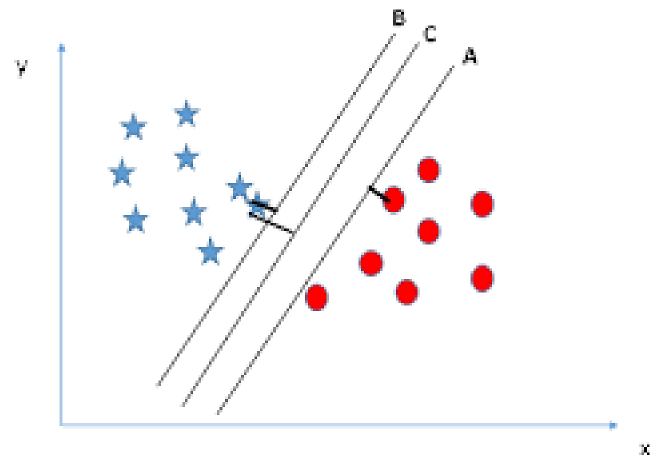
Week 4 Class Notes

Group2: Yaxin Li, Blake Hillier, Joe Puhalla

January 2020

1 Support Vector Machine (SVM)

Suppose we want to classify objects that are not easily separable. We could add dimensions to separate them, and then create a hyper plane between them maximizing the distance between the closest objects. In other words we want to find a point B on a hyperplane π s.t. it is written involving γ_i, the maximum distance between B and the closest object from each class.



Recall:
If $\vec{w}$ is a normal vector of plane $\vec{Px}$, then

$$\begin{aligned} \vec{Px} \perp \vec{w} = \langle \vec{Px}, \vec{w} \rangle = 0 \\ \implies (x-p) \cdot = 0 \\ \implies x\vec{w} - P\vec{w} = 0 \\ \implies wx + b = 0 \end{aligned}$$

Let's define $\gamma_i = y_i(w^T x + b)$. If $y_i = 1$, γ_i is large and positive, but if $y_i = -1$, γ_i is large and negative. Now, we want $y_i(w^T x + b) \geq \gamma_i$ where $\|w\| = 1$, which makes this optimization problem difficult. To solve this, we adjust the equation as follows:

$$\begin{aligned} y_i(\frac{\|w\| w^T x}{\|w\|} + b) \geq & \hat{\gamma}_i \\ y_i(\frac{w^T x}{\|w\|} + \frac{b}{\|w\|}) \geq & \frac{\hat{\gamma}_i}{w} \\ \implies y_i(\hat{w}^T x + \hat{b}) \geq & \hat{\gamma}_i \\ \implies \max & \hat{\gamma}_i \end{aligned}$$

where $\hat{w} = \frac{w}{\|w\|}$, $\hat{b} = \frac{b}{\|w\|}$ and $\hat{\gamma}_i = \frac{\gamma_i}{\|w\|}$. Now without loss of generality, $\gamma = 1$ since $\gamma = \min_i\{\gamma_i\}$. This means

$$\max \hat{\sigma} \implies \max \frac{1}{\|w\|} \implies \min \|w\| \implies \min \frac{\|w\|^2}{2}$$