

11. Übungszettel Mustererkennung WS15/16

Prof. Raúl Rojas, Fritz Ulbrich
Institut für Informatik, Freie Universität Berlin
Abgabe Online bis Mittwoch, 03.02.16, 10 Uhr

Bitte laden Sie ihre Lösung der Aufgaben als **pdf-Datei** hoch.
Quellcode können Sie optional als Archiv anhängen.

1. Aufgabe (7 Punkte):

Trainieren Sie mit dem AdaBoost-Verfahren einen starken Klassifikator $C(x_i)$ für den Datensatz in `sparse_checkboard.csv`. Das Format des Datensatzes ist: x-Koordinate, y-Koordinate, Klasse (1 oder -1). (AdaBoost-Tutorial: http://www.inf.fu-berlin.de/inst/ag-ki/rojas_home/documents/tutorials/adaboost4.pdf)

Verwenden Sie dabei insgesamt 20 schwache Klassifikatoren $k_j(x_i)$ der Form:

$$k(x) = 2 * (x.X > n) - 1 \text{ bzw.}$$

$$k(x) = 2 * (x.Y > n) - 1$$

mit $n = (0 \dots 9)$

(Der Klassifikator soll 1 zurückgeben, falls die entsprechende Koordinate von x kleiner als n ist, sonst -1)

- (2 Punkte) Geben Sie für alle k_j die Erkennungsrate auf dem (ungewichteten) Datensatz an.
- (4 Punkte) Geben Sie für alle k_j die gelernten α_j an.
- (2 Punkte) Geben Sie die Erkennungsrate von $C(x_i)$ auf dem Datensatz an.
- (2 Punkte) Erläutern Sie den Nutzen und die Funktionsweise von einer "cascade of classifiers".
(Siehe z.B. <https://www.cs.cmu.edu/~efros/courses/LBMV07/Papers/viola-cvpr-01.pdf>)