

RMIT University
School of Science
COSC2110/COSC2111 Data Mining
Tutorial Problems Week 6

1. What is the main difference between the filter method and the wrapper method?
2. Describe how C4.5/J48 can be used for feature selection in the filter method and in the wrapper method.
3. Could the IBK classifier be used as filter for feature selection? Why or why not?
4. Consider a data set with 100 attributes, $A_{001} \dots A_{100}$. Assuming a 1-nearest neighbour classifier:
 - (a) Show the operation of the wrapper method with forward selection.
 - (b) Identify a combination of attributes that would be evaluated in an exhaustive search but not in a forward search.
 - (c) Show the operation of the wrapper method with backward elimination
 - (d) Show how the two methods can be combined in a bi-directional approach.
 - (e) How many times will the classifier be evaluated with an exhaustive search?
 - (f) How many times will the classifier be evaluated with a greedy forward search?
 - (g) How many times will the classifier be evaluated with a greedy backward search?
5. Explain how could OneR be used for feature selection.
6. Consider the use of OneR as the classifier in a wrapper. Is this sensible?