

Klausurthemen

- MapReduce
 - Einfache Algorithmen als MapReduce-Prozedur
 - Replikationsrate und Reducer-Size
 - Map-Tasks, Reduce-Tasks, Combiner
- Clustering: K-Means und hierarchische Verfahren
- Dimensionsreduktion: Hauptkomponentenanalyse und Singulärwertzerlegung
- Empfehlungssysteme: Inhaltsbasierte Analyse und kollaboratives Filtern
- Assoziationsregeln
 - Häufige Elementmengen, Support, Confidence
 - A-Priori-Algorithmus
 - PCY-Algorithmus
- LSH
 - Jaccarddistanz
 - Min-Hashing
 - Rate der False Negatives
- Supervised Machine Learning
 - Entscheidungsbäume, (bedingte) Entropie
 - Maximal Margin Classifier, Support Vector Machines
- Social Network Graphs
 - Girvan-Newman, Edge-Betweenness
 - Spectral Partitioning
- PageRank (Definition, Probleme)
- Datenströme
 - DGIM-Methode
 - Bloom-Filter
 - Flajolet-Martin-Algorithmus
 - Alon-Matias-Szegedy -Algorithmus