FUNCTIONS OF CAMASSCLASS PACKAGE

WEDNESDAY, SEPTEMBER 28, 2005

msc.project.run Glossary • data set - features msc.biomarkers. read.mzXML by samples data msc.project.read msc.biomarkers read.mzXML data sets – data - For each data set: .read.csv msc.rawMS. with different msc.rawMS. features for the same read.mzXML and read.csv msc.classifier.test samples read.mzXML /or preprocessing – steps used to lower - Save data as binary - For each crossdimensionality of the validation iteration: data, performed - For each data set : - split labeled data into without use of class - Load binary data temporary train and test set labels • labels - describe sample.split msc.preprocess.run samples (for example "cancer" msc.baseline.subtract - train & test a classifier and "normal") bslnoff msc.classifier.run - operate on features msc.mass.cut msc.feature.select colAUC msc.mass.adjust msc.features.remove msc.peaks. write.csv msc.peaks.find msc.peaks. msc.feature.scale write.mzXML msc.peaks. write.mzXML read.csv - train R classifier (svm, msc.peaks.align nnet, rpart, etc.) on msc.peaks.clust msc.peaks. train samples read.mzXML msc.biomarkers. - predict labels of write.csv read.mzXML test samples msc.biomarkers.fill msc.rawMS. write.mzXML msc.copies.merge write.mzXML - compare predicted & true msc.sample. Legend labels of test samples to correlation calculate accuracy - For unlabeled data: Function - internal steps msc.classifier.run Concatenate preprocessed data sets Sub-function Sub-function Single data set; Predicted labels:

Classification labels:

mzXML

output

Classifier Accuracy

Page 1