

## **Fragen von euch**

1. Was ist inductive Bias von candidate elimination und ID3?
2. Was ist das naïve am Naïve-Bayes-Klassifikator und welchen Nachteil hat es noch im Vergleich mit einer Support Vektor Maschine?
3. Beschreiben sie die intuitive Bedeutung der beiden Schritte des EM-Algorithmus.
4. Was ist der Unterschied zwischen Klassifikation und Regression und nennen sie jeweils ein Beispielfahren.
5. Was ist der Unterschied zwischen MAP und ML-Hypothese?
6. Was ist das Verhältnis zwischen AUC-Wert und ROC-Kurve?
7. Welche ROC-Kurve ergibt sich, wenn die Entscheidungsfunktion für alle Beispiele konstant ist?
8. Was ist der Unterschied zwischen überwachtem und unüberwachtem Lernen?
9. Wie kann man die SVM als MAP-Hypothese auffassen?
10. Gibt es frequentistische Wahrscheinlichkeiten, die man nicht auch als Bayesschewahrscheinlichkeit interpretieren kann?
11. Welche Optimierungsverfahren gibt es alternativ zum Margin-Perzeptron?
12. Spamfilter, Wort=Dimension, Hauptproblem dieser Merkmalsrepräsentation?
13. Trainieren auf Trainingsdaten, Fehler schätzen auf Testdaten, danach noch mal trainieren auf allen Daten. Optimistisch oder pessimistisch?
14. Vergleich, ML, MAP, Bayes-Hypothese, Welche Voraussetzung müssen erfüllt sein, jeweils?
15. Erläutern sie den Unterschied zwischen primaler und dualer SVM?

## **Fragen von mir**

16. Wie sieht die Precision-Recall-Kurve eines perfekten Klassifikators aus?
17. Bei welchem einfachen Beispiel versagt der Roccio-Klassifikator aber ein lineares Perzeptron nicht?
18. Wie groß sind Support und Confidence (genau, höchstens oder mindestens) von  $AB \Rightarrow CD$  wenn  $s(AB)=2$ ,  $s(BC)=3$  und  $s(CD)=4$  ist?
19. Was ist der induktive Bias und was der Language Bias von Candidate Elimination?

20. Was ist die "naive" Annahme beim Naive-Bayes Klassifikator und was sind Vor- und Nachteile von Naive-Bayes verglichen mit einer Support-Vektor-Maschine?
21. Warum ist eine Fehlerschätzung mit einmaligem Aufteilen in Trainings- und Testdaten pessimistisch und ungenau? Gibt es Alternativen? Welche?
22. Was ist die Entropie von zwei Münzwürfen, wenn die Münzen nicht verbogen sind?
23. Welchen Vorteil hat SLIQ gegenüber dem einfachen C4.5 Algorithmus?
24. Was ist der induktive Bias und was der Language Bias von C4.5?  
Wie sieht die Entscheidungsgrenze von C4.5 grafisch aus?
25. Unter welchen Umständen ist ein Rocchio-Klassifikator besser geeignet zum Lernen eines Klassifikators als ein lineares Perzeptron?
26. Was ist der Unterschied zwischen einem primalen und dualen Perzeptron?
27. Was ist ein Large-margin Klassifikator? Welchen Vorteil hat ein Large-margin Klassifikator gegenüber dem normalen Perzeptron-Algorithmus?
28. Wie würden sie den Fehler eines Zeitreihenmodells messen? Kann man hier Cross-Validation verwenden?
29. Wie viele semantisch unterschiedliche Entscheidungsbäume gibt es, wenn unsere Zielklasse binär ist und wir  $k$  binäre Eingabeattribute haben?
30. Was versteht man unter einer Soft-Margin-Support-Vektor-Maschine, was ist der Unterschied zu einer Hard-Margin-Support-Vektor-Maschine?
31. Man kann  $k$ -Means auch als spezielle Variante eines EM-Algorithmus betrachten, was sind hier der E- und M-Schritt?
32. Welche intuitive Bedeutung hat der AUC-Wert eines Klassifikators?
33. Was ist der F-Measure und welche intuitive Bedeutung hat er?
34. Was ist der Unterschied zwischen einer Entscheidungsfunktion und einem Klassifikator?
35. Warum ist die Fehlerschätzung bei Training-and-Test pessimistisch und warum ungenau?  
Was kann man verbessern?
36. Was ist der Vorteil von ID3 gegenüber Candidate Elimination?
37. Geben Sie eine intuitive Beschreibung für Precision und Recall.