

## Übungen zum Modul: Algorithmisches Lernen, Part 2

SS 2009 Blatt 4

Ausgabe: 17.06.2009, Abgabe: 24.06.2009

### Aufgabe 4.1 [0 Punkte] Präsenzübung: Mandelbrotmenge:

### Aufgabe 4.2 [0 Punkte] Implementierung der PCA und Klassifikation:

Implementieren Sie ein Verfahren, das eine PCA auf Bildern durchführt vorzugsweise in MATLAB oder mit Hilfe eines Matrix-Algebra Pakets (z.B. ZLinalg). Berechnen Sie die PCA auf den, auf der Web-Seite der Übung, (diesmal auch wirklich) zur Verfügung gestellten Bildserien und klassifizieren Sie anschliessend Bilder aus der Testmenge.

### Aufgabe 4.3 [0 Punkte] Einarbeitung in LIBSVM:

Laden Sie sich von der Seite [www.csie.ntu.edu.tw/~cjlin/libsvm/](http://www.csie.ntu.edu.tw/~cjlin/libsvm/) (oder über google "libsvm") die libsvmtool herunter. Machen Sie sich mit folgenden Tools vertraut: svm-train, svm-predict, svm-scale, subset.py (in /tools) Informationen hierzu liefert auch die Einführung "A practical guide to SVM classification".

### Aufgabe 4.4 [0 Punkte] Lernen der XOR-Funktion:

Erzeugen Sie sich im Texteditor Ihrer Wahl einen Datensatz zur XOR-Funktion. Diese Textdatei hat folgenden Inhalt:

```
0 1:0.000 2:0.000
0 1:1.000 2:1.000
1 1:1.000 2:0.000
1 1:0.000 2:1.000
```

Die erste Zahl in der Spalte gibt die Klasseneinteilung an, die Parameter sind mit 1,2,... durchnummeriert, die Werte sollten normiert werden  $[-1/+1]$ .

Trainieren sie nun diese Funktion mit den verschiedenen Kernels und testen Sie, indem Sie das erzeugte Modell auf den Ursprungsdatensatz anwenden.

Testen Sie das Modell mit manuell verrauschten Daten.

### Aufgabe 4.5 [0 Punkte] Test mit echten Daten:

Probieren Sie einen oder mehrere Test-Datensätze von der Seite aus. Benutzen Sie das Werkzeug subset.py zur Aufteilung in Trainings- und Messdatensatz. Untersuchen Sie folgende Aspekte:

- Welches Verhältnis Trainingsdaten/Messdaten erscheint bei Ihrem gewählten Datensatz sinnvoll?
- Welche Auswirkung hat vorheriges Skalieren?
- Welche Kernel bringen ein gutes Ergebnis?
- Kann Cross-Validation das Ergebnis verbessern?

