**Filter** - pre classifiear.

Log[a]b = (log[10]b/log[10]a)

-information gain

"how much information about the

Classification the feature provides"

-variance of individual features

""

-correlation among features

**Wrapper** -

**Filter Methods**

Pros: Fast

Cons: Chosen filter might

not be relevant for a specific

kind of classifier.

Doesn’t take into account

interactions among features

Often hard to know how

many features to select.

**Wrapper Methods**

Pros: Features are evaluated

in context of classification

Wrapper method selects

number of features to use

Cons: Slow

**Embedded methods**

-Result is that most of the weights go to zero, leaving a small subset of the weights.

information gain. How it's used for feature selection.



