Memoria Trabajo Spark

Lo primero será conectarse a haoop:

$ ssh [mdat20076629@hadoop.ugr.es](mailto:mdat20076629@hadoop.ugr.es)

Para ver el estado de la máquina:

http://hadoop.ugr.es:8079

Abrimos Scala

$ /opt/spark-1.6.2/bin/spark-shell --packages JMailloH:kNN\_IS:3.0, mdat20076629:spark-infotheoretic-feature-selection:1.3.1, mdat20076629:spark-MDLP-discretization:1.2.1

Hacemos los import de las librerías

> import org.apache.spark.SparkContext

import org.apache.spark.SparkConf

import org.apache.spark.rdd.\_

import org.apache.spark.mllib.linalg.Vectors

import org.apache.spark.mllib.regression.LabeledPoint

import org.apache.spark.mllib.evaluation.MulticlassMetrics

import org.apache.spark.mllib.evaluation.BinaryClassificationMetrics

import org.apache.spark.mllib.classification.kNN\_IS.kNN\_IS

import utils.keel.KeelParser

import scala.collection.mutable.ListBuffer