

Investigator

Load data

- Multiple sessions
 - Add/remove/copy
 - Save/load
- <https://data.mendeley.com/datasets/46htwnp833/1>
 - This data set contains NIR absorbance spectra of the wavelength range 309 - 1149 nm of mango mesocarp with corresponding dry matter content values (DM ~ ripeness).
- Load mango CSV dataset in LibreOffice Calc
- Copy section, add Clipboard source
- Show selecting class from table
- Change class attribute selection from menu
 - use ByName -> DM
- Load mango ARFF datasets (cal, val-ext)

Preprocess

- quick change attribute types
- remove attributes manual, undo
- split dataset
- batch filter (Remove, select atts for range)
- save datasets
- revert modified dataset

Data

- Sort by DM, plot column
- Select range
 - save selected
 - as new dataset
 - filter "Temp" using "low"
 - remove filter
- Select single row, plot row (13-last)

Classify

- select train/test set, select datasets
- additional attributes SampleID
- classifier: FilteredClassifier
 - (LWLSynchro: 250 neighbors, GPD)
 - (MultiFilter: Remove 12-last inv, PLSFilter)
- evaluate
- output generators: add Errors, regenerate
- click on crosses in Errors plots (-> SampleID col)
- result history: export output as zip
- output generators: add Predictions (-> SampleID col)
- result history: export output as zip

Classify

- copy tab
- switch to cross-validation
- on data tab:
 - show repeats
 - select two rows, ArrayStatistics: ArrayDifference 13-last
- use grouped CV on SampleID
- additional atts: SampleID
- evaluate
- use regular CV
- evaluate
- select both, compare output: errors

Flow editor

The Interface

- left: list of open flows
- center: current flow
- right: tabs (eg actor help, actors, current parameters)

Train/test - mango-1

- FileSupplier for mango ARFF
- load dataset
- set class
- cross-validate (with grouped)
- callable actor with classifier setup
 - FilteredClassifier
 - (LWLSynchro: GPD)
 - (MultiFilter: Remove 12-last inv, PLSFilter)
- output
 - eval summary
 - errors
 - predictions

CV - mango-2

- FileSupplier for mango ARFF
- load dataset
- set class
- cross-validate (with grouped)
- callable actor with classifier setup
 - FilteredClassifier
 - (LWLSynchro: GPD)
 - (MultiFilter: Remove 12-last inv, PLSFilter)
- output
 - eval summary
 - errors
 - predictions

Interactive - mango-3

- Let user select dataset, output dir
 - EnterManyValues (via template)
 - input_file
 - output_dir
- Store plot/text in output files
 - use Tabview
 - create vars for output files
 - CallableActorScreenshot for errors