

Anwendungsprobleme

Grundideen

Vorteile

Nachteile

Begrenzungen

Prüfungsschwerpunkte, Teil II

Data Mining

- Aufgabenstellung und Teilschritte
- Datentypen und Datenvorverarbeitung
 - Ausreißerbehandlung
 - Dimensionalitätsreduktion
- Klassifikation
 - Abstands-basierte Verfahren
 - Evaluation
- Clustern
 - Clustern durch Partitionierung
- Assoziationsregeln

11. Data Warehouses and OLAP

- Einsatzgebiete und Besonderheiten für Data Warehouses
- Dimensionsmodellierung, Datenbankschemata für OLAP
- OLAP Operationen

Prüfungsschwerpunkte, Teil II

13. Semi-Strukturierte Daten

- Markup-Sprachen, was und wozu?
- Dokumenttypdeklaration
- Schemadefinition
- Abfragesprachen

14. Document Retrieval

- Probleme
- Vektorraummodell
- Evaluation