

Praxisübung zur 8. Übung „Knowledge Discovery“

Wintersemester 2008/2009

Dichtebasiertes Clustering mit OPTICS

Ziel dieser Übung ist das eigenständige Implementieren des Clustering-Algorithmus OPTICS als RapidMiner-Plugin.

Nutzen Sie die Datei `KDDOptics.java` als Grundgerüst für Ihre Implementierung. Behalten Sie Klassen- und Paketnamen bei und implementieren Sie Ihren Algorithmus in der Methode `private List<OpticsData> optics(final List<OpticsData> examples)`. Selbstverständlich können Sie ggf. weitere Methoden in der Klasse anlegen.

Beachten Sie bitte zusätzlich folgende Formalitäten:

- Abgabe der Praxisübung: bis **27.1.2009, 23:59 Uhr MEZ** per E-Mail an jaeschke@cs.uni-kassel.de. Spätere Einreichungen werden nicht berücksichtigt!
- Alle benötigten Methoden, Attribute, etc. müssen sich in der Java-Datei `KDDOptics.java` befinden, deren Grundgerüst Sie hier finden: <http://www.kde.cs.uni-kassel.de/lehre/ws2008-09/kdd/daten/KDDOptics.java>.
- Am 29.1.2009 werden Sie uns Ihren Algorithmus kurz vorführen, erklären und ggf. Fragen beantworten. Termine (pro Teilnehmer ca. 5 Minuten) werden wir vorher vereinbaren.
- Für eine positive Bewertung muss ihr Code problemlos compilieren, einen Durchlauf durch den Datensatz `optics.csv` von der Vorlesungs-Webseite ohne Fehler bestehen, eine nachvollziehbare Clusterung liefern (siehe auch Abbildung 1) und es muss im Gespräch erkennbar sein, dass Sie *Ihre Implementierung* des Algorithmus verstanden haben.

Sollten Sie Fragen haben, wenden Sie sich bitte an Robert Jäschke (jaeschke@cs.uni-kassel.de). Zusätzlich bieten wir Ihnen an, dass sie immer donnerstags vor und nach der Übung kurz Fragen stellen können. Weitere Hinweise und Dateien finden Sie auch auf der Webseite <http://www.kde.cs.uni-kassel.de/lehre/ws2008-09/kdd/uebung/rapidminer.html>.

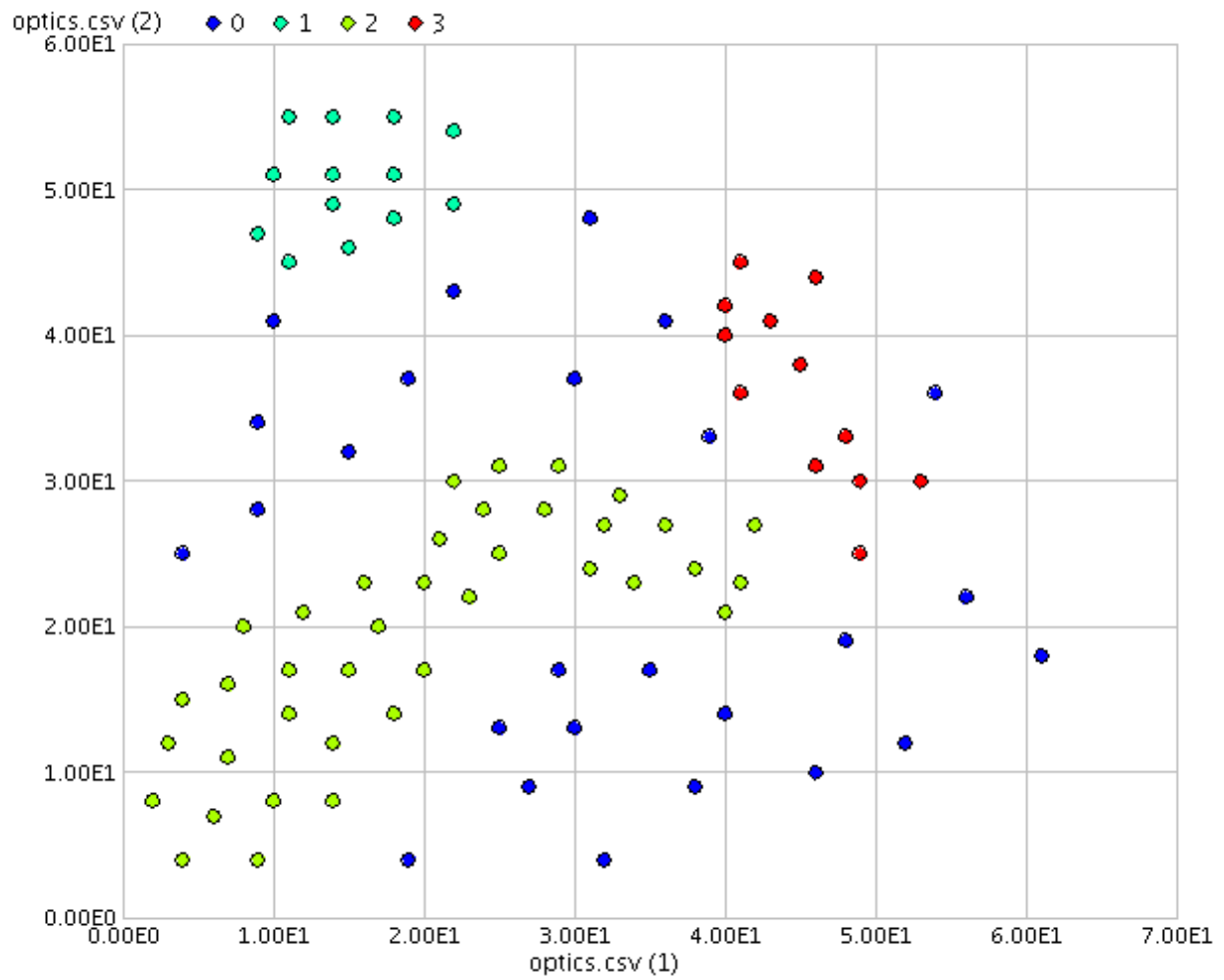


Abbildung 1: Eine OPTICS-Clustering mit den Parametern $minPts = 5$, $\varepsilon = 6.0$, $\varepsilon' = 6.0$.