

24. Mai 2007

Übung zur Vorlesung

Data Mining mit CI-Methoden
(SS 2007)

Blatt 3

Aufgabe 3 (10 Punkte)

Gegeben seien N Paare jeweils bestehend aus einem Attribut-Tupel und einer zugehörigen Klassenbezeichnung. Trainieren Sie ein *radiales Basisfunktionsnetz* mit N_T Datenpaaren, um anschließend die Güte des Klassifikators mit den verbleibenden $N - N_T$ Datenpaaren zu testen.

- a) Verwenden Sie den Datensatz `iris.data` ($N = 150$). Die letzte Spalte enthält den Klassennamen. Wählen Sie $N_T = 100$ Datenpaare zufällig für das Training aus. Dokumentieren Sie die Güte des Klassifikators.
- b) Verwenden Sie den Datensatz `ionosphere.data` ($N = 351$). Die letzte Spalte enthält den Klassennamen. Wählen Sie die ersten 200 Datenpaare in der Datei für das Training aus. Dokumentieren Sie die Güte des Klassifikators.

Die Datensätze sind dem *UCI Machine Learning Repository* entnommen. Weitere Informationen finden sich in den zugehörigen Dateien mit der Endung `*.names`, die Sie über den Link `Info` auf der Webseite der Übung erhalten.

Abgabetermin: Donnerstag, 31. Mai 2007, spätestens zur Übung.

Bitte dokumentieren Sie (für sich), wieviel Zeit Sie für die Aufgaben benötigt haben.