CS 6780: Advanced Machine Learning (Submitted on: 3/28/2019)

Submission Assignment #3

Instructor: Thorsten Joachims *Name:* Molly Ingram, Julien Neves, *Netid:* msi34, jmn252

Problem 1

(a)

First, let Σ^d be space representing the collection of all possible strings with alphabet Σ of lenght d. Since $|\Sigma| = a$, we have that $|\Sigma^d| = a^d$. Additionally, we can enumerate the elements of Σ^d, in the following way $\Sigma^d = \{s_1, \dots, s_{a^d}\}$.

We can then imagine a new feature space where for x_i, characterized by the function $\phi_d(x)$ that maps x to the space $\mathbb{N}^{a^d}$ where j^th component of $\phi_d(x)$ is given by the number of occurences of s_j in x.

Then, it is easy to see that $K_d(x_i, x_j) = \phi_d(x_i)\phi_d(x_j)$. Note that the minimum dimensionality of $\phi_d(\cdot)$ is a^d, thus as d increases, the feature space increases exponentially in d.

(b)

(c)

It is straightforward to see that for $K(x_i, x_j) = \prod_{d=1}^{D} K_d(x_i, x_j)$, the implicit feature space of $K(x_i, x_j)$ would simply be the composition of feature spaces of $K_d(x_i, x_j)$ definided in (a), i.e.,

$$\phi(\cdot) = \phi_1(\cdot) \times \cdots \times \phi_D(\cdot)$$

Since, $\phi_d(\cdot)$ maps into $\mathbb{N}^{a^d}$, we have that $\phi(\cdot)$ would map into $\mathbb{N}^{a^1} \times \cdots \times \mathbb{N}^{a^D} = \mathbb{N}^{a+a^2+\cdots+a^D}$. Hence, the dimensionality of the feature space of $\phi(\cdot)$ is given by $\sum_{d=1}^{D} a^d$.

(d)

Let $a, b \in \Sigma$ be distinct elements in Σ. Since $|\Sigma| \geq 2$, a, b exists. Now, we can create the two following strings of length d:

$$u = \{a, a, \dots, a\}$$
$$v = \{b, b, \dots, b\}$$

Then, we can create the following three strings of length $n = 2d$ by joining the string u and v:

$$x_1 = uu$$
$$x_2 = uv$$
$$x_3 = vv$$

Then, we have that $K'(x_i, x_j)$ for x_1, x_2, x_3 is given by

	x_1	x_2	x_3
x_1	1	1	0
x_2	1	1	1
x_3	0	1	1

which is not positive semidefinite matrix.

More precisely, if look at the general K' and take $z^T K' z$ where $z = \{1, -1, 1, 0, 0, \dots\}$, i.e., $z = -1$ at the position of x_2, $z = 1$ at the position of x_1 and x_3, and 0 otherwise, then $z^T K' z = -1 < 0$.

Problem 2

(a)
(b)
(c)

Problem 3

See Jupyter notebook.