

Prüfungsdatum: 21.08.2008.

Dauer: 45 Minuten.

Vorlesungen: Introduction to Databases (4 SWS), Indexstrukturen in Datenbanken (4 SWS), Data Mining Algorithms (4 SWS).

Note: 1.0

1) Datenbanken: :

- Wenn man eine Datenbank erstellen möchte, wie geht man überhaupt vor?
- Wir haben ja verschiedene Datenmodelle, welche von denen können Sie nennen?
- Sie haben gesagt, dass heutzutage meistens Relationale Datenbanken und OO-Datenbanken benutzt werden. Welche Gemeinsamkeiten existieren denn?
- Was gibt's denn noch Unterschiede zwischen den Relationellen Datenbanken und Objekt Orientierten Datenbanken?
- Wie können wir denn die Daten in den Datenbanken zugreifen?
- Wie sieht eine allgemeine Form von SQL (SELECT) aus?
- Wie kann man dies denn erweitern? (SELECT DISTINCT, GROUP BY, HAVING sind die einfachste Statements, die man dazu benutzen kann.)
- Wozu benutzt man überhaupt SELECT DISTINCT?
- Wie funktioniert GROUP BY und HAVING? Was muss man darauf achten, dass GROUP BY ohne Fehler funktioniert? (Die Attribute, die ohne Aggregatsfunktionen in SELECT Statement vorkommen, sollen auch in GROUP BY Statement erscheinen.)
- *Malt ein Relationsschema Studenten mit MatrNr, Name, Fach, Semester, und noch ein Attribut. * Schreiben Sie nun ein GROUP BY Statement.

2) Indexstrukturen in Datenbanken

- Brauchen wir Indexstrukturen überhaupt? Warum?
 - Was ist die Zeitkomplexität für N Datensätze ohne Indexstrukturen?
 - Sie haben gerade gesagt, dass lineares Scan normalerweise länger dauert als Suche durch Indexstrukturen. Es gibt aber ein Verfahren, das auf linearem Scan basiert ist. Wie nennt man dieses Verfahren? (VA)
 - Ist VA ein Dimensionsreduktion? (Nein)
-

- Worauf basiert es denn?
- Wie kann es sein, dass der Vector Appr. File kleiner ist als Vector File? Geben Sie ein Beispiel dafür.
- Sie haben gerade gesagt, dass es hierarchische Strukturen wie Bäume existieren. Was gibt es denn noch für 1-D Daten?
- Ja, Hashing gibt's. Aber wir haben ja noch was!
- Bitmap Index! Wozu brauchen wir es denn?
- Wir haben die Werte Montag bis Sonntag für ein Attribut in einer Relation. Was ist die Idee von standarden und Bereichskodierten Bitmap-Indexen?
- Schreiben Sie die Anfrage für "Mittwoch".
- Schreiben Sie die Anfrage für "Dienstag bis Samstag".
- Welche Bitvektoren werden gelesen? Warum?
- Wie viele Bitvektoren haben wir denn für ein Attribut?
- Schreiben Sie die gleiche Anfragen bzgl. Bereichskodierten Bitmap-Index.
- Was ist der Unterschied zwischen den beiden Strukturen (bzgl. Effizienz und Anfragentypen)?

3) Data Mining Algorithms

- Ok, you would like to do this part in English, right? Let's start. Can you explain the main steps of Data Mining?
 - What is actually the difference between Clustering and Classification?
 - In the classification part of the lecture, we have seen that we can use decision trees. How does the structure look like and what are the elements of the tree? (Here, I have explained everything on the example of "Shall we play tennis today?".)
 - Ok, how does the training phase work for decision trees?
 - Is this lazy learner?
 - What is lazy learner?
 - How does the classification work on a decision tree?
-

- We have one more interesting classification method: NN classifiers. Can you tell me about the decision rules? (Here, I again explained everything on an appropriate example gathering difference outputs from each decision rule.)
- What does this classifier do in the training phase?
- So, what is the difference between decision trees and NN-classifiers regarding their training phase?
- What is generalization?
- How does the method Attribute Oriented Induction work?
- What are the thresholds used for this method? (This was a long discussion as I have talked about 2 thresholds but I guess these were not comprising the one that he wanted to hear about. If I were you, I would absolutely have a look at them in detail before the exam day!)

Da ich das Protokoll nicht direkt nach der Prüfung, sondern während des Flugs in Urlaub geschrieben habe, konnte ich mich nicht an alle Fragen erinnern. Ich wünsche euch viel Erfolg und viel Glück bei euren Prüfungsvorbereitungen.
