**4) Support Vectors Machines (SVM) [12]**

a) What is the margin for a SVM hyperplane? Why do SVM models maximize the margin? What are support vectors? [4]

Margin means the width of the slab parallel to the hyperplane that has no interior data points [1.5] not mentioning “no interior points” at most 0.5 points

…maximize the margin to better handle noise/to become more fault tolerant [1]

Support vectors are the data points nearest to the hyperplane, the points of a data set that, if removed, would alter the position of the dividing hyperplane. [1.5]

b) There has been a lot work in designing new kernels in machine learning including using kernels in conjunction with support vector machines. What do kernels do? Why do most support vector machine approaches employ non-linear kernels? What do you believe is the reason that support vector machines in conjunction with kernels accomplish quite high accuracies for challenging datasets? [5]

Kernel map the dataset into a different, usually higher dimensional space [1]

To deal with datasets whose classes are not linearly separable [2]

By mapping the data to a higher dimensional datasets there are more ways to separate the examples of the 2 classes, improving the potential to obtain higher accuracies [2]

Other answers for the second or third question might deserve partial credit!

c) Assume we have a dataset with numerical attributes x, y, and an attribute c where c is a class variable which we assume takes value in {0,1}. Give the equation of a hyperplane that the SVM learning could potentially learn for this dataset! [3]

e.g. xy

no partial credit!

**3) SVMs [9]**

a) What are the characteristics of hyperplanes that support vector machines learn from a training set? [3]

b) The soft margin support vector machine solves the following optimization problem:



What does the first term minimize? Depict all non-zero ξi in the figure below! Depict all support vectors in the figure below---if example j is a support vector what is its value for ξj. What is the advantage of the soft margin approach over the linear SVM approach?



All other points

have **ξi** values

of 0!

width

width

## What is a support vector?



[Image: nlp.stanford.edu](https://nlp.stanford.edu/IR-book/html/htmledition/support-vector-machines-the-linearly-separable-case-1.html)

Support vectors are the **data points nearest to the hyperplane**, the points of a data set that, if removed, would alter the position of the dividing hyperplane. Because of this, they can be considered the critical elements of a data set.

c) Explain how examples are classified by SVMs!