**README**

Deliverables

1. Source code for ImagePrediction in Spark
2. Report
3. The log files (stderr) for the above programs on AWS
4. Output file(s) of the above programs run on AWS.

Running the source code on AWS:

1. Copy the src folder from the SourceCode folder.
2. Create a new Maven project in the IDE, add Scala nature, and add the extracted folder into the newly created Maven project and add the hierarchy to the Java build path. Ex: The project should have the following folder structure src/main/scala configured in the Java build path. Place the imagepredictionproject folder provided in the same hierarchical folder structure in the zipped file.
3. Copy the pom.xml and MakeFile to the newly created Maven Project.
4. Create a folder “input” in the project for the input files. In the input folder create three folders “training\_data”, “testing\_data”, “prediction\_data” and place the training, test and unlabeled data respectively into these folders.
5. For local execution, please uncomment the configuration line in the \*.scala program where the setMaster has argument local[\*] and comment out the one with argument yarn.
6. Please type the below commands to run the programs after installation of AWS CLI:

make cloud11 – for executing the training model on AWS

make cloud – for executing the prediction model on AWS

make alone – for local execution of the training model

make alone-prediction – for local execution of the prediction model

1. The output will be created in folder named as follows:

<input\_argument[1]>/prediction\_output – will contain the predictions for the unlabeled validation data

Report:

Report contains the analysis and implementation of the image prediction project along with answers to questions raised in the project questionnaire.

Log and Output Files:

BuchankrishnamurthyPradhan.csv – is the prediction file for the input unlabeled data.

The logs files are present in the following folders:

logs\prediction\_logs - syslog for the prediction program.

logs\training\_logs - syslog for the model training program.