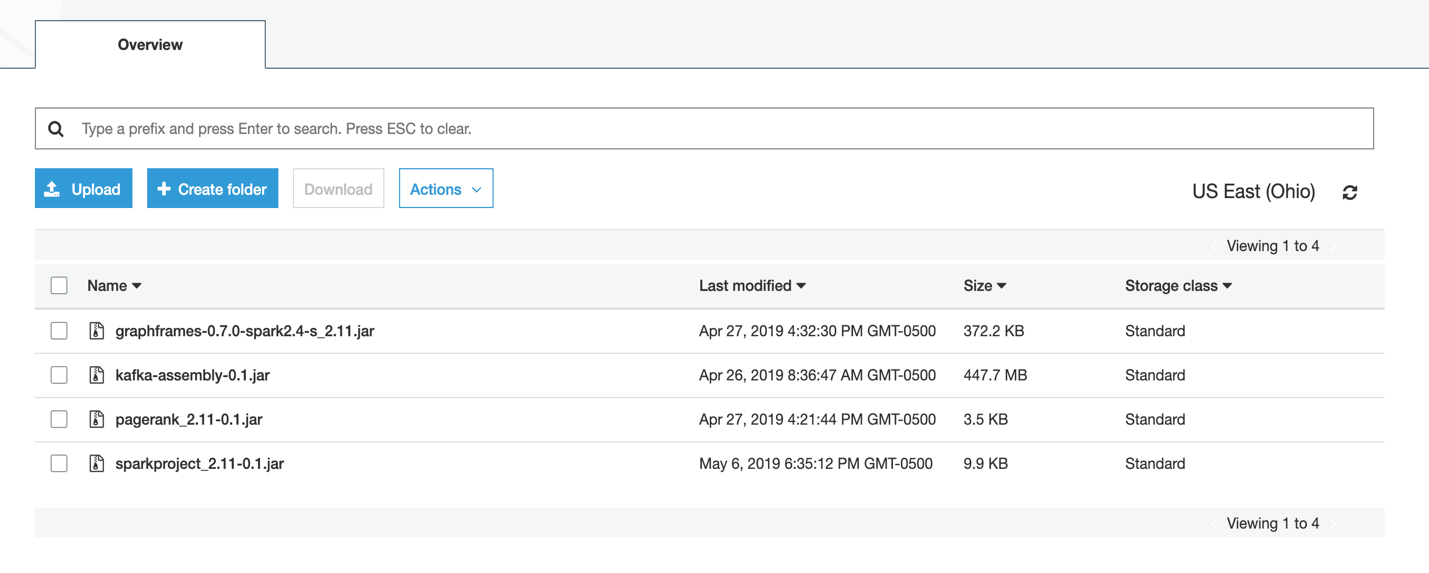
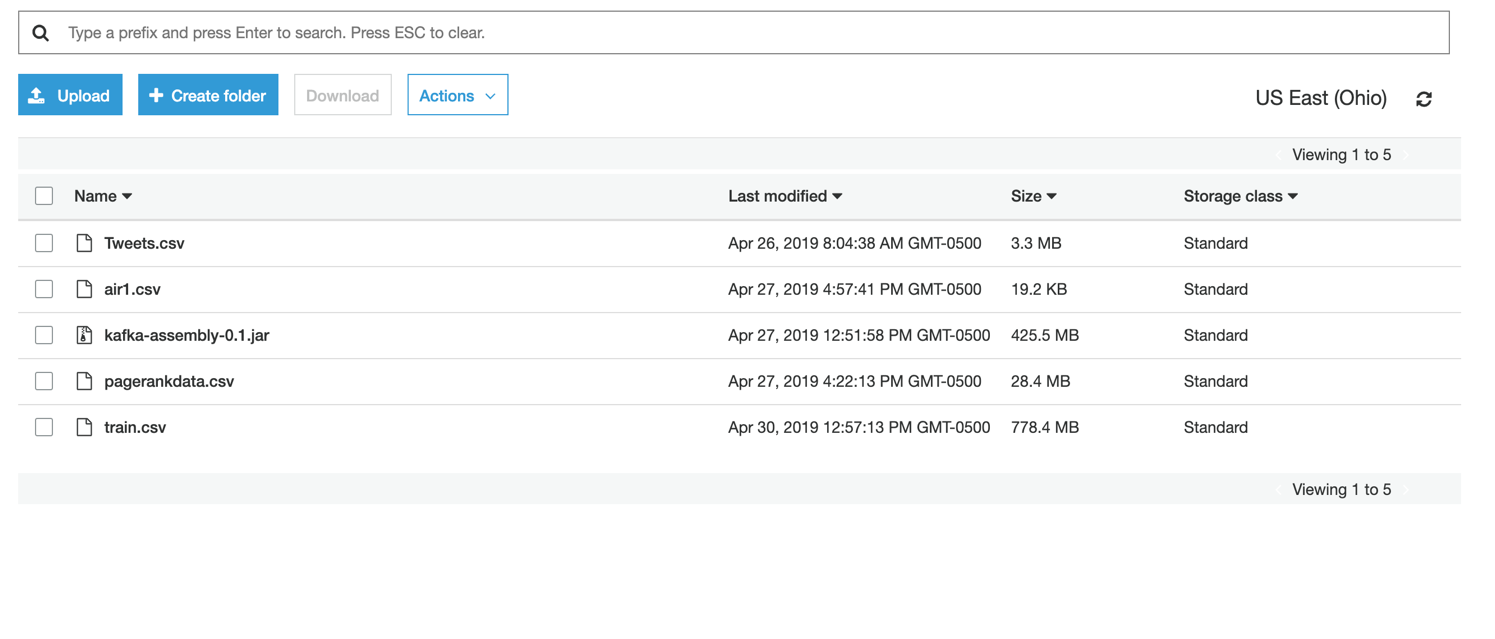
Big Data Project ReadME

**Steps for Jigsaw Unintended Bias in Toxicity Classification**

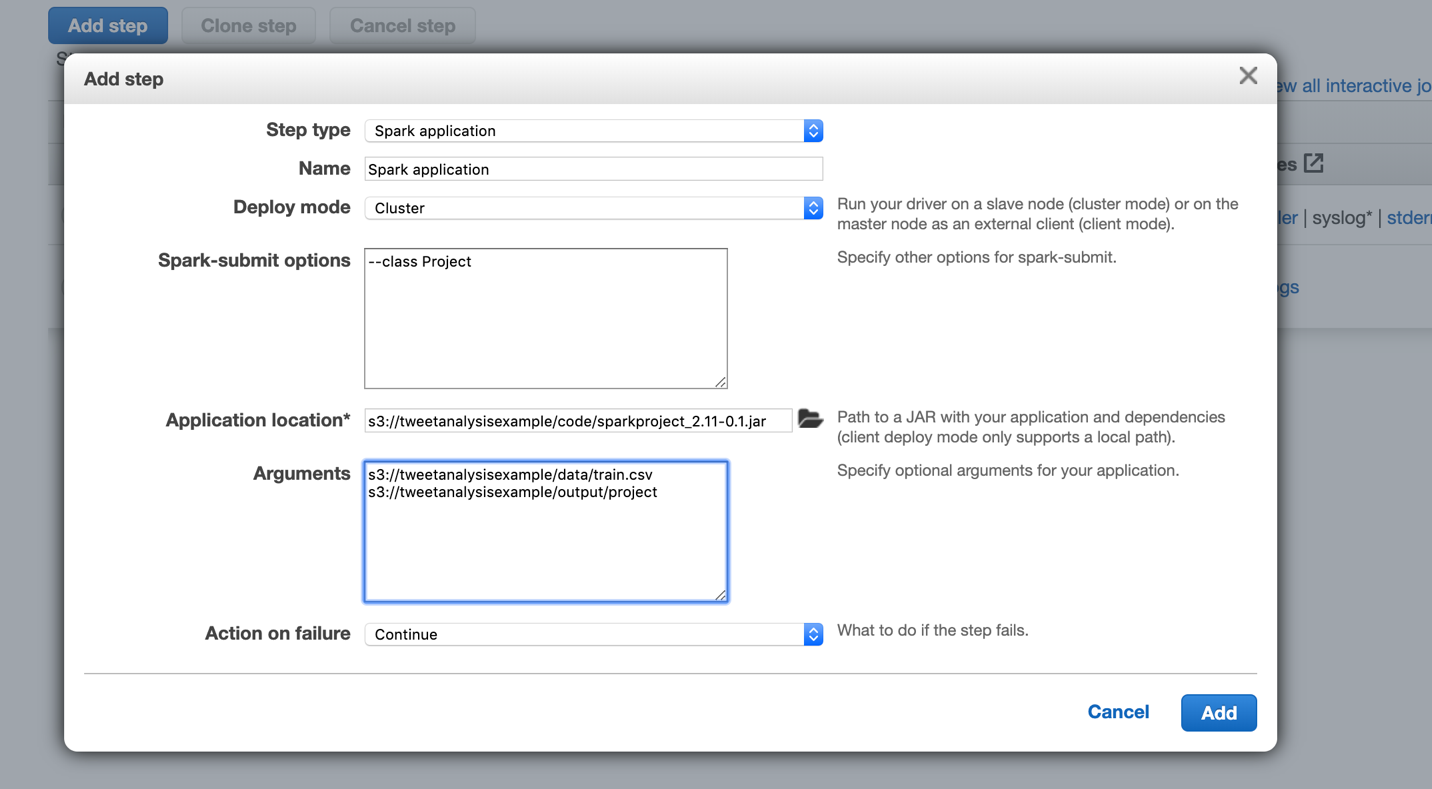
1. Open AWS S3 console and upload sparkproject\_2.11-0.1.jar to S3 bucket.



1. Add train data from jigsaw to S3 Bucket

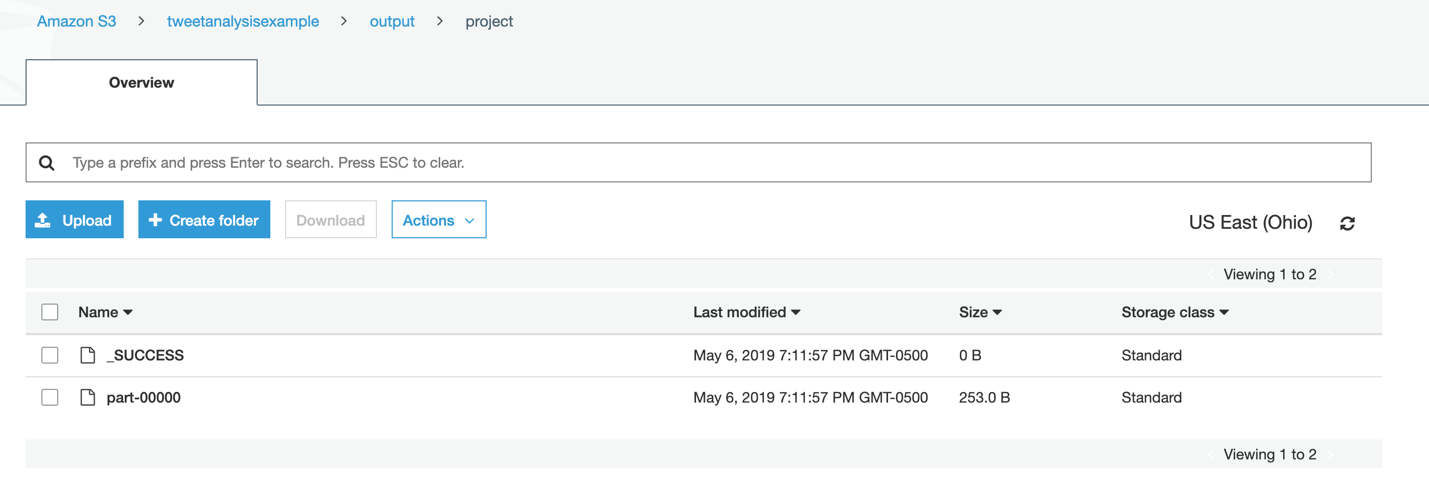


1. Open AWS EMR console select a cluster and start the application form the add step.



Run the class Project and pass the arguments as show.

1. Once the job is done we can see the result in S3 as Output/project folder



We can see the test results of the toxicity analysis of the given data with different Estimation metrics