Wrappers

- The wrapper approach can be used to select subsets of features or do parameter tuning.
- The idea is to utilize the learning algorithm in the process of feature selection.
- Each set of features can be evaluated by a fivefold cross-validation utilizing the learning algorithm.
Wrappers

- The wrapper approach conducts a search space in the space of possible parameters.
- Let a feature set be a node in the search space.
- Clearly, neighboring states would be all states with one more or one less feature.
Four features $x_1, x_2, x_3, \text{ and } x_4$. With start state $s_1$. 
How to search?

• Use a best first search algorithm.
• Use either forward selection which is a search that begins with an empty set features.
• Or use backward elimination which is a search that begins with a full set features.
• Generate successor states by adding or subtracting n features.
• We can have special operators which add more than 1 or subtract more than one feature.
Algorithm

1. Put the initial state on the open list, the closed list is the empty set and Best <- initial state.

2. Let $v = \arg \max f(w_i)$ (get the state from OPEN with maximal $f(v)$).

3. Remove $v$ from OPEN, add $v$ to CLOSED.

4. If $f(v) - \text{eps} > f(\text{BEST})$, then BEST <- $v$. 
Algorithm (continued)

5. Expand $v$: apply all operators to $v$, giving $v$’s children.

6. For each child not in the CLOSED list, evaluate and add to the OPEN list.

7. If BEST changed in the last k expansions, go to 2

8. Return BEST.
Evaluation function

• You use the accuracy of a fivefold cross validation to evaluate a particular feature set.

• Forward selection tends to give small feature sets that are as accurate or more accurate with decision trees.

• Backward elimination tends to give larger feature sets which are even more accurate at the cost of running times that are five times longer than forward selection.
References

