ICS-E4030 Kernel methods in machine learning, course exam 15.12.2014 / Examiner: Juho Rousu

Instructions:

- *You have 3 hours to complete exam.*
- *You are allowed to use one two-sided cheat-sheet (A4 page, both sides hand-written), which you have to submit together with the exam paper.*
- *You may use a scientific calculator.*
- *No additional material is allowed.*

Questions

1. (6 points) Give short (a few sentences) definitions of the following concepts.

   (a) Duality gap
   (b) Line search
   (c) Random walk kernel
   (d) Kernel alignment
   (e) Kendall's distance
   (f) Pre-image problem

2. (6 points) Prove the following proposition:

   Proposition. Given a finite input space $X = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$ and $K(\boldsymbol{x}, \boldsymbol{z})$ a symmetric function on X. Then $K(\boldsymbol{x}, \boldsymbol{z})$ is a kernel function if and only if the matrix

   $$\boldsymbol{K} = (K(\boldsymbol{x}_i, \boldsymbol{x}_j))_{i,j=1}^n,$$

   is positive semi-definite.

3. (6 points) Explain in detail what are string kernels and how they can be used in machine learning

4. (6 points) Derive the dual problem of the structured output optimization problem (soft-margin case):

   $$\min_{\mathbf{w}, \xi_i \geq 0} \quad \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^m \xi_i$$
   $$s.t. \quad \langle \mathbf{w}, \phi(x_i, \mathbf{y}_i)\rangle - \langle \mathbf{w}, \phi(x_i, \mathbf{y})\rangle \geq \ell(\mathbf{y}_i, \mathbf{y}) - \xi_i, \forall i, \mathbf{y} \in \mathcal{Y}$$

   $\varepsilon_i \geq 0$

   $(x^2, \sqrt{2}x, 1)\begin{pmatrix} y^2, \\ \sqrt{2}y, \\ 1 \end{pmatrix}$

   $= \cancel{xy+1}$

   $= x^2y^2 + 2xy + 1$

5. (6 points) Explain in detail how convex optimisation algorithms can be used to train support vector machines

$(xy+1)(xy+1)$
$= x^2y^2+2xy+1$

$sA \quad sK = sA + s^2A^2 + \cdots + s^nA^n + \ldots$

$K = 1 + sA + \cdots + s^{n-1}A^{n-1} + \ldots \qquad (xy, \sqrt{2}xy, 1)$

$\cancel{1 - s^nA^n = (1-sA)k}. \qquad x^2y^2 + 2xy + 1$