

# Supervised model comparison

---

Linus Sjöbro

2020-03-10

# Business understanding

- The problem
  - What properties decides the quality of a wine?
- Project goals
  - Evaluate datamining models
  - Accuracy of the models

- Wine quality dataset for red wines
  - 11 inputs
  - 1 output
    - Quality

- Remove weak correlations
  - Residual sugar
  - pH
  - Free sulfur dioxide

- Decision tree
- KNN
- Hold out
  - 30% training data
  - 70 % test data

# Evaluation

- Confusion Matrix
- Decision tree
  - Accuracy: 85.5 %
- KNN
  - Accuracy: 87.0 %
  - With K = 12

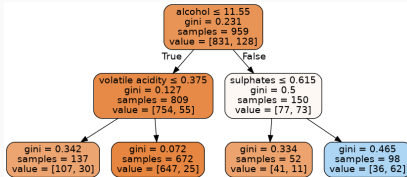


Figure 1: Decision tree

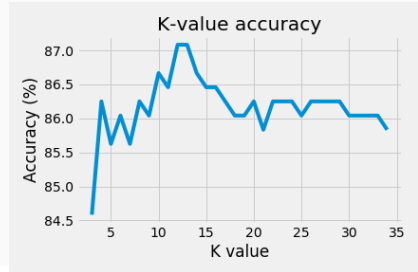


Figure 2: K value distribution

- Good and useful result
- Decision tree might be better

**Thanks for listening!**