Exercises for the course
Machine Learning 2
Summer semester 2019

Abteilung Maschinelles Lernen
Institut für Softwaretechnik und theoretische Informatik
Fakultät IV, Technische Universität Berlin
Prof. Dr. Klaus-Robert Müller
Email: klaus-robert.mueller@tu-berlin.de

Exercise Sheet 8

Exercise 1: Multiple Kernel Learning (25 P)

Let $\boldsymbol{x}, \boldsymbol{x}' \in \mathbb{R}^d$ be two data points. Consider a list of positive semi-definite kernels $(k_l(\boldsymbol{x}, \boldsymbol{x}'))_{l=1}^L$. Each of these kernels have an associated feature map $\phi_l : \mathbb{R}^d \to \mathbb{R}^h$, such that

$$k_l(\boldsymbol{x}, \boldsymbol{x}') = \phi_l(\boldsymbol{x})^T \phi_l(\boldsymbol{x}').$$

We now consider an MKL kernel defined as:

$$k(\boldsymbol{x}, \boldsymbol{x}') = \sum_{l=1}^{L} \beta_l \, k_l(\boldsymbol{x}, \boldsymbol{x}'),$$

where $\beta_1, \ldots, \beta_L$ is a list of parameters to be learned.

(a) *Show* that when $\beta_1, \ldots, \beta_L \geq 0$, then the kernel $k(\boldsymbol{x}, \boldsymbol{x}')$ is positive semi-definite.

(b) *Find* a feature map $\phi(\boldsymbol{x})$ associated to this kernel, such that

$$k(\boldsymbol{x}, \boldsymbol{x}') = \phi(\boldsymbol{x})^\top \phi(\boldsymbol{x}')$$

for all $\boldsymbol{x}, \boldsymbol{x}'$.

Exercise 2: Structured Output Kernels (25 P)

Let $\boldsymbol{x}, \boldsymbol{x}' \in \mathbb{R}^d$ be two data points and $y, y' \in \{1, \ldots, C\}$ their respective class, with C the number of classes. Consider the structured output kernel

$$k_{\text{struct}}((\boldsymbol{x}, y), (\boldsymbol{x}', y')) = k(\boldsymbol{x}, \boldsymbol{x}') \cdot 1_{[y=y']},$$

where $k(\boldsymbol{x}, \boldsymbol{x}')$ is a positive semi-definite kernel, with associate feature map $\phi(\boldsymbol{x})$.

(a) *Show* that the kernel $k_{\text{struct}}((\boldsymbol{x}, y), (\boldsymbol{x}', y'))$ is positive semi-definite.

(b) *Find* a feature map $\phi_{\text{struct}}(\boldsymbol{x}, y)$ associated this kernel, such that

$$k_{\text{struct}}((\boldsymbol{x}, y), (\boldsymbol{x}', y')) = \phi_{\text{struct}}(\boldsymbol{x}, y)^T \phi_{\text{struct}}(\boldsymbol{x}', y')$$

for all $(\boldsymbol{x}, y)$ and $(\boldsymbol{x}', y')$.

Exercise 3: Programming (50 P)

Download the programming files on ISIS and follow the instructions.