WEKA
Waikato Environment for Knowledge Analysis

Attribute Selection

Prof. Pietro Ducange
Attribute Selection (1)

- Attribute selection involves searching through all possible combinations of attributes in the data to find which subset of attributes works best for prediction.
- Attribute selection consists basically of two different types of algorithms:
  - **evaluator** – determines the merit of single attributes or subsets of attributes
  - **search algorithm** – the search heuristic
Attribute Selection (2)
Attribute Selection: First Example (1)

Evaluates the worth of a subset of attributes by considering the individual predictive ability of each feature along with the degree of redundancy between them.

Searches the space of attribute subsets by greedy hill climbing augmented with a backtracking facility. Forward or Backward Search can be selected.
Attribute Selection: First Example (2)

=== Attribute Selection on all input data ===

Search Method:
  - Best first.
  - Start set: no attributes
  - Search direction: forward
  - Stale search after 5 node expansions
  - Total number of subsets evaluated: 12
  - Merit of best subset found: 0.887

Attribute Subset Evaluator (supervised, Class (nominal): 5 class):
  - CFS Subset Evaluator
  - Including locally predictive attributes

Selected attributes: 3,4 : 2
  - petallength
  - petalwidth
Evaluates the worth of an attribute by measuring the information gain with respect to the class.

Ranks attributes by their individual evaluations. A selection threshold can be fixed.
Attribute Selection: Second Example (2)

=== Attribute Selection on all input data ===

Search Method:
  Attribute ranking.

Attribute Evaluator (supervised, Class (nominal): 5 class):
  Information Gain Ranking Filter

Ranked attributes:
  1.418  3 petallength
  1.378  4 petalwidth
  0.698  1 sepallength
  0.376  2 sepalwidth

Selected attributes: 3,4,1,2 : 4
Attribute Selection: Wrapper Method (2)

Evaluates the worth of a set of attributes by using a specific classifier.
Attribute Selection: Wrapper Method (2)

--- Attribute Selection on all input data ---

Search Method:
- Best first.
- Start set: no attributes
- Search direction: forward
- Stale search after 5 node expansions
- Total number of subsets evaluated: 11
- Merit of best subset found: 0.947

Attribute Subset Evaluator (supervised, Class (nominal): 5 class):
- Wrapper Subset Evaluator
- Learning scheme: weka.classifiers.trees.J48
- Scheme options: -C 0.25 -M 2
- Subset evaluation: classification accuracy
- Number of folds for accuracy estimation: 5

Selected attributes: 4 : 1
- petalwidth
Attribute Selection as a Filter
Attribute Selection as a Filter (setting parameters)
Attribute Selection as a Filter (results)

Now we can switch to the classify module and perform a cross validation analysis....

...Is it a correct way to act?
Dimensionality of training and test data is reduced by attribute selection before being passed on to a classifier.
Selected attributes: 3,4 : 2
   petallength
   petalwidth

--- Evaluation on test split ---
--- Summary ---

<table>
<thead>
<tr>
<th></th>
<th>Value</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correctly Classified Instances</td>
<td>49</td>
<td>96.0784 %</td>
</tr>
<tr>
<td>Incorrectly Classified Instances</td>
<td>2</td>
<td>3.9216 %</td>
</tr>
<tr>
<td>Kappa statistic</td>
<td>0.9408</td>
<td></td>
</tr>
<tr>
<td>Mean absolute error</td>
<td>0.0396</td>
<td></td>
</tr>
<tr>
<td>Root mean squared error</td>
<td>0.1579</td>
<td></td>
</tr>
<tr>
<td>Relative absolute error</td>
<td>8.8979 %</td>
<td></td>
</tr>
<tr>
<td>Root relative squared error</td>
<td>33.4091 %</td>
<td></td>
</tr>
<tr>
<td>Total Number of Instances</td>
<td>51</td>
<td></td>
</tr>
</tbody>
</table>
Command Line Attribute Selection: An Example

- Generating new training and test files
  ```
  java weka.filters.supervised.attribute.AttributeSelection
  -E "weka.attributeSelection.CfsSubsetEval"
  -S "weka.attributeSelection.BestFirst -D 1 -N 5"
  -b
  -i <Training.arff>
  -o <TrainingSel.arff>
  -r <Test.arff>
  -s <TestSel.arff>
  ```

- Generating and testing a classifier
  ```
  java weka.classifiers.trees.J48 -t TrainingSel.arff -T TestSel.arff
  ```
The structure of the filter is based exclusively on the training data and test instances will be processed by the filter without changing their structure.
Exercise II

Perform the classification by using the three different meta classifiers (select a classification algorithm and three different attribute selection methods) and the Pima Diabetes dataset with a 5-fold cross validation.

Which is the best attribute selection method?