

Generate the datasets A and B in \mathbb{R}^2 , each consisting of 2000 data points from a normal distribution. The dataset A and B have been drawn from the $N(\mu_1, \Sigma_1)$ and $N(\mu_2, \Sigma_2)$. Let us fix the $\mu_1 = [-1, 1]$ and $\mu_2 = [2, 1]$ and $\Sigma_1 = \Sigma_2 = \begin{bmatrix} 0.7 & 0 \\ 0 & 0.3 \end{bmatrix}$.

Separate the 250 data points from each class as a testing set. Plot the optimal Bayesian decision boundary.

1. Write a function implementing the standard SVM with a linear kernel using the gradient descent method. Obtain the best accuracy on the test set by tuning the value of the parameter c . Plot the decision boundary obtained by the standard SVM mode with linear kernel. Compare it with the Bayesian decision boundary.
2. Consider the two-moon dataset. Divide the training and testing point in the ratio of 4:1. Train the standard SVM model with RBF kernel and plot the optimal separating surface obtained by the SVM model by tuning the parameter c and kernel parameter σ . Report Precision, Recall, F-measure, and accuracy on the testing set.
3. Consider the Iris dataset. The dataset contains three types of flowers described by the four features. Consider only the data points with labels 1 and 2. Divide the dataset into training, testing, and validation in the ratio 8:1:1. Use the training set to train the SVM model with a linear kernel. Use the validation set to tune the

parameter value . Finally, obtain the accuracy of the test set.