Start by downloading spark locally

http://spark.apache.org/downloads.html

Spark is built on Scala 2.10. To build Spark and its example programs, enter directory and run:

./sbt/sbt assembly

This takes a long time.

To run the logistic regression example

./bin/run-example mllib.JavaLR data/mllib/lr\_data.txt 1 100

Standalone application for AMIDST

First we need to install maven and make sure that it builds. If you already have done this, you can skip the next step.

Install Maven: http://maven.apache.org/download.cgi

Put in ~/.profile:

source $HOME/workspace/core/doc/readMe/maven\_startup.sh

It is also important to put the hugin code in a certain place. I recommend to put it on core, (BUT DO NOT ADD to git repository as I am not sure that Hugin wants the jar file to be on the open source code.

Copy hgapi82\_amidst-64.jar to workspace/core, so Maven can find it (ref. pom.xml). Remember to add to project structure in IntelliJ.

Add hgapi82\_amidst-64.jar to your local .gitignore so that the file is not shown as untracted.

Create package, ignoring tests that fail:

mvn -Dmaven.test.skip=true package

I recommend that you all do not submit failing tests to core. Keep this on local projects till test are not failing. In essence, code should “always” build on master.

The final command to run JavaLR2 in the amidst project is:

spark-submit --class "eu.amidst.core.sparkExamples.JavaLR2" --master local[4] target/core-1.0-SNAPSHOT.jar