

## Übungen zum Modul WPM6: Algorithmisches Lernen

### SS 2011 Blatt 2

Ausgabe: 22.06.2011, Besprechung: 06.07.2011

#### Aufgabe 2.1 Einarbeitung in libsvm / Python-Bindings:

Arbeiten Sie sich in die Python-Bindings ein. Führen Sie das Demo-Programm aus der Readme-Datei aus.

#### Aufgabe 2.2 Mehr-Klassen SVM:

Obwohl in libsvm bereits Unterstützung für Mehrklassen-SVMs integriert ist, soll in dieser Aufgabe diese Unterstützung manuell implementiert werden. Suchen Sie sich einen geeigneten Datensatz aus. Vorschlag zur Vorgehensweise: Erzeugen Sie aus dem Datensatz  $N$  (=Anzahl Klassen) Datensätze, die jeweils nur 2 Klassen haben. (Mitglied/ Nichtmitglied). Trainieren Sie  $N$  Modelle und evaluieren Sie einen Satz Testdaten. Hinweis: `svm_predict` gibt einen decision value zurück. Anhand dieser Werte kann bei Mehrdeutigkeiten die passendste / am wenigsten unpassende Zuordnung gewählt werden.

#### Aufgabe 2.3 Veranschaulichung der PCA: Für diese Aufgabe kann ein Tool Ihrer Wahl (z.B.

Python, Matlab...) verwendet werden. Gegeben ist die folgende zweidimensionale Datenmenge:

$x_1$	6	10	9	5	6	8	14	7	12	10	10	15	18	8	12	11	14	18	8
$x_2$	4	5	6	7	7	8	8	9	9	10	11	11	11	12	12	13	13	13	14
$x_1$	17	11	13	16	9	19	11	22	13	17	21	15	17	20	22	25	18	24	21
$x_2$	14	15	15	15	16	16	17	17	18	18	18	18	20	20	20	20	21	21	22

- Zeichnen Sie die Punktwolke dieser 38 Datenpunkte.
- Führen Sie auf der gegebenen Datenmenge eine PCA durch.
- Zeichnen Sie die aus der PCA gewonnenen Eigenvektoren und interpretieren Sie das Ergebnis.

#### Aufgabe 2.4 [0 Punkte] Rücktransformation: Gegeben seien die Daten aus Aufgabe 1

- Geben Sie die allgemeingültige Formel zur Rücktransformation der dimensionsreduzierten Daten, die aus einer PCA gewonnen wurden, in den Ausgangsraum an.
- Berechnen Sie zu der gegebenen Datenmenge jeweils den dimensionsreduzierten sowie den rücktransformierten Wert wenn Sie zur Dimensionsreduktion nur den Eigenvektor zu dem größten Eigenwert verwenden.
- Zeichnen Sie die rücktransformierten Werte in ein Koordinatensystem ein.
- Bestimmen Sie für jeden Datenpunkt den Fehler der bei der 1-dimensionalen Rücktransformation durch die Dimensionsreduzierung entsteht.

