

Aufgabenserie 8 zur Vorlesung "Computergestützte Datenanalyse"

1. Analysieren Sie die Datensätze "iris", "wein", "banknoten", "Tabak1", "spam", "vokale", "znervensystem", "weizen", "urin", "margarine3" und "gestein" mit Hilfe der Diskriminanzanalyse. Geben Sie das entsprechende Modell (Klassifikationsfunktionen) an. Welcher Anteil an Fällen wird jeweils mit diesem Modell richtig klassifiziert? Lassen Sie sich eine 2D-Grafik der Daten ausgeben, bei der 2 Merkmale ausgewählt werden. Führen Sie die kanonische Diskriminanzanalyse durch. Alternativ kann mit folgenden Datensätzen gearbeitet werden: "aepfel2", "bordeaux", "entwicklungsind", "diabetis", "glas", "grundwasser", "pimaindianer", "rohoel", "sojamaisoel", "weizen", "toepferware", "wuehlmaeuse", "urin".

2. Analysieren Sie die Datensätze "wstruktur", "usuntern", "klausuren" und "ostdlandkreise" mit Hilfe der Faktoranalyse. Welche Faktoren können extrahiert werden?