Ludwig-Maximilians-Universitaet Muenchen 19.05.2014
Institute for Informatics
Prof. Dr. Volker Tresp
Gregor Jossé

Machine Learning and Data Mining

Summer 2014

Exercise Sheet 4

Presentation of Solutions to the Exercise Sheet on the 22.05.2014

Aufgabe 4-1 Soccer Ball PCA

The Caltech 101 dataset consists of more than 9000 images which have each been assigned to one out of 102 classes. We consider down-scaled images ((32×32)-thumbnails) taken from the classes `soccer ball` and `faces easy`.

a) Conduct a PCA on the 64 `soccer ball` images. Can the images be reconstrcted losslessly with only part of the principal components?

b) Now consider the `faces easy` dataset, consisting of 435 image. Can this dataset be adequately reconstructed using the principal components from part a)?

c) Now consider the dataset of part a) and the principal components of the dataset of part b). Does reconstructing the soccer balls from the faces' components work?

Aufgabe 4-2 Probability Calculus

Let X and Y be random variables with the following data:

		Y		
		1	2	3
X	1	0,1	0,15	0,25
	2	0,05	0,3	0,15

Compute

a) the marginal distributions $P(X = x_i)$ and $P(Y = y_i)$

b) the expectancy values $E(X)$, $E(Y)$

c) the variances $var(X)$, $var(Y)$ as well as the covariance $cov(X,Y)$.

d) the correlation $\rho = \frac{cov(X,Y)}{\sqrt{var(X) \cdot var(Y)}}$

e) if the variables X,Y are independent.

Aufgabe 4-3 Interpretation of the Standard Deviation

Sketch the graph of the standardized normal distribution with the following parameters $f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$; mit $\mu = 0$ und $\sigma = 1$ in the intervall $x \in [-4, 4]$. Mark and intepret the intervalls $0 \pm \sigma;\ 0 \pm 2\sigma;\ 0 \pm 3\sigma$

Aufgabe 4-4 Kernelcombinations

In order to use an own kernel $k(\mathbf{x}_i, \mathbf{x}_j)$ für $\mathbf{x}_i, \mathbf{x}_j \in \mathbb{R}^n$, it must be shown that it is indeed a legitimate kernel. It can be quite complex to show that the *Mercer Theorem* holds for k. Therefore, often the explicit mapping of the implicit basis transformations is stated: $k(\mathbf{x}_i, \mathbf{x}_j) = \phi(\mathbf{x}_i)^T \phi(\mathbf{x}_j)$.

Another popular method of showing the validity of a kernel is representing a kernel, $k(\mathbf{x}_i, \mathbf{x}_j) = k_1(\mathbf{x}_i, \mathbf{x}_j) \circ k_2(\mathbf{x}_i, \mathbf{x}_j)$, as a combination of legitimate kernels combined through valid basis operations.

Show that for valid kernel $k_l(\mathbf{x}_i,\mathbf{x}_j)$, where $l \in \mathbb{N}_+$, holds:

a) **Scaling**: For $a > 0$: $k(\mathbf{x}_i,\mathbf{x}_j) := a \cdot k_1(\mathbf{x}_i,\mathbf{x}_j)$ is a kernel.

b) **Sum**: $k(\mathbf{x}_i,\mathbf{x}_j) := k_1(\mathbf{x}_i,\mathbf{x}_j) + k_2(\mathbf{x}_i,\mathbf{x}_j)$ is a kernel.

c) **Linear combination**: For $w \in \mathbb{R}^d_+$: $k(\mathbf{x}_i,\mathbf{x}_j) := \sum_{l=1}^{d} w_l \cdot k_l(\mathbf{x}_i,\mathbf{x}_j)$is a kernel.

d) **Product**: $k(\mathbf{x}_i,\mathbf{x}_j) := k_1(\mathbf{x}_i,\mathbf{x}_j) \cdot k_2(\mathbf{x}_i,\mathbf{x}_j)$ is a kernel.

e) **Power**: For a $p \in \mathbb{N}_+ : k(\mathbf{x}_i,\mathbf{x}_j) := (k_1(\mathbf{x}_i,\mathbf{x}_j))^p$ is a kernel.