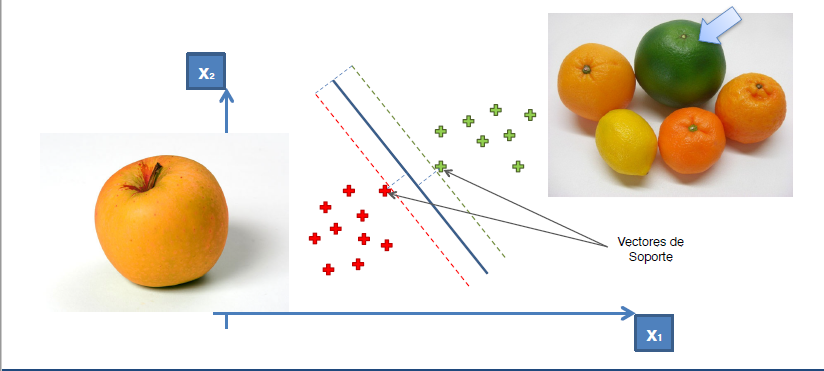
Universidad Galileo

Maestría en Ciencia de Datos

SVM



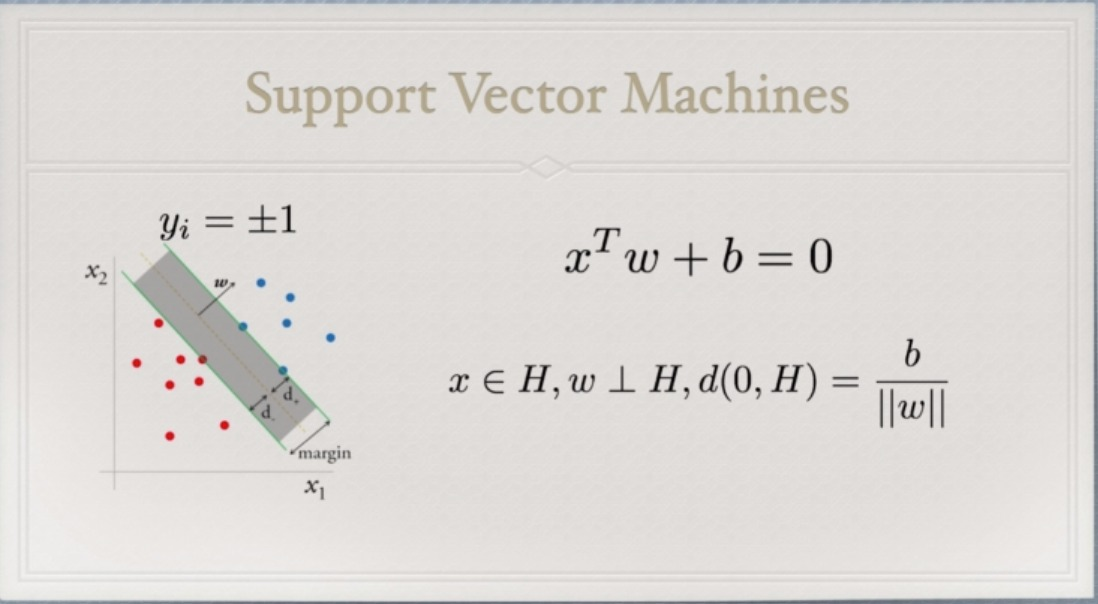
Jaime Ariel Chitay

19002065

Luis Leal

# Hipótesis SVM

Para el caso linealmente separable, SVMs determina el hiperplano óptimo que separa el conjunto de datos. Para este propósito, linealmente separable requiere encontrar el par (w, b) tal que clasique correctamente los vectores.



# Función de Costo

El vector de soporte es una muestra que está clasificada incorrectamente o una muestra cercana a un límite. Mirando la trama a continuación. Las muestras con círculos rojos son exactamente el límite de decisión. En SVM, solo los vectores de soporte tienen un impacto efectivo en el entrenamiento del modelo, es decir, eliminar el vector sin soporte no tiene ningún efecto en el modelo. ¿Por qué? Lo resolveremos a partir de su función de costo.

La función de pérdida de SVM es muy similar a la de Regresión logística. Mirándolo por y = 1 e y = 0 por separado en la siguiente gráfica, la línea negra es la función de costo de la Regresión logística, y la línea roja es para SVM. Tenga en cuenta que el eje X aquí es la salida del modelo en bruto, θᵀx. Recuerde que poner la salida del modelo en bruto en la función Sigmoide nos da la hipótesis de la regresión logística. ¿Cuál es la hipótesis para SVM? Es simple y directo. Cuando θᵀx ≥ 0, predice 1, de lo contrario, predice 0.