

Masterarbeit/Bachelorarbeit

Analyse von Assoziationsregeln zwischen Sub-Spaces von hoch-dimensionalen Daten

Hoch-dimensionale Daten treten häufig auf, und sind daher ein wichtiges Studienobjekt der Informationsvisualisierung: Eine beliebige Exeltabelle stellt bereits einen hoch-dimensionalen Datensatz dar, wobei die Spalten die unterschiedlichen Dimensionen repräsentieren.

Betrachten wir eine Teilmenge dieser Dimensionen, dann hat sich die Bezeichnung Sub-Space diesbezüglich etabliert. Eine interessante Analyseaufgabe ist Assoziationen zwischen (disjunkten) Sub-Spaces zu finden: Assoziationen sind in einfachster Form Wenn-Dann-Relationen zwischen gleichartigen Objekten in unterschiedlichen Sub-Spaces, wie Beispielsweise „Wenn Kunde A Produkt B kauft, dann kauft Kunde A zu 90 % auch Produkt C“.

Diese Abschlussarbeit soll ein visuelles Werkzeug für die Analyse von Assoziationen in hoch-dimensionalen Daten entwickeln, implementieren, und evaluieren. Dabei kommen Methoden des Sub-Space Clustering von Dimensionen zum Einsatz, aber auch Methoden zum Clustern von gleichartigen Objekten ansich. Es ist wünschenswert - nicht aber zwingend erforderlich - wenn Sie Kenntnisse im Bereich Clustering mitbringen. Sie können für ihre Arbeit auf ein vorhandenes Framework (C++) zurückgreifen, und dieses ausbauen.

Sie sind motiviert und haben Interesse an der Visualisierung von Daten? Dann ist dies genau das richtige Thema für sie. Melden sie sich einfach bei mir elektronisch unter dirk@isg.cs.uni-magdeburg. Ich freue mich darauf sie persönlich kennenzulernen.

Aufgaben:

- Entwickeln eines visuellen Werkzeugs zur Assoziationsanalyse von hoch-dimensionalen Daten
- Evaluierung des Werkzeugs anhand realer Daten
- Integration des Werkzeugs in ein vorhandenes Framework (C++)

Kontakt:

Dr.-Ing. Dirk Joachim Lehmann
Office: 29/243 Spezialgerätelabor (234 alternativ)
Mail: dirk@isg.cs.uni-magdeburg.de

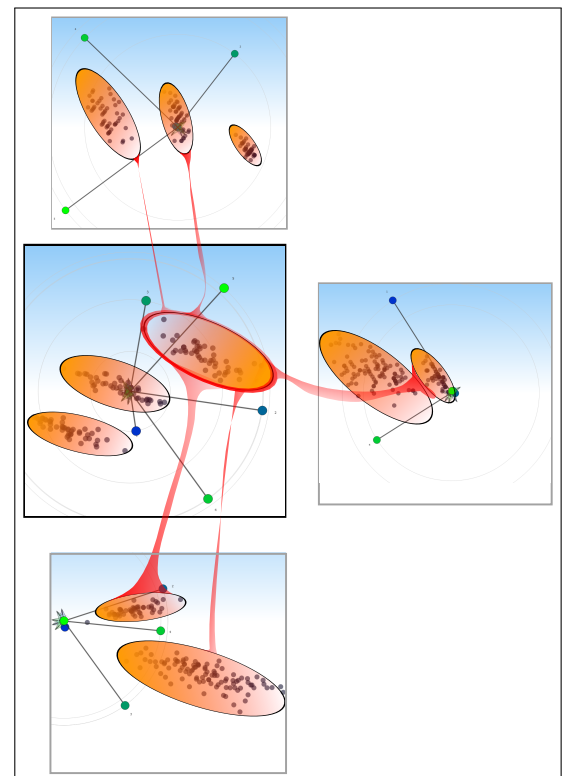


Abbildung 1: Assoziationsanalyse