Machine Learning

This page describes how to use the machine learning (aka mining) bundle.

Operators

- Clustering
- Classify
- Classification_learn
- FrequentItemset
- GenerateRules

Ensembles

You can simply combine all operators with other operators in Odysseus to create ensembles.

If we have, for examaple a stream with windspeed and power (speedandpower) and another one containing the windspeed (e.g. a forecast).

Then, it is possible to create different regressionfunctions, use them and weight the regression results by an aggregation.

One example:

```
/// create the first classifier by using SMO
smo = CLASSIFICATION_LEARN({
          class='power',
          learner = 'weka',
          algorithm = ['model'='SMO-REGRESSION']
        },
        speedandpower
      )
/// create the second classifier by using gaussian processes
gaussian = CLASSIFICATION_LEARN({
                class='power',
                learner = 'weka',
                algorithm = ['model'='GAUSSIAN-PROCESSES']
              },
              speedandpower
            )
/// create the thirs classifier by using a linear regression
linear = CLASSIFICATION_LEARN({
              class='power',
              learner = 'weka',
              algorithm = ['model'='LINEAR-REGRESSION']
            },
            speedandpower
          )
/// union them all into one stream
unioned = UNION(smo, gaussian, linear)

/// then, classify them - each tuple will be classified by using all three classifiers
ensemble = CLASSIFY(speed, unioned)

/// aggregate the clazz using average - which allows a weighted kind of voting
agg = AGGREGATE({
          aggregations=[
            ['AVG', 'clazz', 'powerForecast', 'DOUBLE']
          ]
        },
        ensemble
      )
```

Examples

to be added