Support Vector Machines (SVMs)

**Task 1:**

**Q1: What is the margin and support vectors?**

The margin is the distance/decision boundary between two vectors aligned at the closest positive and negative point. The support vectors are the data points which lie on the borders of the margin and are used to maximize the distance/width of margin.

**Q2: How does SVM deal with non-separable data?**

SVM uses two tricks for non-separable data: soft margin and kernel trick. The soft margin tries to find a line to separate data groups but allows room for some misclassified data points. Kernel trick is the use of existing features and transforming them into new features to new dimensions.

**Q3: What is a kernel?**

A kernel is a function which maps n-dimension input to a higher dimension in a less computationally expensive manner.

**Q4: How does a kernel relate to feature vectors?**

A kernel allows a quick “mapping” or efficient calculation process of the feature vectors transformation which allows a different way to maximize the margin.

**Task 2: Construct a support vector machine that computes the kernel function. Use four values of +1 and -1 for both inputs and outputs:**

**•[−1, −1] (negative)**

**•[−1, +1] (positive)**

**•[+1, −1] (positive)**

**•[+1, +1] (negative)**

**Map the input [x1, x2] into a space consisting of x1 and x1x2. Draw the four input points in this space, and the maximal margin separator. What is the margin?**

Scatter chart

Description automatically generated with medium confidence

•[−1, −1] (negative) = [-1, +1]

•[−1, +1] (positive) = [-1, -1]

•[+1, −1] (positive) = [+1, -1]

•[+1, +1] (negative) = [+1, +1]

The maximal margin separator is where x1x2 = 0, and the margin is of width 1.

**Task 3: Recall that the equation of the circle in the 2-dimensional plane is (x1 − a)2 + (x2 − b)2 − r2 = 0. Please expand out the formula and show that every circular region is linearly separable from the rest of the plane in the feature space (x1, x2, x12, x22).**

Once expanded:

Following the linear SVM model in format (): we can see the weight vector is (1, 1, 2a, 2a) and intercept is which makes it linearly separable in this feature space.

**Task 4: Recall that the equation of an ellipse in the 2-dimensional plane is c(x1 − a)2 + d(x2 − b)2 − 1 = 0.**

**Please show that an SVM using the polynomial kernel of degree 2, K(u, v) = (1 + u · v)2, is equivalent to a linear SVM in the feature space (1, x1, x2, x12, x22, x1x2) and hence that SVMs with this kernel can separate any elliptic region from the rest of the plane.**

Once expanded:

Similar to previous question, we can see the weight vector w=(1, 1, 2ac, 2bd) and intercept is . Thus, linearly separable in this feature space.

**Task 5: Consider the following training data**

**Chart, scatter chart

Description automatically generated**

**(a) Plot these six training points. Are the classes {+, −} linearly separable?**

Yes, it is linearly separable yet small margin.

**(b) Construct the weight vector of the maximum margin hyperplane by inspection and identify the support vectors.**

From the graph above, using the points on the left and right, we can see that the slope of the hyper plane will be -2/2 or -1.

**Task 6: Consider a dataset with 3 points in 1-D:**

**(a) Are the classes {+, −} linearly separable?**

No, the classes are not linearly separable.

**(b) Consider mapping each point to 3-D using new feature vectors φ(x) = [1, sqrt(2)x,x^2] . Are the classes now linearly separable? If so, find a separating hyperplane.**

Transforming the points using φ(x) =[(1, 0, 0), (1, , 1), (1, , 1)]

**Task 7: Learning SVMs on the Titanic dataset. Please report your five-fold cross validation classification accuracies on Titanic training set, with respect to the linear, quadratic, and RBF kernels. Which kernel is the best in your case?**

This case, linear kernel gave a better accuracy than the rest.

Additional Questions:

•Approximately how many hours did you spend on this assignment?

•Which aspects of this assignment did you find most challenging? Were there any significant stumbling blocks?

•Which aspects of this assignment did you like? Is there anything you would have changed?

Please submit a PDF report. In your report, please answer each question with your explanations, plots, results in brief. DO NOT paste your code or snapshot into the PDF.At the end of your PDF, please include a website address (e.g., Github, Dropbox, OneDrive, GoogleDrive)that can allow the TA to read your code.