

Gedächtnisprotokoll: Praktische Informatik

- Datamining Algorithms (SS07), nach Folienskript (Buch von Kamber ist auch nützlich)
- Modelle der Datenexploration (WS06/07), nach Folienskript
- Introduction to Artificial Intelligence (WS06/07), Russel/Norvig Buch, Kapitel nach Absprache
- Prüfer: Prof. Seidl
- Datum: 28.02.08
- Dauer: ca. 45min, 15min pro Gebiet
- Hier sind jetzt nur die Fragen in etwa...teilweise sind die Fragen hier durch Antworten schon vorweg genommen.

Datamining:

1. Was macht das Datamining? Wo gliedert sich Datamining in den KDD Prozess ein?
2. Was ist Clustering, was ist Klassifikation? Worin unterscheiden diese sich?
3. Was kann man da noch so machen -> Assoziationsregeln
4. Was genau macht man nun da?
5. Was ist so eine Assoziationsregel?
6. Was ist hierbei Support und Confidence?
7. Welche Algorithmen gibt es da so? -> Apriori, FP-Growth
8. Apriori Schritt machen, Einzelheiten erklären.
9. Frequent Patterns aus dem Ergebnis ableiten.
10. Wieviele gibts zu einem k-Itemset?
11. Wie kann man Apriori verbessern?
12. FP-Growth erklären.
13. Was ist das Problem bei Subspace-Clustering?

Modelle der Datenexploration:

1. Was macht man hier so? -> Ähnlichkeitsmodell und Anfragen.
2. Was für Anfragen gibt es denn so?
3. Was genau ist hier "Ähnlichkeit"?
4. Was für Ähnlichkeitsmodelle hatten wir denn?
5. Bilder, wie kann man die denn vergleichen?
6. Histogramme, was ist das?
7. Wie vergleicht man die jetzt?
8. Quadratische Form, was ist das denn hier genau?
9. Was ist dabei die Ähnlichkeitsmatrix?
10. Wie kann man die bestimmen?
11. Wie sieht die Matrix aus wenn man erstmal nichts über die Ähnlichkeiten weiß?
12. Cross-Bin-Similarity was kann man da machen?
13. EMD, wo ist da der Unterschied zur Quadratischen Form? Distanz vs. Ähnlichkeit.
14. Quadratische Form, was gibts denn da für Filter und wie bekommt man da den Zusammenhang zu Distanzen hin?
15. Filter, was können die denn so?
16. Lowerbounding-Property.
17. Kriterien für gute Filter.

Introduction to Artificial Intelligence:

1. Hier gibts ja auch Suchprobleme, wie sind die denn definiert?
2. Wie sieht denn der Suchraum so aus?
3. Vergleichen sie das mal mit der Suche in Datenbanken.
4. Was für Suchen gibts denn so?
5. Uninformierte Suchen...
6. Komplexitäten, Optimalität, Vollständigkeit, alles mit Begründung.
7. Iterative Deepening, Komplexität, warum ist die doch nicht so schlecht?
8. Welche Level werden hier warum wieoft expandiert? (Wurzel d-mal, Ziellevel 1-mal)
9. Spiele, was ist das, was kann man hier machen?
10. Minmax, Baum usw. Explizit erklären was hier Nullsummenspiel bedeutet.
11. Baum ist viel zu groß, was kann man hier machen?
12. Suche vorher abbrechen, Bewertungsfunktion für "neue" Endzustände, Zusammenhang zu Schach.
13. Alpha-Beta-Pruning, erklären, Beispiel, Zusammenhang zu Schach.
14. Wie sieht der Baum nachher aus?
15. Wann genau macht man das Pruning? (Nach Bewertung der "neuen" Endzustände)

Prüfung an Sich ziemlich Angenehm und Prof. Seidl als Prüfer nur zu Empfehlen. Trotz hier und da einiger Holprigkeiten ist nachher eine 1,7 rausgesprungen, Minmax und der Nullsummenkram ist mir irgendwie zum Verhängnis geworden, urgs.