

## Classify Tab

Using the Classify Tab, it is possible to load and unload classifiers and classify data. It is also possible to view the build report of the active classifier and set normalization options to be used when classifying.

The screenshot shows the 'Classify Tab' interface. At the top, there are five tabs: 'Data', 'Method', 'Options', 'View', and 'Cluste'. Below these, the 'Classifier' section contains a 'Load' button, a dropdown menu showing 'rf\_25', an 'Unload' button, and a 'Show Report' button. The 'Normalization Mode' section contains two radio buttons: 'As Dataset' (selected) and 'As Classifier'. An 'Apply' button is located to the right of the radio buttons.

### Controls

- **Load:**  
Load a classifier from file.
- **Unload:**  
Unload the currently selected classifier.
- **Show Report:**  
Show the build report for the currently selected classifier. See [Build Report](#).
- **Normalization Mode:** Determine whether to use the normalization mode determined by the selections in the [Method Tab](#) ("As Dataset") or use the normalization parameters stored in the classifier ("As Classifier").
- **Apply:**  
Classify the active data set using the selected classifier. Successful classification results in the creation of a new sample annotation. The name of the annotation is shown in a dialog after completed classification.

### Pre-requisites For classification

In order for the classification of data to be successful, a number of pre-requisites must be met. An error or warning will be displayed in case the classification fails, or if there is a risk of poor classification results. For successful classification, consider the following:

- **Normalization mode:**  
The normalization modes of the classifier (i.e. the mode used on the data set when building the classifier) should match the normalization mode of the data set being classified. If there is a mis-match between data set and classifier, a warning dialog is displayed. It is still possible to classify, but the results may be negatively affected by the mis-match. Also note that the Normalization option "As Classifier" should be used with care, since good classification performance using this option requires the data set under classification to have similar characteristics as the training data set. This option is mainly intended for situations where it is desirable to classify data consisting of a single sample, when normalization of the data therefore is not possible.
- **Missing variables in the data set**  
When building a classifier, the variables selected for use by the classifier is stored in the classifier. The dataset that is to be classified *must* contain at least those variables. Furthermore, the variables must be unique in the data set (this can be accomplished by using "collapse"). Note that when building or loading a classifier, a variable list containing the variables used by the classifier is created. This list can be used to verify that the dataset contains the correct (sub-)set of variables for classification.