

# Fragenkatalog ASS

1. Was ist Machine learning (ML) im Allgemeinen? Wovon hängt ein ML Algorithmus am stärksten ab und wodurch kann dieser daher verbessert werden? Nennen Sie Problemstellungen, die durch ML gelöst werden können.
2. Erkläre die MAPE loop. Was kann dadurch vermieden werden? Was passiert in den einzelnen Schritten?
3. Was ist ein Peer-To-Peer Netzwerk? Welcher Peer hat besonderen Stellenwert in einen P2P Netzwerk und wer übernimmt die Organisation? Welche Vorteile hat P2P gegenüber Client-Server Systeme?
4. Nennen Sie je einen Vorteil, Nachteil von P2P System. Erklären Sie diese konkret an einem Anwendungsfall und stellen Sie einen Vergleich zu Client-Server Architektur.
5. Was ist Ziel einer Distributed Hash Table? Vergleichen Sie die Anzahl der Zugriffe und 1 Vor-+ 1 Nachteil mit P2P Systemen ohne DHT.
6. Wie funktioniert DHT (Chord) und welche Funktionen gibt es?
7. Wie funktioniert der Insert in DHT? Skizzieren Sie die Schritte mit einem  $m = 2$  Ring.
8. Wie funktioniert Join in DHT und welche Möglichkeiten gibt es als Node eine DHT wieder zu verlassen?
9. Wie funktioniert Destination Sequenced Distance Vector? Welche Daten müssen pro Node gespeichert werden?
10. Wie funktioniert Link State Routing? Warum sind berechnete Routen in LSR immer optimal?
11. Wie funktioniert das optimierte Link-State-Routing? Was wird in OLSR optimiert und wodurch? Skizzieren Sie ein kleines Beispiel indem ein OLSR sinnvoll verbessert werden kann (im Gegensatz zu LSR).
12. Wie funktioniert ad-hoc on-demand distance vector? Was unterscheiden AoDDV grundsätzlich zu DSDV?
13. Was ist Dynamic Source Routing? Wodurch unterscheidet es sich von anderen besprochenen Mechanismen? Erklären Sie das Prinzip von RREQ und RREP.
14. Nennen Sie 2 Strategien für Kommunikation und vergleichen Sie diese. Nennen Sie je ein Beispiel für eine Anwendung und nennen Sie ein Beispiel, wo Sie die Kommunikation nicht einsetzen würden. Begründen Sie ihre Antwort.

15. Welche Klassifizierung gibt es bei Data-Space Kommunikation? Geben Sie 2 Beispiele (gesamt) mit Begründung warum diese Kategorie an Kommunikation hier gut passt.

16. Welche JavaSpaces Operationen werden unterschieden? Erklären Sie wie die CRUD (Create Retrieve Update Delete) Actions auf einen Space durchgeführt werden können.

17. Was sind die Vor- und Nachteile von Code-Migration? Welchen SPOF (Single Point of Failure) kann man hier identifizieren?

18. Wie ist der klassische Ablauf einer Service Discovery? Welche Integrationsstrategien für Ad-Hoc Kommunikation von Geräten gibt es?

19. Was ist Data Mining? Was ist der Unterschied/Zusammenhang zu Machine Learning?

20. Woraus besteht der Lernprozess in Machine Learning? Wie kann das Model verbessert werden?

21. Was bedeutet Bias und was bedeutet Over-Fitting? Wie stehen die beiden Werte zueinander? Wie kann man Overfitting verhindern?

22. Welche Schritte im Prozess sind anzuwenden um mit machine learning Algorithmen zu arbeiten?

23. Was sind Features und Examples in machine learning? Beschreiben Sie diese und erläutern Sie im Anschluss anhand eines konkreten Beispiels.

24. In welche 2 Kategorien können machine learning Algorithmen eingeteilt werden? Nennen Sie je ein Beispiel für eine konkrete Anwendung.

25. Welches Ziel hat die Regressionsanalyse? Gehen Sie sowohl auf die Regression im Allgemeinen als auch auf die lineare Regression ein. Was passiert bei der Berechnung linearer Regression wenn bei den Messdaten exakt 1 Messfehler auftritt und statt einem durchschnittlichen Wert ein extrem hoher Wert gemessen wird (Beispielsweise statt 15 kommt 15.000.000)?

26. Was ist Ziel des OLS (Ordinary least Square; de: Methode der kleinsten Quadrate) – Vorgehens? Erklären Sie dabei auch den Begriff Residuum.

27. Was besagt der „correlation coefficient“? Erklären Sie dabei

a) welchen Sinn der Korrelationskoeffizient hat,

b) welche Einheit,

c) welchen Wertebereich,

d) was die Extremwerte und 0 bedeuten,

e) unter welchen Randbedingungen er Anwendung findet und

f) ob der Korrelationskoeffizient zwischen X und Y

gleich dem Korrelationskoeffizienten zwischen Y und X ( $r_{xy} = r_{yx}$ ).

28. Was ist „Entropy“? Erklären Sie in Folge die Entropie anhand des konkreten Beispiels „Würfeln mit Spezialwürfel“. Der Spezialwürfel hat dabei 2 blaue Seiten und 4 rote. Erklären Sie „Information gain“.

29. Warum sollte man einen Entscheidungsbaum kürzen (=„pruning“)? Gehen Sie auf Pre- und Post-Pruning Algorithmen ein.
30. Was ist die grundlegende Idee von Regression Trees? Was wird wie partitioniert? Warum heißt es "Recursive Partitioning"? Was haben Elemente innerhalb einer Partition gemeinsam? Was ist der Unterschied (im Prinzip, in der Vorgehensweise) zwischen Regression und Decision Tree?
31. Wie bildet man Klassen für k-NN Algorithmus und wie funktioniert er? Wie werden unklassifizierte Examples klassifiziert? Worauf ist zu achten wenn man k klein oder k groß ansetzt? Welche Ansätze gibt es, um den k-NN Algorithmus zu verbessern?
32. Was versteht man unter k-fold cross validation? Was versteht man unter data splitting?
33. Wozu gibt es AOP im Allgemeinen? Was wird dadurch gelöst, was in "herkömmlichem" OOP nur mit sehr viel Aufwand verbunden wäre.
34. Wozu dient ein Aspekt und aus welchen Bestandteilen setzt er sich zusammen?
35. Was ist ein Aspekt Weaver? Welche Strategien gibt es und wodurch unterscheiden sich diese?
36. Welche technischen Komponenten hat AspectJ? Nennen Sie ein weiteres AOP-Produkt.
37. Nennen Sie einen Vor- und einen Nachteil von AOP und dessen wichtigste Begriffe/Konzepte mit deren Zusammenhang zueinander.
38. Welche Arten von Advices gibt es? Welche Bestandteile hat ein Advice?
39. Was ist der Unterschied zwischen einem Call und einem Execution Pointcut? Kann überall wo ein Call Pointcut angewendet wird, auch ein Execution Pointcut angewendet werden? Begründen Sie Ihre Antwort.
40. Was unterscheidet einen Advice von einer Methode? Welche Unterschiede haben before/after Advices zu around Advices?
41. Welche Methoden gibt es für Strukturänderungen?