Machine learning Homework- Soft-Margin SVM and Kernels

Abinav Ravi Venkatakrishnan - 03694216 and Abhijeet Parida - 03679676

December 9, 2018

Problem 1:

No it will not be the correct label. The training sample depends on the distance from the hyperplane decision boundary ξ. If $\xi < 1$ for the training sample it gets classified correctly else it gets mis-classified.

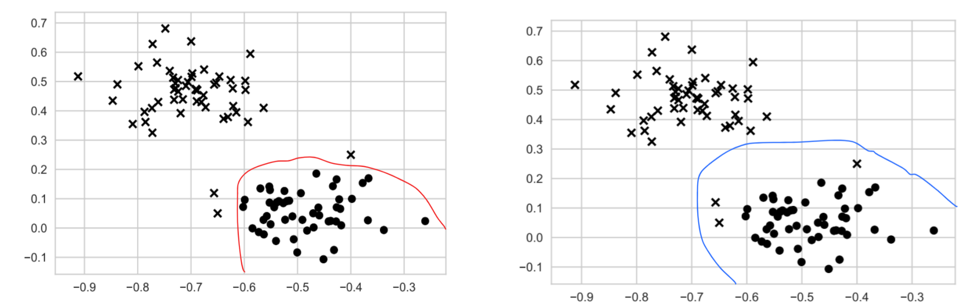
Problem 2:

The cost function for soft-margin SVM is

$$min f_0(\mathbf{w}, b, \xi) = \frac{1}{2}\mathbf{w}^T\mathbf{w} + C\sum_{i=1}^{N}\xi_i \tag{1}$$

C is a penalizing factor on ξ.
case 1: when C = 0 there is no restriction on ξ values.
case 2: when C ¡ 0 it encourages higher values of ξ and hence encouraging mis-classification.

Problem 3:



(a) $C = 10^{10}$ (b) $C = 10^{-10}$

(a) (b)

Figure 1

The first case tends to have hard margin of SVM since the C value is very big and C tends to ∞. The second figure on the right has soft margin since the C value is very less and this can take some mis-classification.

Problem 4:

We know that $\sum_{i=0}^{N} a_i({x_i}^T x_j)^i$ is a polynomial kernel. by the kernel preserving operation we know that sum of valid kernel is also a valid kernel and hence $k(x1, x2)$

Problem 5:

Problem 6:

a) The algorithm is trying to find the matching characters in the given string.
b) The alphabets are mapped from SxS space to real space. Now the SxS space has infinite dimension and gets

mapped to real space. The real number is actually a mapping of s¿0 which means that the transformation is a valid Kernel.

Problem 7:

For linearly separated points the SVM is given by

$$y_i(w^T\phi(x_i) + b) > 0 \tag{2}$$

this equation is transformed into a different forms by using the kernel trick for non linear separations. But this is possible only when the limit of σ tends to zero or very low values of σ