

PA196: Project summary

Richard Trebichavský, 27. 10. 2015

Dataset: Wine Quality (<http://archive.ics.uci.edu/ml/datasets/Wine+Quality>)

- 12 attributes
- unbalanced (much more average samples than extremes)
- could be used for regression or classification

Goal

- create / train reasonable classifier, that will predict wine quality (last attribute, ranging from 1 to 10) from remaining 11 attributes

Planned approach

- balance / normalize dataset if necessary (e.g. if method takes to count quantity of observed values)
- perform feature selection (probably not all attributes are relevant)
- evaluate multiple methods for classification
 - compare with results published in original [paper](#)
 - Multilayer Perceptron
 - SVM
 - try some unused methods, i.e.:
 - kNN
 - LDA
 - ???