Christian Shadis

Dr. Alsallami

CS-383

Due 5/2/2021

Final Project Report 1

# Progress

I have developed an r script that classifies an image with 55% accuracy on a small subset of the training data. I am now working on scaling the analysis up to EMR/Spark and trying to parallelize processes to be able to train the model using all 1800+ pictures in the training set. This, along with parameter tweaking, will hopefully yield a model that will be much more accurate.

# Challenges/Changes

Figuring out how to scale the script up to AWS has been a challenge that I have spent the entire weekend trying to solve. I was originally planning on using the R Hadoop tool, but I believe Spark on an EMR cluster may be a better fit.

# Next Steps

Over the coming week I will finish scaling the model up to AWS and increase the accuracy of the model by increasing the amount of training data used, and by changing parameters. I will also develop the 3-page report and begin to work on my presentation.