Ludwig-Maximilians-Universitaet Muenchen 20.05.2014
Institute for Informatics
Prof. Dr. Volker Tresp
Gregor Jossé
Johannes Niedermayer

Machine Learning and Data Mining

Summer 2015

Exercise Sheet 5

Presentation of Solutions to the Exercise Sheet on the 27.05.2014

Exercise 5-1 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

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

b) Compute the expected values $E(X)$, $E(Y)$

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

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

e) Compute if the variables X,Y are independent.

Exercise 5-2 Conditional Probability I

Assume that a certain country's population is equally male and female (and that there exist no other sexes). Furthermore, assume that 10% of all men are color blind, but only 1% of all women.

(a) Compute the probability that a person is color blind.

(b) Compute the probability that a color blind person is male.

Exercise 5-3 Conditional Probability II

If screening for a disease, there are several possible outcomes. Let $T+$, $T-$ denote the events that the test is positive and negative, respectively, and D, $\neg D$ denote the events of having and not having the disease, respectively. There are two major criteria to evaluate tests by:

- Sensitivity: Probability (in practice more likely: ratio) of positively tested people having the disease, i.e., $P(T+ \mid D)$.
- Specificity: Probability (or ratio) of negatively tested people not having the disease, i.e., $P(T- \mid \neg D)$.

Now, assume a (realistic) test for HIV with a sensitivity and specificity of 99.9%. Suppose that a person is randomly selected from a population where 1% are infected with HIV and tested with the aforementioned test. Compute the probability that the person has HIV if:

(a) The test is positive.

(b) The test is negative.

Exercise 5-4 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 $0 \pm \sigma;\ 0 \pm 2\sigma;\ 0 \pm 3\sigma$

Exercise 5-5 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.