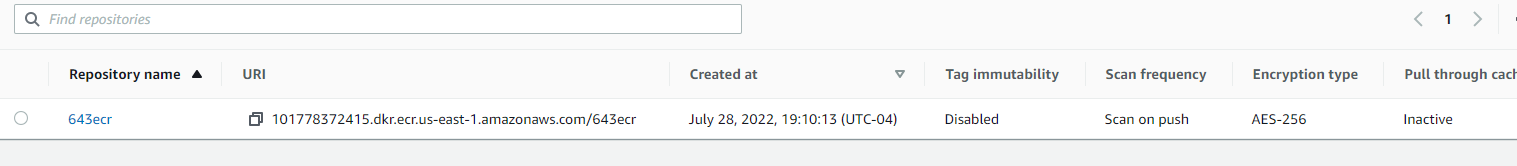
Description automatically generated

1. Next I set up a docker container on AWS to run the task of making the prediction. It is linked to an ECR.

Graphical user interface, text, application, email

Description automatically generated

ECR Configs:



1. Overall python code is the basis for the docker contained to make a prediction.