Data Mining with Weka

Using a filter

Ian H. Witten

Using a filter

Use a filter to remove an attribute

- Open weather.nominal.arff (again!)
- Check the filters
  - supervised vs unsupervised
  - attribute vs instance
- Choose the unsupervised attribute filter *Remove*
- Check the *More* information; look at the options
- Set *attributeIndices* to 3 and click OK
- Apply the filter
- Recall that you can *Save* the result
- Press *Undo*

Using a filter

Remove instances where *humidity* is *high*

- Supervised or unsupervised?
- Attribute or instance?
- Look at them
- Select *RemoveWithValues*
- Set *attributeIndex*
- Set *nominalIndices*
- Apply
- *Undo*

Using a filter

Fewer attributes, better classification!

- Open glass.arff
- Run J48 (trees>J48)
- Remove Fe
- Remove all attributes except RI and MG
- Look at the decision trees

- Use right-click menu to visualize decision trees

Using a filter

- Filters in Weka
- Supervised vs unsupervised,
  attribute vs instance
- To find the right one, you need to look!
- Filters can be very powerful
- Judiciously removing attributes can
  - improve performance
  - increase comprehensibility