



CARTXValidation Documentation

Description: Tree-based regression and classification with leave-one-out cross validation

Author: Joshua Gould, gp-help@broad.mit.edu

Summary:

CART builds classification and regression trees for predicting continuous dependent variables (regression) and categorical predictor variables (classification) (Breiman, et al., 1984). It works by recursively splitting the feature space into a set of non-overlapping regions (rectangles in the case of continuous features; subsets of values, in the case of categorical features), and by then predicting the most likely value of the dependent variable within each region.

A classification tree represents a set of nested logical if-then conditions on the values of the features variables that allows for the prediction of the value of the dependent categorical variable based on the observed values of the feature variables. A regression tree also represents a set of nested logical if-then conditions on the features variables, but these are used to predict the value of a continuous response variable instead.

References:

- Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and regression trees*. Monterey, CA: Wadsworth & Brooks/Cole Advanced Books & Software.

Parameters:

Name	Description
data.filename	The data file - .gct, .res
cls.filename	The class file - .cls
prediction.results.file	The name of the output file for prediction results

Output Files:

1. the prediction results

Platform dependencies:

Module type: Prediction
CPU type: any
Language: R