

# **Diplompruefung praktische Informatik**

pruefer: seidl

vorlesungen: Datamining nach Folienskript, Datenbanken nach kempler/eickler, KI nach Norvig/Russel

note: 1.0

dies ist eine grobe zusammenfassung des pruefungsverlaufs und damit nicht vollstaendig. Es gab relativ viel freie diskussion, die prof seidl auch zu gefallen schien

## **DataMining:**

S: fangen wir mal an.. kdd prozess habe ich eigentlich keine lust drauf.. aber mh, sagen sie mal die großen teilgebiete des dataminings

i: classification, clustering, association, generalisierung

s: gut, wie unterscheidet man classification und clustering

i: unterschied erklärt

s: ok, was macht man denn so ganz allgemein bei classification

i: man trainiert parameter eines modells, mit dem man dann ungesehene daten klassifiziert

s: ok, was fuer verfahren/modelle hatten wir denn da und was fuer parameter werden da trainiert?

I: bayes, nn, decisiontrees, svms

s: ja, wie funktionieren denn svms?

I: man sucht ebene die die daten trennt, maximum margin hyperplane, die ist definiert durch die vectoren die auf dem margin liegen

s: ja, und was ist dann das modell bei svms?

I: die supportvectoren mit den gewichten

s: ja, also, eigentlich die ganzen trainingsdaten mit gewichten

i: ja, aber supportvectoren sind dann die mit gewichten ungleich 0

s: ja, stimmt. Kommen wir mal zu anderen klassifikatoren.. wie ist denn die entscheidungsregel bei bayes?

I: aufgeschrieben, eigentlich alles erklärt mit wie man die parameter schätzt bei welchen annahmen/ attributen. Abhängig/unabhängig, hier hat er ein paar mal Kleinigkeiten nachgefragt.

S: und wie sehen Entscheidungsbäume aus?

I: erklärt

s: und was trainiert man bei kNN?

I: eigentlich ja nichts, lazy learner erwähnt, verschieben der evaluation der Trainingsdaten auf testphase

s: gut, wir hatten dann ja noch generalisierung.. was macht man denn da?

I: OLAP und AOI erklärt

s: machen sie mal OLAP mit SQL

i: erklärt welches statement was macht. Seidel hat sich dann selber verhaspelt weil er wohl schon an selection bei RA gedacht hat und dann meine SELECT statements von SQL in den WHERE teil gesteckt hat :)

## **Datenbanken**

s: ja, gut, kommen wir mal zu Datenbanken, welche Modelle gibt's denn da so?

I: blabla

s: wie kann man das denn modellieren?

I: ER und UML erwähnt

s: ja, mh, wann nimmt man denn welches? Achja, das hatten sie ja eigentlich schon erwähnt. mh. Kann man denn UML in ER umwandeln?

I: mh, ueberlegt.. jein, man kann das machen, aber verliert halt die Methoden.

S: ja, noch was?

I: ja, man muss ja irgendwie Schlüssel fuer das ER Modell finden

s: richtig, was ist denn ein Schlüssel ueberhaupt?

I: Menge von Attributen, so dass alle Attribute der Relation davon funktional abhängig sind

s: richtig, das können auch mehrere Attribute sein. Können das denn auch alle sein?

I: ja, aber Sinn macht das ja nicht.

An irgendeiner Stelle hatten wir hier noch ueber kuenstliche Schlüssel geredet. Und wann man die braucht. Habe da was von Vorname, Nachname erzaehlt und dass das ja nicht eindeutig ist, war dann

ok.

s: ok, wie stellt man denn ganz allgemein anfragen an die datenbank?

I: mit einer DML, zb SQL

s: stimmt, sql haben wir ja vorhin schon gemacht, kann man denn mit SQL noch mehr machen?

I: ja, man kann damit auch das datenbank-schema beeinflussen, zB mit create table

s: stimmt, wie heisst das dann?

I: data description language

s: naja, data defenition language eigentlich.. welche anderen möglichkeiten ausser prozeduralen anfragen hat man denn noch so?

I: relationales kalkuel, tupelkalkül erklärt und domänenkalkül, aber nur oberflächlich(variablen werden an xxx gebunden)

s: ja, schreiben sie mal einen ausdruck aus dem tupelkalkül auf.

I: hinschreib  $\{t \mid P(t)\}$ , erklärt wie eine formel aufgebaut ist. Dabei hat mich seidl darauf hingewiesen, dass ueblicherweise auf der linken seite tatsächlich nur eine variable steht, da wollte ich schon widersprechen, habs aber nicht gemacht. Braucht man das nicht fuer kreuzprodukte dass da mehrere variablen stehen?? egal)

s: ja, da haben sie ja schon die logischen operatoren stehen, kommen wir mal zu KI

## ***Kuenstliche Intelligenz***

s: wie ist denn da so ein logischer ausdruck aufgebaut? Was sind syntax und semantik?

I: mh, was ueber syntax und semantik von aussagen logik erzaehlt und prädikatenlogik erwähnt.

S: nachbohr

I etwas mehr erzählt. Syntax und semantik von AL sind ja einfach, bei prädikatenlogik wollte er auf trägermenge hinaus, was ich auch irgendwann gesagt habe. Dann wollte er nochmal die semantik von funktionen wissen, da hab ich das left leg beispiel aus dem buch erzählt weil mir nix eingefallen ist. Letztendlich meinte er dann dass die semantik dann wieder terme sind oderso.

S: ok, es gab da ja ein anderes großes thema in der KI, suche, wie ist denn so ein suchproblem definiert?

I: start, uebergang, pfadkosten, zieltest

s: ja, und die uebergangsfunktion und start zusammen definieren dann den suchraum, gut. Wir hatten ja auch spiele, was fuer spiele denn genau?

I: zweispieler, abwechselnd, perfect information, nullsumme

s: was ist denn ein nullsummenspiel?

I: erklärt

s: gibt es denn auch andere spiele?

I: klar, bischen rumgebeispiel

s: ja, sonst köntne es ja in der industrie keine win-win situationen geben

i: naja, aber man weiss ja nicht ob da nicht doch wer anders draufzahlt und man im endeffekt doch wieder ein nullsummenspiel hat

s: mh, stimmt. Wenn was gutes passiert gibt's vermutlich auch was schlechtes als auswirkungen.

I: ja, und meine oma hat gesagt dass nichts so schlecht ist, dass es nicht auch was gutes hat

s: ja, stimmt, wobei man nicht weiss ob das eine nicht das andere ueberwiegt

i: ja, aber das geht ueber das thema jetzt hinaus

s: ja, was sucht man denn bei spielen?

I: optimale strategie erklärt

s: und wie macht man das?

I: baum aufgemalt, entscheidungen erklärt, alpha-beta pruning erklärt und auf cutofftest hingewiesen. Da dann ein bischen rumerklärt mit utility funktion, quiescent states und horizont problem. Das war dann irgendwann genug und ich durfte raus

### **fazit:**

sehr angenehme pruefung. Prof seidel hat einen angenehmen pruefungsstil und stellt fragen in einer weise, dass man oft selber entscheiden kann wie sehr man ins detail geht. Ich habe meist erst sehr stark zusammenfassend geantwortet und bin dann auf nachfrage ins detail gegangen. Zusätzlich angenehm waren so nachdenk fragen wie UML in ER umwandeln, wo er selber scheinbar auch gerne mitgedacht hat.