COMP 632: Assignment 2

Due on Wednesday, February 18 2015

Presented to Dr. Doina Precup

Geoffrey Stanley
Student ID: 260645907

Question 1

A)

For a function to be considered a kernel function the kernel matrix defined as $K_{ij} = K(x_i, x_j)$ must have two properties:

1. be symmetric
2. be positive semidefinite

As such, a Kernel matrix must abide by the following:

$$K_{ij} = K_{ji} \tag{1}$$

$$z^T K z \geq 0 \tag{2}$$

Where z is an arbitrary vector.

B)

As l increases words will have a tendency of having higher scores when compared with itself then any other words. This will result in a diagonal matrix.

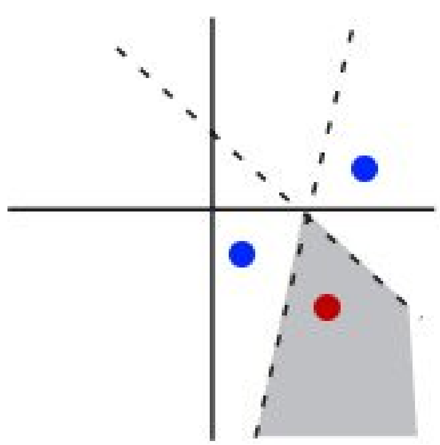
C)

Yes.

D)

Question 2

A)



B)

The VC-dimension of this hypthesis class is 4. This is because it can successfully shatter all configurations of 4 points. However, it would not be able to do so for all configuration of 5 points.

C)

The VC-dimension of any type of boolean combination of 2 linear classifiers is also 4.

Question 3

A)

Given the log-likelihood of a hypothesis h :

$$\log L(h) = \sum_{i=1}^{m} \log P(y_i|x_i, h) \tag{3}$$

And the probability of an example x belonging to class K as being :

$$P(K|x) = 1 - \sum_{i=1}^{K-1} h^i(x) \tag{4}$$

We can derive the log likelihood for a set of hypotheses and a given data set D as:

$$\log L(h) = \sum_{i=1}^{m} \sum_{j=1}^{K} \log \left(1 - \sum_{l=1}^{K-1} h^l(x_i)\right) \tag{5}$$

B)

C)

Question 4

A)

	Folds									
	1		2		3		4		5	
log L Train	-0.340	-1.763	-0.305	-1.772	-0.313	-1.718	-0.340	-1.639	-0.321	-1.712
log L Test	-0.431	-1.632	-0.284	-1.806	-0.319	-1.696	-0.324	-1.561	-0.329	-1.801
Training Accuracy	0.81%		0.81%		0.83%		0.81%		0.80%	
Testing Accuracy	0.69%		0.69%		0.72%		0.87%		0.71%	

B)

C)

	Folds									
	1		2		3		4		5	
log L Train	-0.231	-1.428	-0.174	-1.435	-0.222	-1.560	-0.253	-1.434	-0.223	-1.371
log L Test	-0.410	-1.418	-0.497	-1.627	-0.547	-1.701	-0.549	-1.807	-0.456	-1.664
Training Accuracy	0.82%		0.84%		0.83%		0.81%		0.83%	
Testing Accuracy	0.72%		0.67%		0.69%		0.85%		0.74%	

D)

E)