



FAKULTÄT FÜR
INFORMATIK

Advanced Topics in Machine Learning Programming Task

Jannis Becke, Christian Lausberger and Maik Riestock

Advanced Topics in Machine Learning

June 22, 2015

Data Preparation

49 Attributes, 101.000 Instances

Reasons for attribute removal:

1. **Uniqueness**
2. **Missing Values**
3. **Unequal Distribution**
4. **Single Attribute**

No instance was removed

Choice of Algorithms

Algorithms:

- Naive Bayes
- SVM
- YATSI
- Decision Table

Implementation

Choice of libraries:

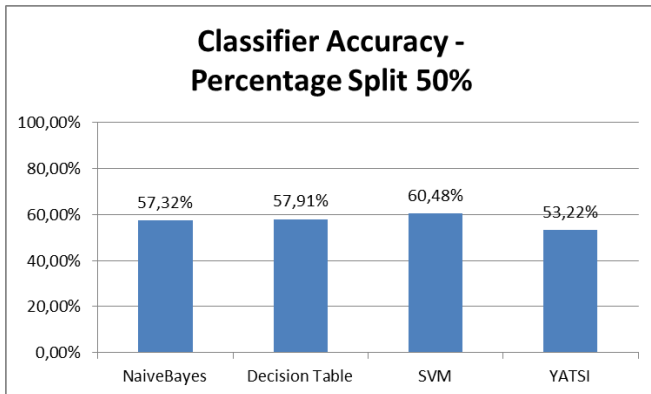
- weka
- libSVM
- collectiveClassifiers

Experiments

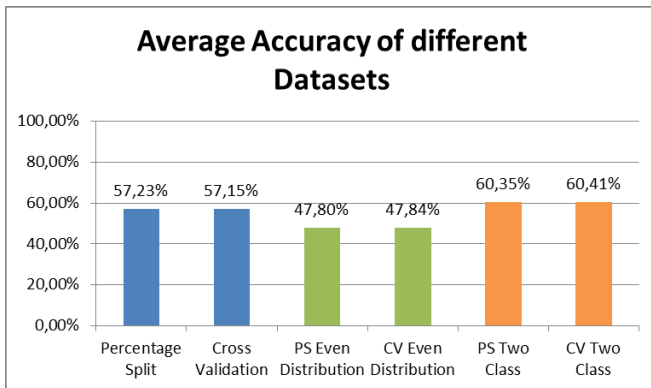
Settings:

- Percentage Split / Cross-Validation
- Accuracy of classification as quality measure
- two additional data sets:
 - even class dataset (equally distributed classification attribute)
 - two class dataset (join label > 30 , < 30)

Results



Results



Thanks for Your attention!

