

Support Vector Machines (SVMs)

Task 1: Q1: What is the margin and support vectors? Q2: How does SVM deal with non-separable data? Q3: What is a kernel? Q4: How does a kernel relate to feature vectors?

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 $[x_1, x_2]$ into a space consisting of x_1 and x_1x_2 . Draw the four input points in this space, and the maximal margin separator. What is the margin?

Task 3: Recall that the equation of the circle in the 2-dimensional plane is $(x_1 - a)^2 + (x_2 - b)^2 - r^2 = 0$. Please expand out the formula and show that every circular region is linearly separable from the rest of the plane in the feature space (x_1, x_2, x_1^2, x_2^2) .

Task 4: Recall that the equation of an ellipse in the 2-dimensional plane is $c(x_1 - a)^2 + d(x_2 - b)^2 - 1 = 0$. Please show that an SVM using the polynomial kernel of degree 2, $K(u, v) = (1 + u \cdot v)^2$, is equivalent to a linear SVM in the feature space $(1, x_1, x_2, x_1^2, x_2^2, x_1x_2)$ and hence that SVMs with this kernel can separate any elliptic region from the rest of the plane.

Task 5: Consider the following training data

class	x_1	x_2
+	1	1
+	2	2
+	2	0
-	0	0
-	1	0
-	0	1

(a) Plot these six training points. Are the classes $\{+, -\}$ linearly separable?

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

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

(class)	x
+	0
-	-1
-	+1

(a) Are the classes $\{+, -\}$ linearly separable?

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

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?

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.