AG Neuroinformatik
Prof. Dr. Helge Ritter

Universität Bielefeld
Technische Fakultät

Advanced Machine Learning

Problem Sheet 1
due Friday, Oct-30 2015

Problem 1 (Kernel Trick for polynomial SVM) Consider the polynomial Kernel $K(\boldsymbol{u}, \boldsymbol{v}) = (1 + \boldsymbol{u} \cdot \boldsymbol{v})^2$ where $\boldsymbol{u}, \boldsymbol{v} \in [-1, 1]^d$.

a) show that for d=2 that $\boldsymbol{K}(\boldsymbol{u}, \boldsymbol{v})$ can be written as a scalar product $\boldsymbol{\phi}(\boldsymbol{u}) \cdot \boldsymbol{\phi}(\boldsymbol{v})$ in a *six-dimensional state space* and provide explicit expressions for the feature transform $\boldsymbol{\phi}(\boldsymbol{u})$.

b) generalize your result for general (integer) d>1.

c) demonstrate for this case explicitly the validity of Mercer's condition, if the integration range is chosen as the hypercube $[-1, 1]^d$.

d) use your results to explain why the "Kernel Machine"

$$y(\boldsymbol{x}) = \sum_{l=1}^{L} c_l K(\boldsymbol{x}_l, \boldsymbol{x})$$

with $\boldsymbol{x}, \boldsymbol{x}_1, ..\boldsymbol{x}_L \in [-1, 1]^d$ is equivalent to a perceptron

$$y(\boldsymbol{x}) = \sum_{k=1}^{D} w_k \phi_k(\boldsymbol{x})$$

$with$ D linearly independent features $\phi_k(\boldsymbol{x}), k = 1..D$, and specify D

Problem 2 (Linear filter features for time series prediction) Consider the learning a predictor for the simple time sequence

$$x_t = sin(\omega \cdot t).$$

a) show that for a linear predictor that uses only the previous time step value as its only feature, i.e.,

$$x_t = a \cdot x_{t-1} + b$$

no accurate prediction becomes possible (Hint: consider how much b has to minimally vary about a constant value for a given value of a and show from this that no choice of a can make this variation zero).

b) show that using the values x_{t-1} and x_{t-2} of *two* consecutive previous time steps as features allows to find a linear predictor

$$x_t = a_0 + a_1 \cdot x_{t-1} + a_2 \cdot x_{t-2}$$

that predicts the time sequence *perfectly*. What are the values for the coefficients? (Hint: express x_{t-1} and x_{t-2} in terms of $x_t = \sin(\omega \cdot t)$ and $\cos(\omega \cdot t)$ and solve for x_t).