

Information Mining - winter semester 2020/21**Exercise sheet 3****Exercise 1: Naive Bayes**

Define the Bayes rule and explain in your own words the prior probability. What do you understand under statistical independence? And also explain in your own words the terms evidence and event.

$P(A|B) = P(B|A)P(A)/P(B)$
independence : $P(A) = P(A|B)$, $P(B) = P(B|A)$
in $P(A)$, A is a event , in $P(A|B)$, B is a evidence

Exercise 2: Transformation of rules to decision trees

Following rules are given:

```
if speed = average and colour = blue
then car_from_cologne
if speed = fast and colour = red
then car_from_stuttgart
default car_from_wolfsburg
```

- (a) Create a decision tree based on these rules.

The attribute **speed** can have the values **fast**, **average**, **slow**. The attribute **colour** consists of the values **red**, **blue** or **other**.

- (b) How are decision trees transformed to rules?
1. each leaf is a rule
2. Pruning to remove redundant rules
- (c) Why is the transformation of rules to decision trees considered more complex than the transformation of decision trees to rules?

Tree cannot easily express disjunction between rules.
replicated subtree problem.

**Exercise 3: Rules and decision trees
in *RapidMiner***

- (a) Create two processes, which will create a decision tree and rules based on the given data¹. The attribute **label** is the class.

Hint: The creation of the decision tree and the induction of the rules may take several minutes on slower PCs.

¹dataA7.csv, please download this from Moodle