Ayudantía 11

SVM - Support Vector Machines

Por Blanca Romero y Gonzalo Fuentes

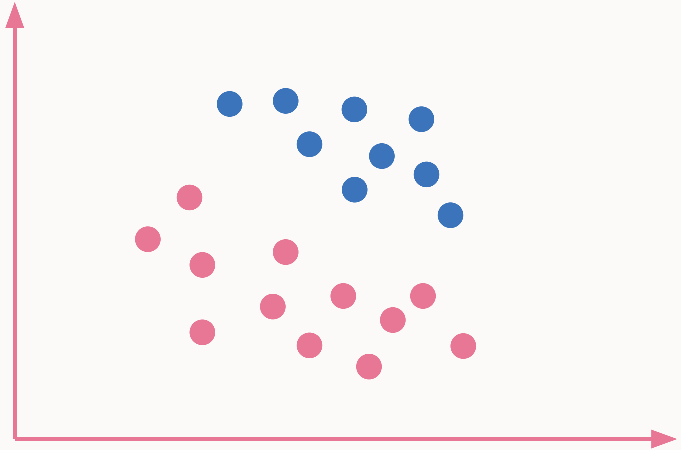
8 de noviembre 2024

¿Qué es una SVM?

¿Qué es una SVM?

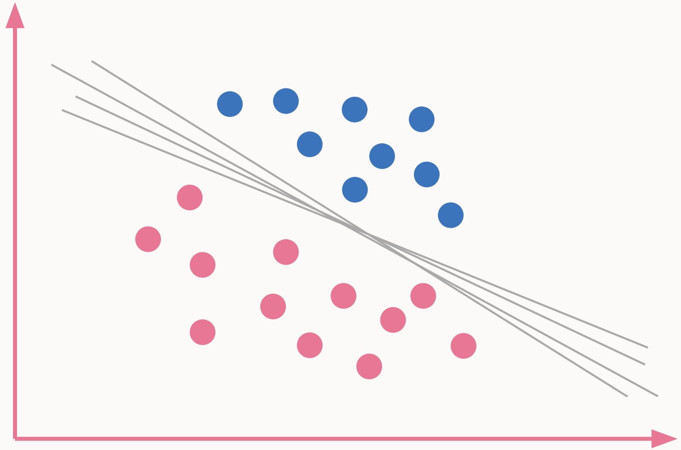
Queremos encontrar una **recta** que **separe los datos**

Para luego poder **clasificar** nuevos datos



¿Qué es una SVM?

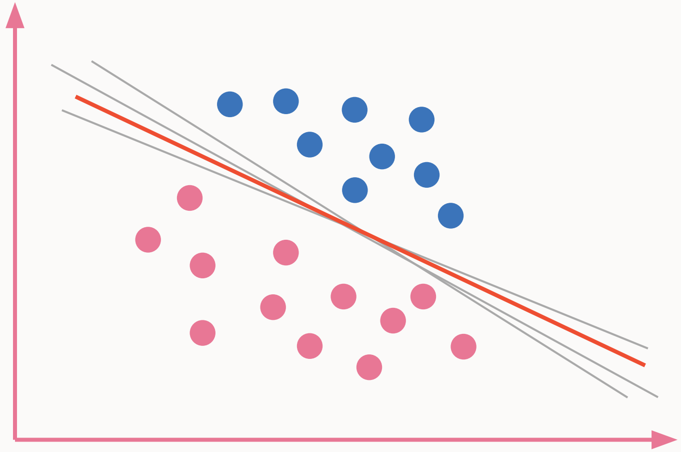
Existen **infinitas** rectas que logran separar los datos:



¿Qué es una SVM?

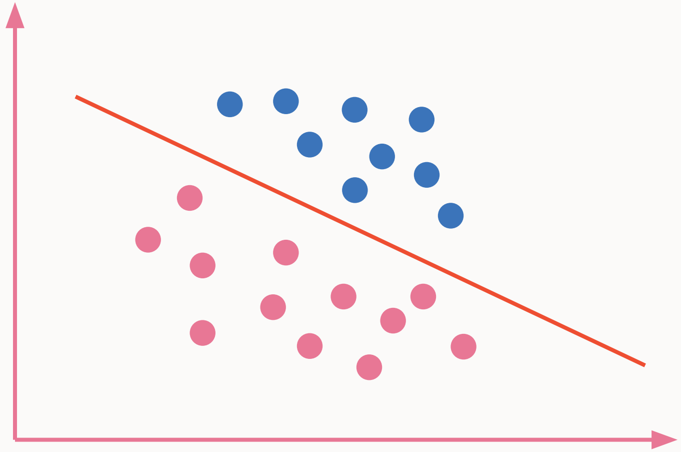
Existen **infinitas** rectas que logran separar los datos:

¿Cómo elegimos la mejor?



¿Qué es una SVM?

SVM nos permite encontrar la mejor recta:

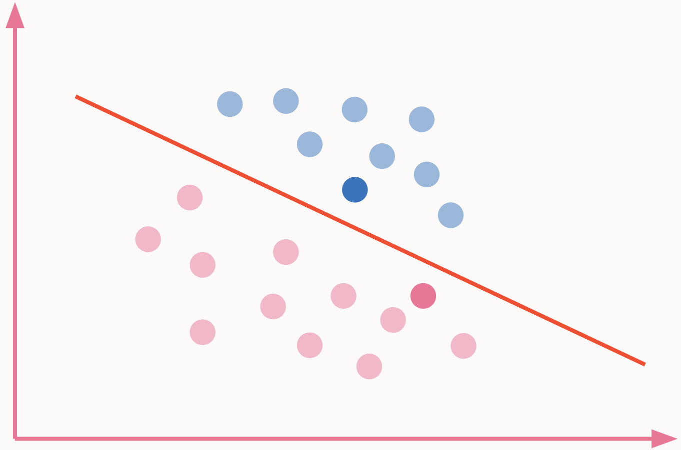


¿Qué es una SVM?

SVM nos permite encontrar la mejor recta:

Busca **maximizar** la distancia de los datos **más cercanos a la recta** (de cada clase)

● y ● son los *vectores de soporte*

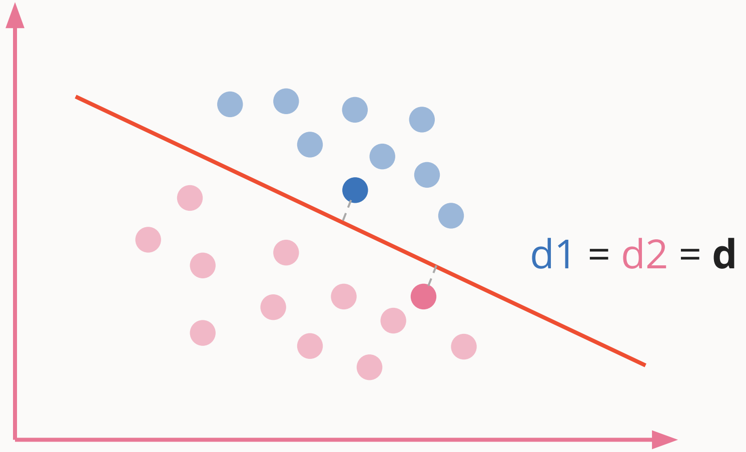


¿Qué es una SVM?

SVM nos permite encontrar la mejor recta:

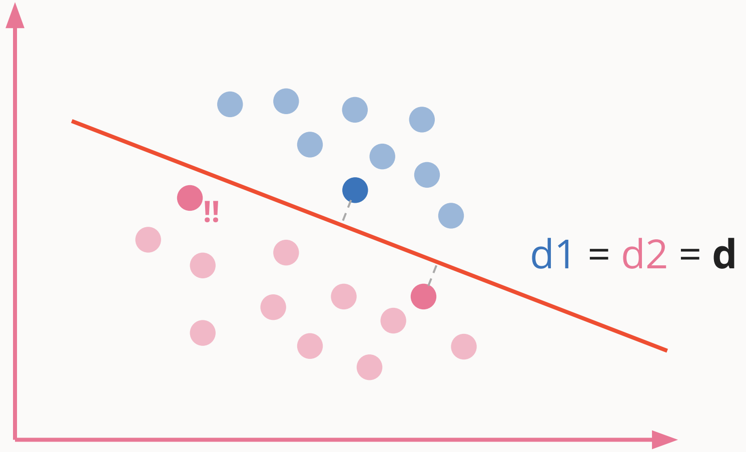
Busca iterativamente para **maximizar d**

Y trabaja siempre con rectas **equidistantes** a los *vectores de soporte*

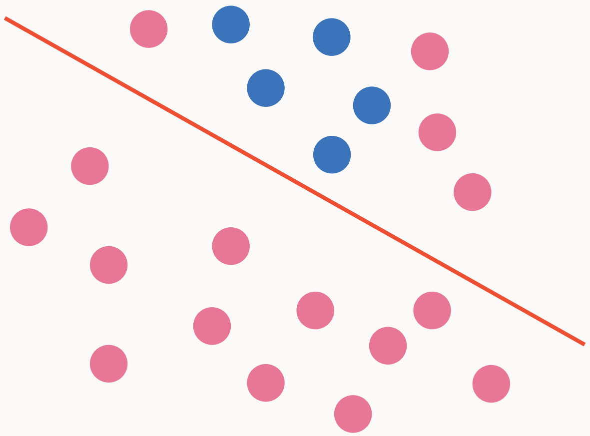


¿Qué es una SVM?

Debemos procurar que cada **nueva recta** considere los **nuevos vectores de soporte** que se podrían generar

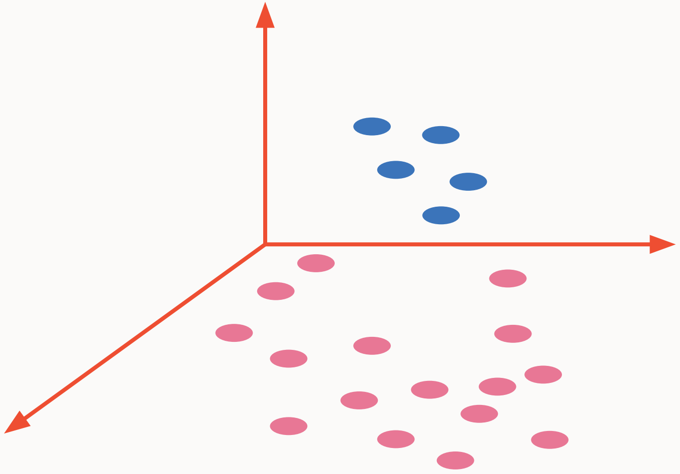


¿Y si no puedo hacer una clasificación lineal?



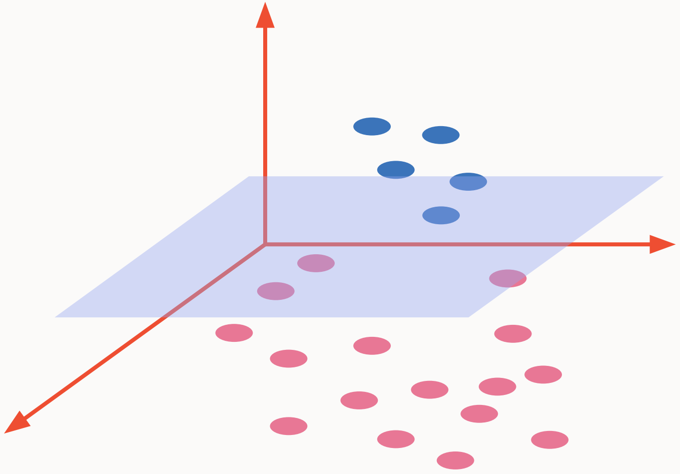
¿**Qué podemos hacer** al respecto?

¡Transformamos los datos!



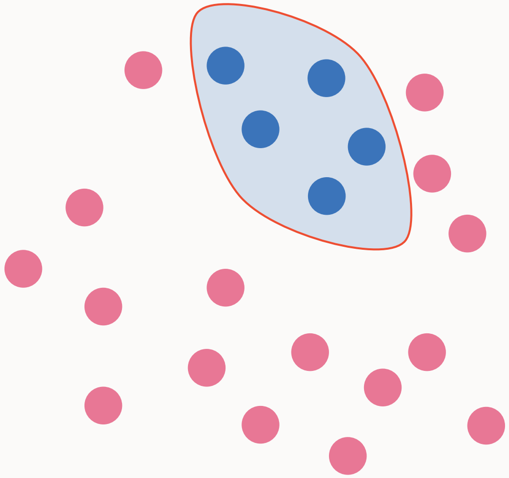
Transformamos los datos a un **nuevo** **espacio geométrico**

¡Transformamos los datos!



Así podemos tomar **un plano** para separar los datos

¡Transformamos los datos!



La **proyección** del plano es un **segmento** del espacio inicial

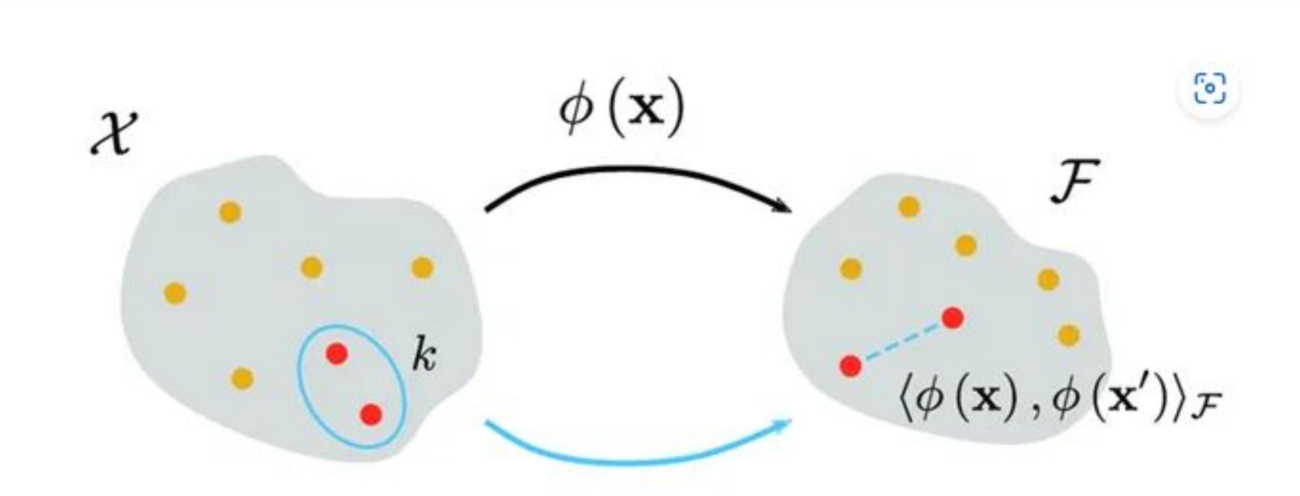
¿Cómo puedo transformar los datos a otro espacio?

La idea nace de ver SVM dual

$$\max_{\alpha} \sum_{i=1}^{M} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{M} y^{(i)} y^{(j)} \alpha_i \alpha_j x^{(i)} . x^{(j)}$$

$$\text{subject to } 0 \leq \alpha_i \leq C \text{ for all } i, \text{ and } \sum_{i=1}^{M} \alpha_i y^{(i)} = 0.$$

¿Cómo puedo transformar los datos a otro espacio?



¿Cómo puedo transformar los datos a otro espacio?

$$K(x, x') = \langle \phi(x), \phi(x') \rangle$$

Lineal

$$K(x, x') = x^T x'$$

Polinomial

$$K(x, x') = (x^T x' + c)^d$$

RBF

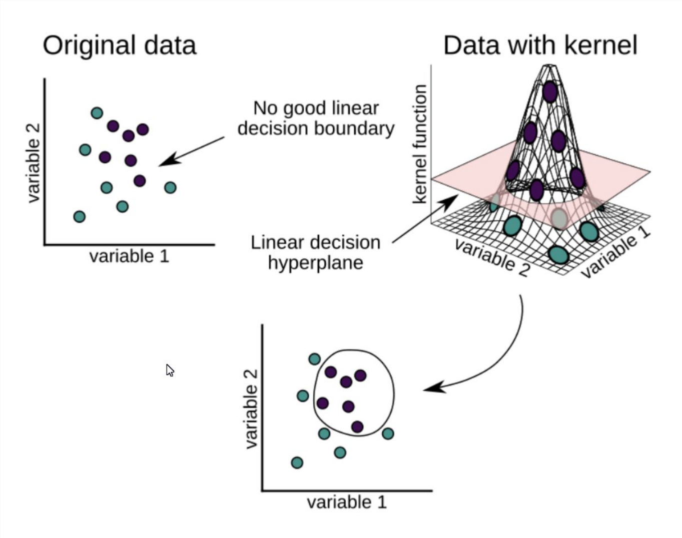
$$K(x, x') = \exp\left(-\gamma \|x - x'\|^2\right)$$

¿Cómo puedo transformar los datos a otro espacio?

$$\max_{\alpha} \sum_{i=1}^{M} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{M} y^{(i)} y^{(j)} \alpha_i \alpha_j K(x^{(i)}, x^{(j)})$$

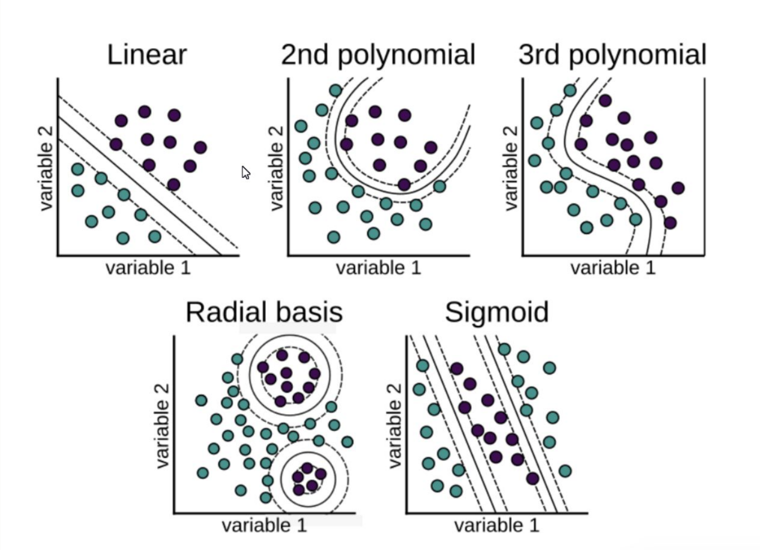
$$\text{subject to } 0 \leq \alpha_i \leq C \text{ for all } i, \text{ and } \sum_{i=1}^{M} \alpha_i y^{(i)} = 0.$$

Función de Kernel



Conjunto de funciones que permite **transformar** el espacio de características con el que trabajamos

Función de Kernel



Existen varios tipos de Kernel para **distintas aplicaciones**