A Multi-dimensional Measure Function for Classifier Performance
by Niklas Laveson and Paul Davidson

a talk given by Adrian Schröter
contact me via schroeter@st.cs.uni-sb.de
What’s it all about?

• want to measure how good our model will perform

• want to compare different models
kNN
k – Nearest Neighbors

- Red circle: no swim
- Green circle: swim
kNN
k – Nearest Neighbors

- Red circles: no swim
- Green circles: swim
kNN
k – Nearest Neighbors

3 – NN
kNN
k – Nearest Neighbors

Intuitive and fast training but slow predictions.
Decision Tree

- Temperature?
  - Warm
    - Windy?
      - Yes
        - Swim
      - No
        - No swim
  - Cold
    - Windy?
      - Yes
        - No swim
      - No
        - Swim
Decision Tree

- **temperature?**
  - **warm**
    - **windy?**
      - **yes** (swim)
      - **no** (swim)
  - **cold**
    - **windy?**
      - **yes** (no swim)
      - **no** (no swim)
Intuitive and \textit{fast predictions} but \textit{slow training}. 

**Decision Tree**

- **Temperature?**
  - Yes
    - Intuitive and \textit{fast predictions} but \textit{slow training}.
  - No
    - Swim

**Legend:**
- Red: no swim
- Green: swim
SVM

support vector

support vector

separating hyper plane

red: no swim

green: swim
SVM

Not Intuitive and slow training but fast predictions.
Assessing Models

• Cross Validation (state of the art)
  – k-fold CV
  – Jackknife / Leave One Out CV
  – 10-fold CV
Parameter Tuning

- k - number of neighbors
- distance measure
- decision borders
- pruning factor
- penalization of outliers
- kernel parameters
Cross Validation

partition dataset into 10 folds
Cross Validation

Select 9 out of 10 folds.

- training data
- testing data
Cross Validation

Select 9 out of 10 folds.
Cross Validation

Select 9 out of 10 folds.
Cross Validation

Select 9 out of 10 folds.
Cross Validation

Select 9 out of 10 folds.
Cross Validation

Select 9 out of 10 folds.
Cross Validation

Select 9 out of 10 folds.
Cross Validation

Select 9 out of 10 folds.
Cross Validation

Select 9 out of 10 folds.
Cross Validation

Select 9 out of 10 folds.

After 10 iterations the results are averaged.
Disadvantage

relies only on subset fit
Subset Fit

training data

testing data
How To Measure?

Precision = TP / (TP + FP)

Recall = TP / (TP + FN)
More Measurements

• Similarity
  – similar observations should be classified similarly

• Simplicity
  – the simpler the model the better
Simplicity

Model A

Model B

Which model is simpler?
How To Measure?

Model A

Model B
Over Fitting

Model A

Model B
Measure Function

\[ a \times \text{subset fit} + b \times \text{similarity} - c \times \text{simplicity} \]
Conclusion

Lots of models and lots of parameters to tune.

State of the Art

Cross Validation

Measure Function

\[ a \times \text{subset fit} + b \times \text{similarity} - c \times \text{simplicity} \]