Support Vector Machines

The goal of a Support Vector Machine (SVM) classifier is to find a linear hyperplane (decision boundary) that separates the data in such a way that the margin is maximized. A linear separation between two classes is accomplished through the function $\boldsymbol{w} \cdot \boldsymbol{x} + \boldsymbol{b} = 0$. We define a function

$$f(x) = \begin{cases} 1 & \text{if } \boldsymbol{w} \cdot \boldsymbol{x} + \boldsymbol{b} \geq 1 \\ -1 & \text{if } \boldsymbol{w} \cdot \boldsymbol{x} + \boldsymbol{b} \leq -1 \end{cases}$$

that can classify items of being of class $+1$ or -1 as long as they are separated by some minimum distance from the class separation function. If the items are not linearly separable we can decide to turn the svm into a soft margin classifier by introducing a slack variable. The most common kernel functions are the family of Radial Basis Function (RBF)

$$L(w) = \frac{\|\boldsymbol{w}\|}{2} + C \sum_{i=1}^{N} \boldsymbol{\varepsilon}$$

$$f(x) = \begin{cases} 1 & \text{if } \boldsymbol{w} \cdot \boldsymbol{x} + \boldsymbol{b} \geq 1 - \boldsymbol{\varepsilon} \\ -1 & \text{if } \boldsymbol{w} \cdot \boldsymbol{x} + \boldsymbol{b} \leq -1 + \boldsymbol{\varepsilon} \end{cases}$$