Bootcamp

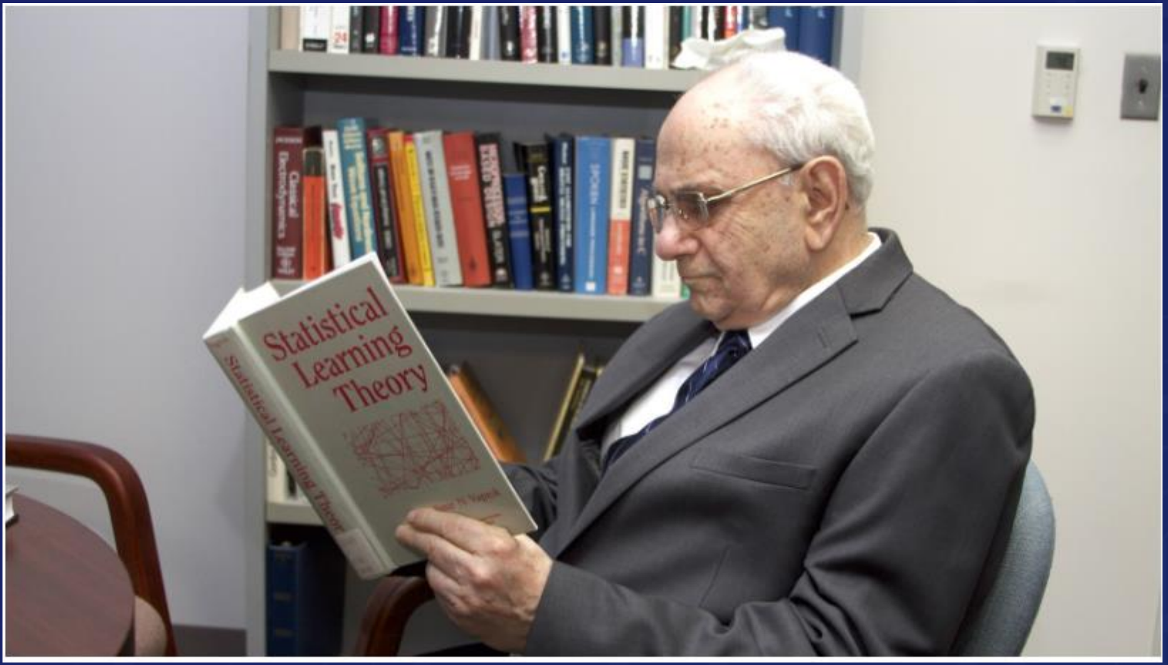
Inteligencia Artificial

Nivel (Intermedio)
Docente: Víctor Viera Balanta
Fecha 23/07/2024

UT TALENTOTECH

Tabla de contenidos

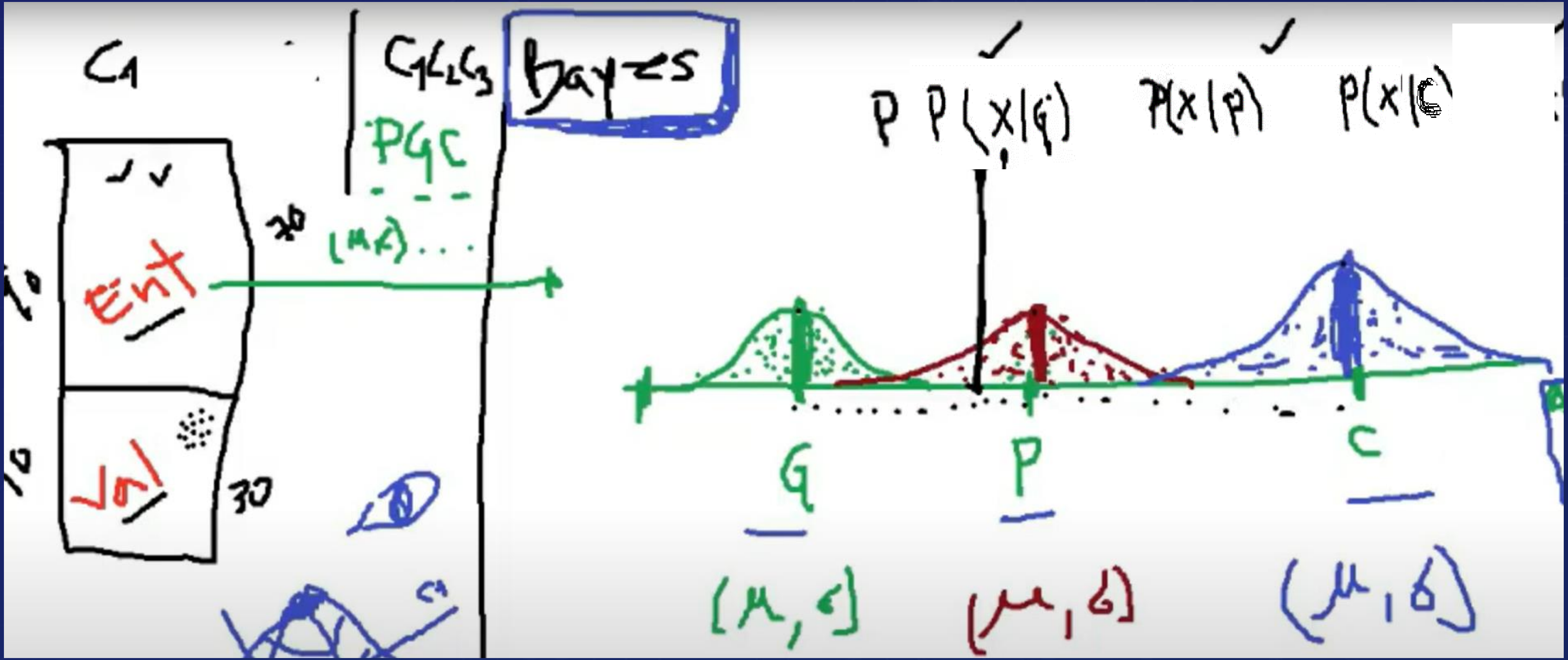
Profesor: Víctor Viera Balanta



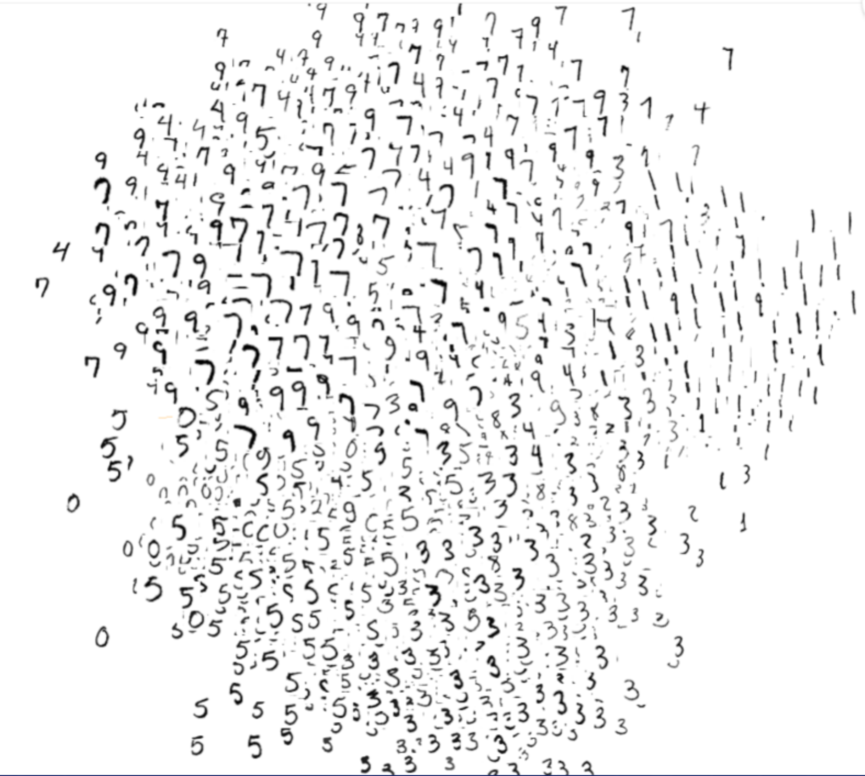
1990 desarrolla los SVM

Vladimir Vapnik

Profesor: Víctor Viera Balanta


C_1
$C_1C_2C_3$
bayes
PGC
Ent
Val
30
$P(x|G)$
$P(x|P)$
$P(x|C)$
G
P
C
(μ,σ)
(μ,σ)
(μ,σ)

Minist *Modified National Institute of Standards and Technology*



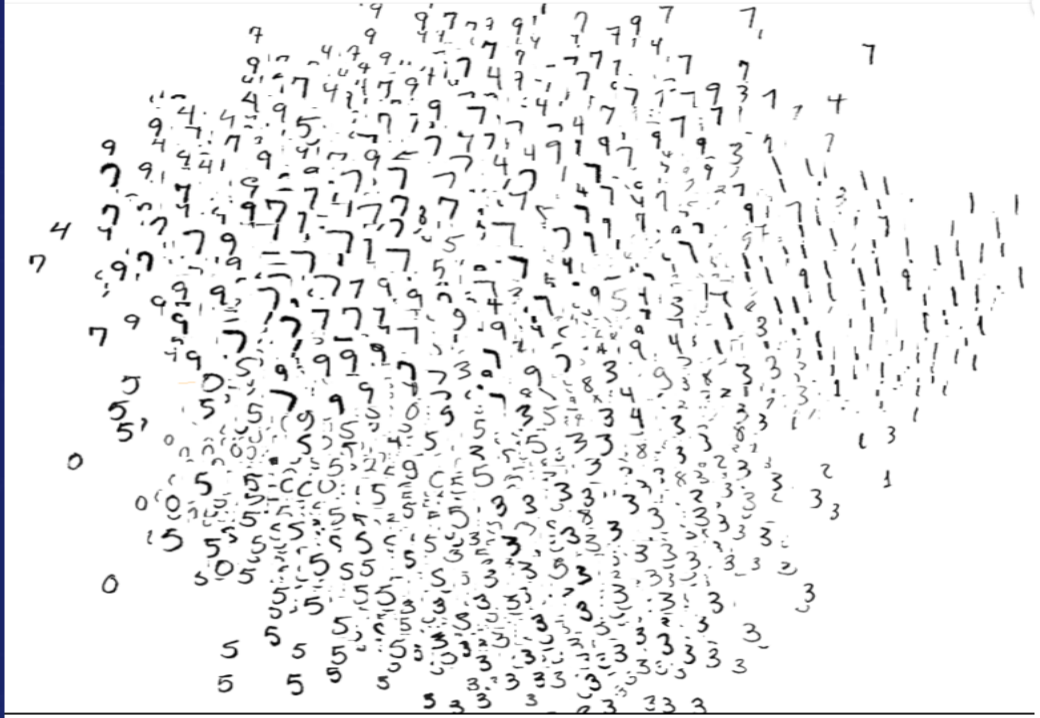
Creado en 1994, con imágenes de dígitos entre 0 y 9, escritos a mano.

60,000 imágenes entrenamiento
10,000 imágenes de prueba.

https://projector.tensorflow.org/

Profesor: Víctor Viera Balanta

Minist *Modified National Institute of Standards and Technology*



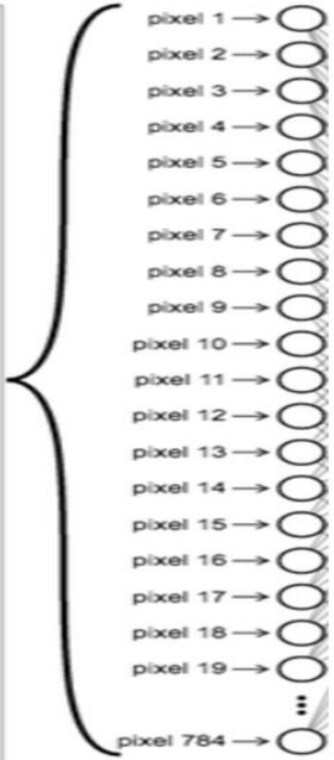
https://projector.tensorflow.org/

Creado en 1994, con imágenes de dígitos entre 0 y 9, escritos a mano.

60,000 imágenes entrenamiento
10,000 imágenes de prueba.

Minist Modified National Institute of Standards and Technology

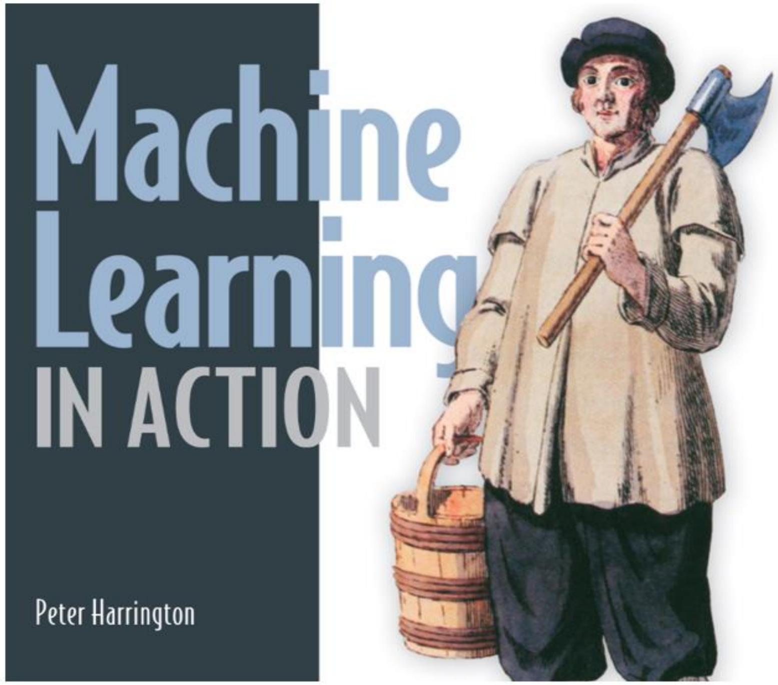
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	4	62	146	182	254	254	181	176	139	15	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	34	186	253	217	208	136	136	136	166	232	99	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	61	242	208	111	3	0	0	0	0	0	18	32	107	43	0	0	0	0	0	0
0	0	0	0	0	0	0	0	156	242	23	0	0	0	0	0	0	0	13	191	181	6	0	0	0	0	0	0
0	0	0	0	0	0	0	0	121	255	98	3	0	0	0	0	0	8	194	225	12	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	169	253	120	3	0	0	0	0	128	247	51	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	3	111	244	169	19	0	14	131	249	117	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	59	241	235	72	142	229	66	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	25	218	254	231	36	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	133	253	221	33	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	19	237	111	196	217	19	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	174	138	0	23	193	204	18	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	96	224	0	0	0	25	218	169	3	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	215	138	0	0	0	0	86	253	99	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	215	97	0	0	0	0	3	162	214	11	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	215	97	0	0	0	0	0	118	253	68	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	185	157	0	0	0	0	0	40	254	98	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	50	244	61	0	0	0	0	112	244	58	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	174	251	142	59	83	167	244	111	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	6	133	253	253	253	169	61	3	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0



https://ml4a.github.io/ml4a/es/neural_networks/

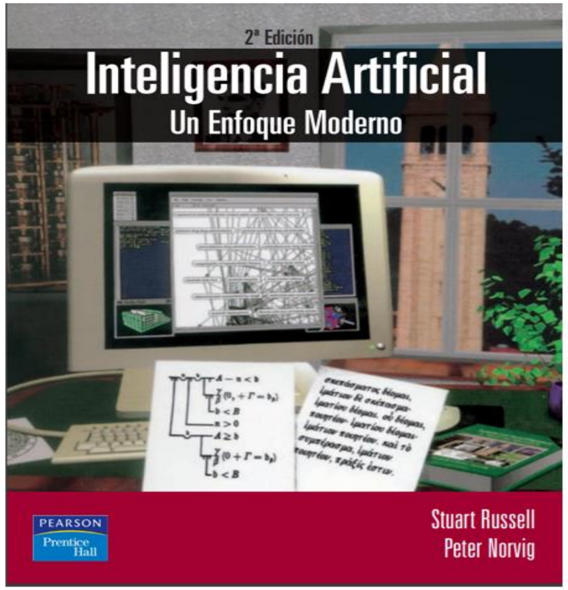
Víctor Viera Balanta

Máquinas de Soporte Vectorial (SVM)



{cb} Codificando Bits

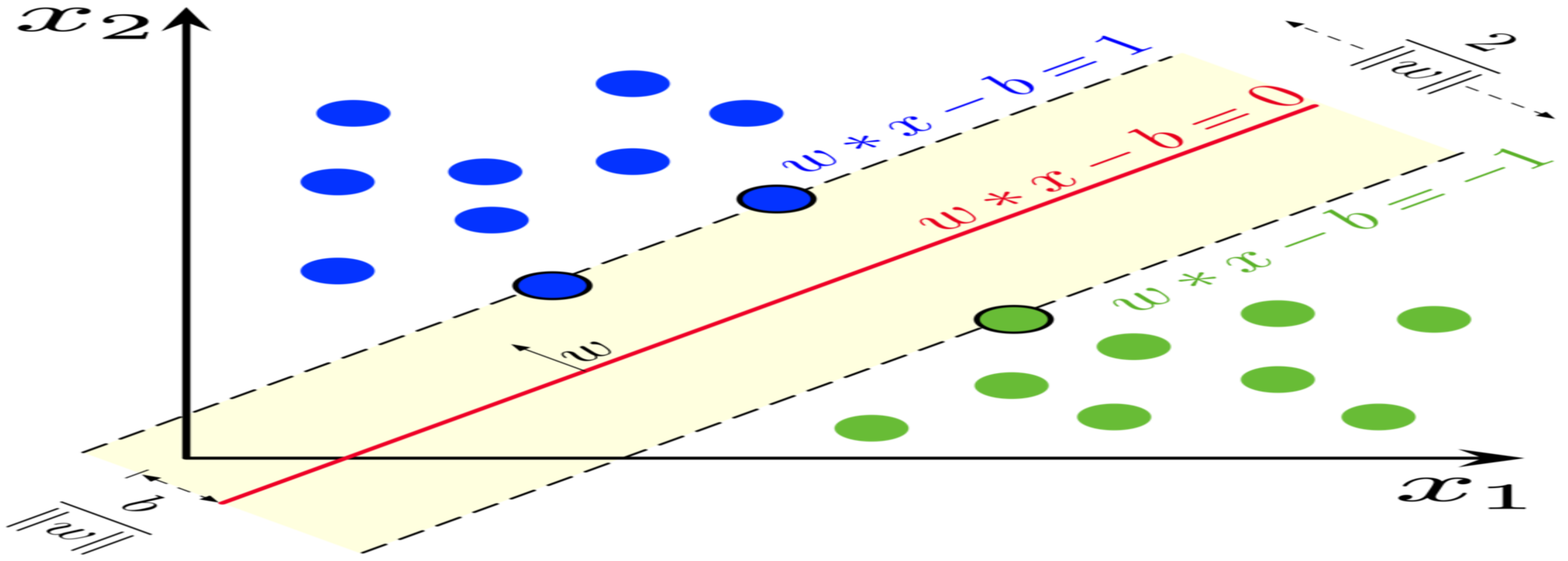
https://youtu.be/Xbd8T-JoGPQ?si=1_aiW50BcfvvPBjp



Minist *Modified National Institute of Standards and Technology*

```
train_images =datasets.load_digits()
train_images.data
train_images.target
```

Máquinas de Soporte Vectorial (SVM)



Linealmente separable. Aprendizaje Supervisado

Máquinas de Soporte Vectorial (SVM)

Algoritmos clásicos de Machine Learning.

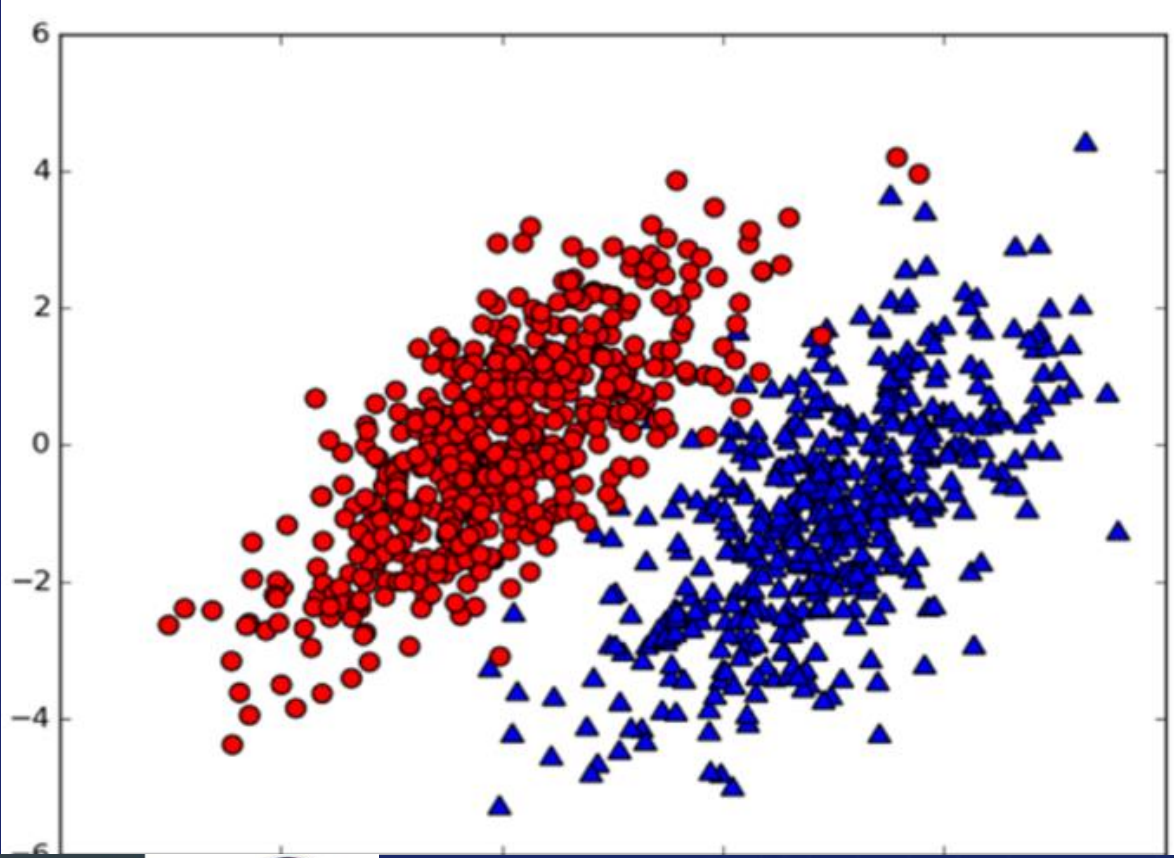
Muy usados para clasificación Binaria.

Kernelized Support Vector Machines

Tienen mecanismos de Kernel.

También se pueden usar para clasificación multiclase.

Problemas de optimización.

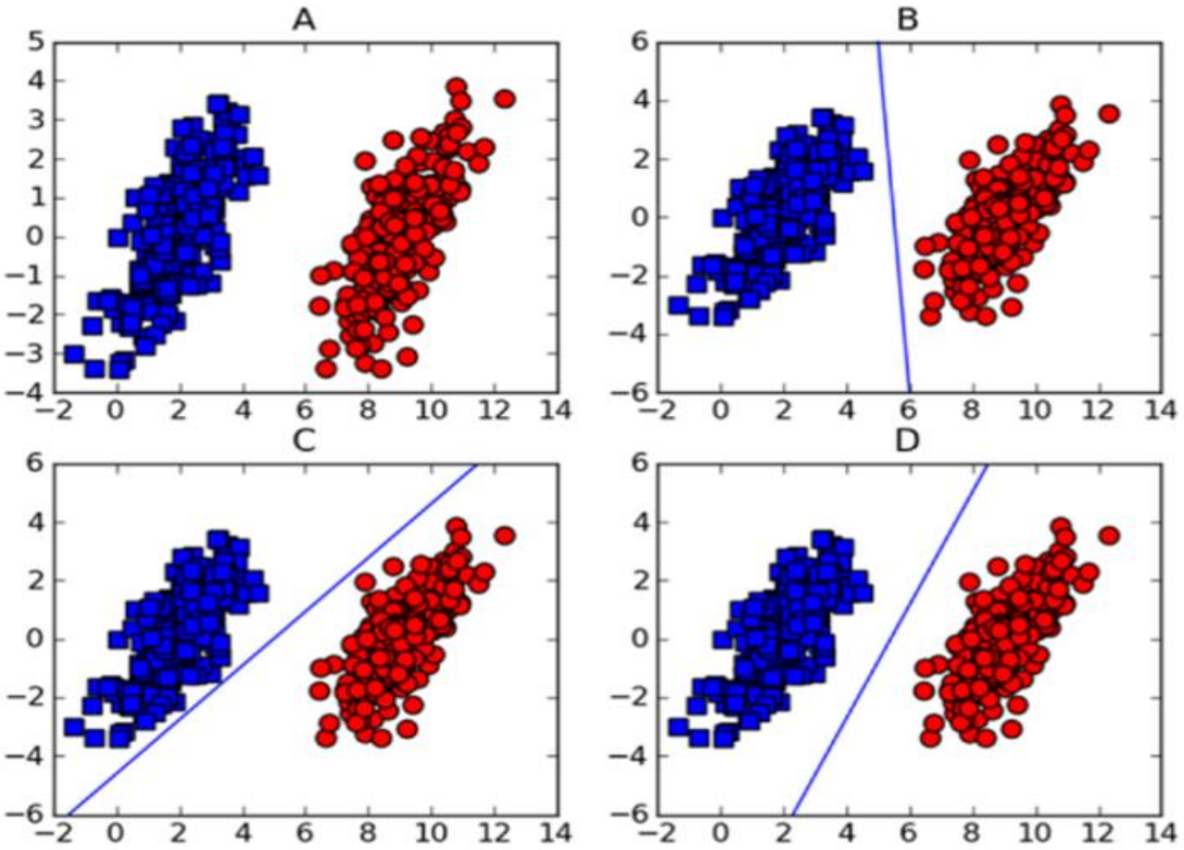


Separating data with the maximum margin

$$\{x_i, d_i\}_{i=1}^{N}$$

Variables independientes con dependientes

Profesor: Víctor Viera Balanta

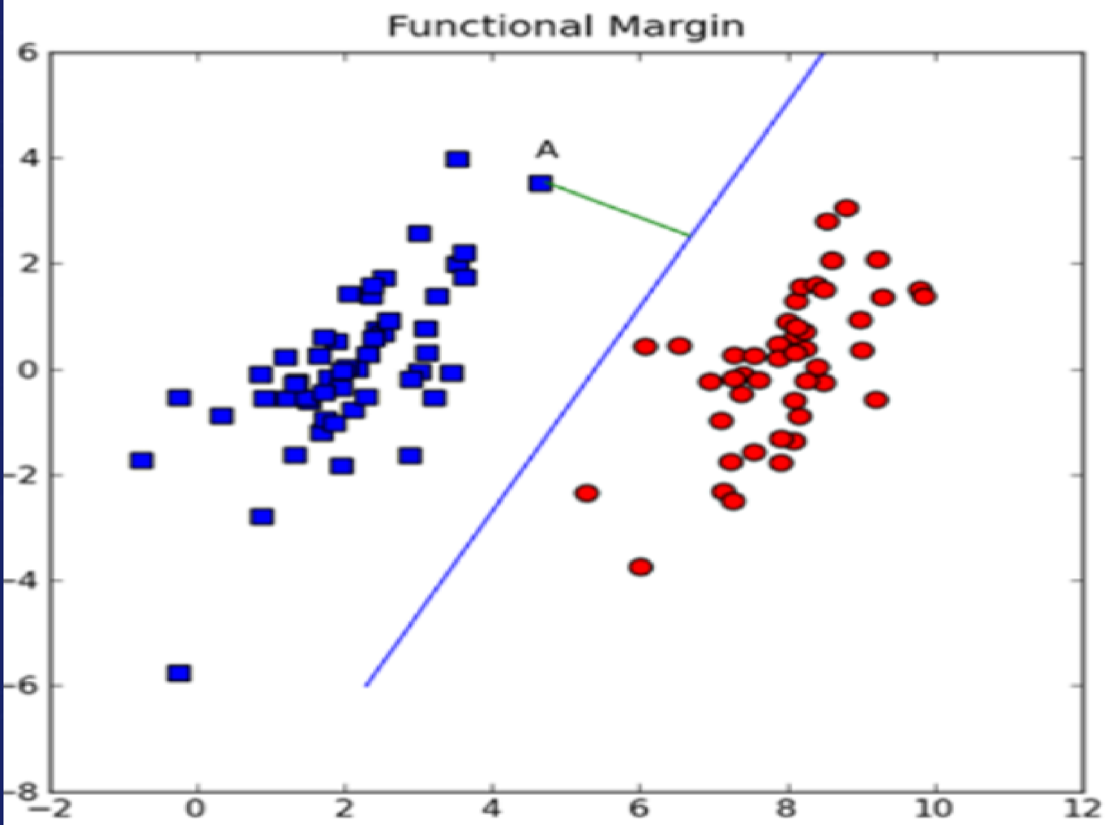


Separating data with the maximum margin

$$\{x_i, d_i\}_{i=1}^{N}$$

Variables independientes con dependientes

Profesor: Víctor Viera Balanta

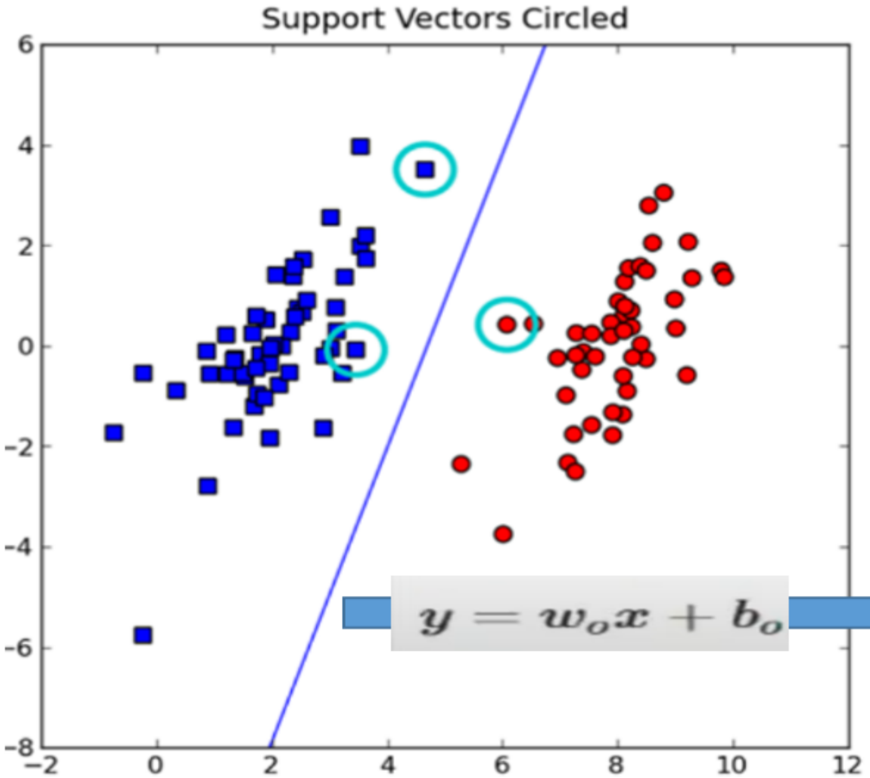


Separating data with the maximum margin

La distancia desde el punto A hasta el plano de separación se mide por una línea normal, al plano de separación.

Profesor: Víctor Viera Balanta

Máquinas de Soporte Vectorial (SVM)



$$y = w_o x + b_o$$

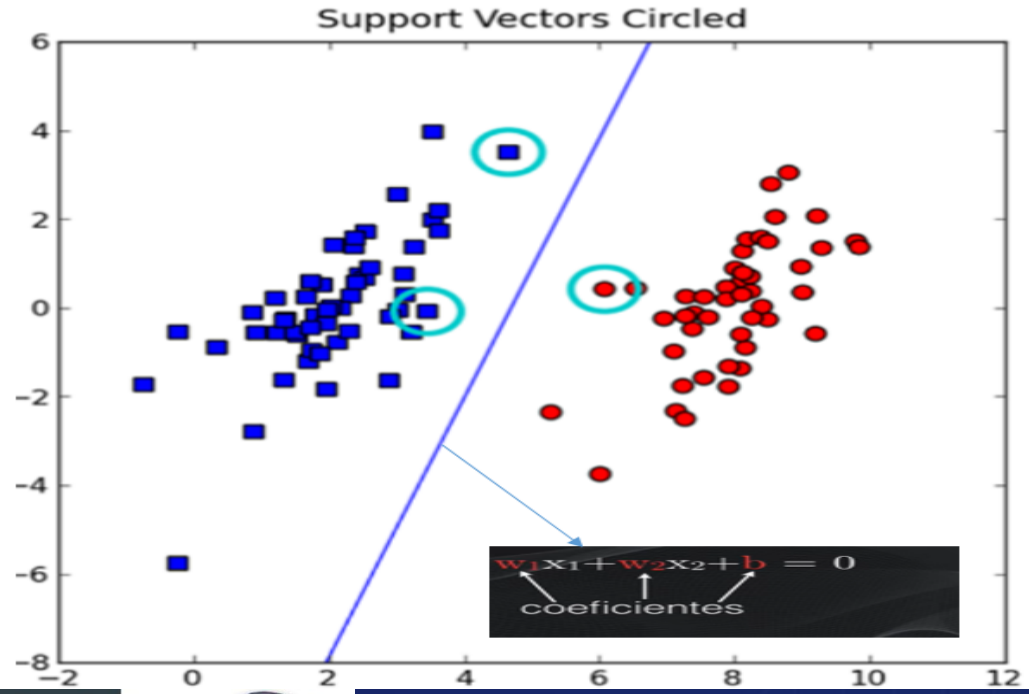
Separating data with the maximum margin

Conjunto de datos de muestra de mostrando los vectores de soporte rodeados por un círculo y el hiperplano de separación

El hiperplano es
nuestro límite de decisión. Todo lo que está en un lado pertenece a una clase y todo
del otro lado pertenece a una clase diferente.

Profesor: Víctor Viera Balanta

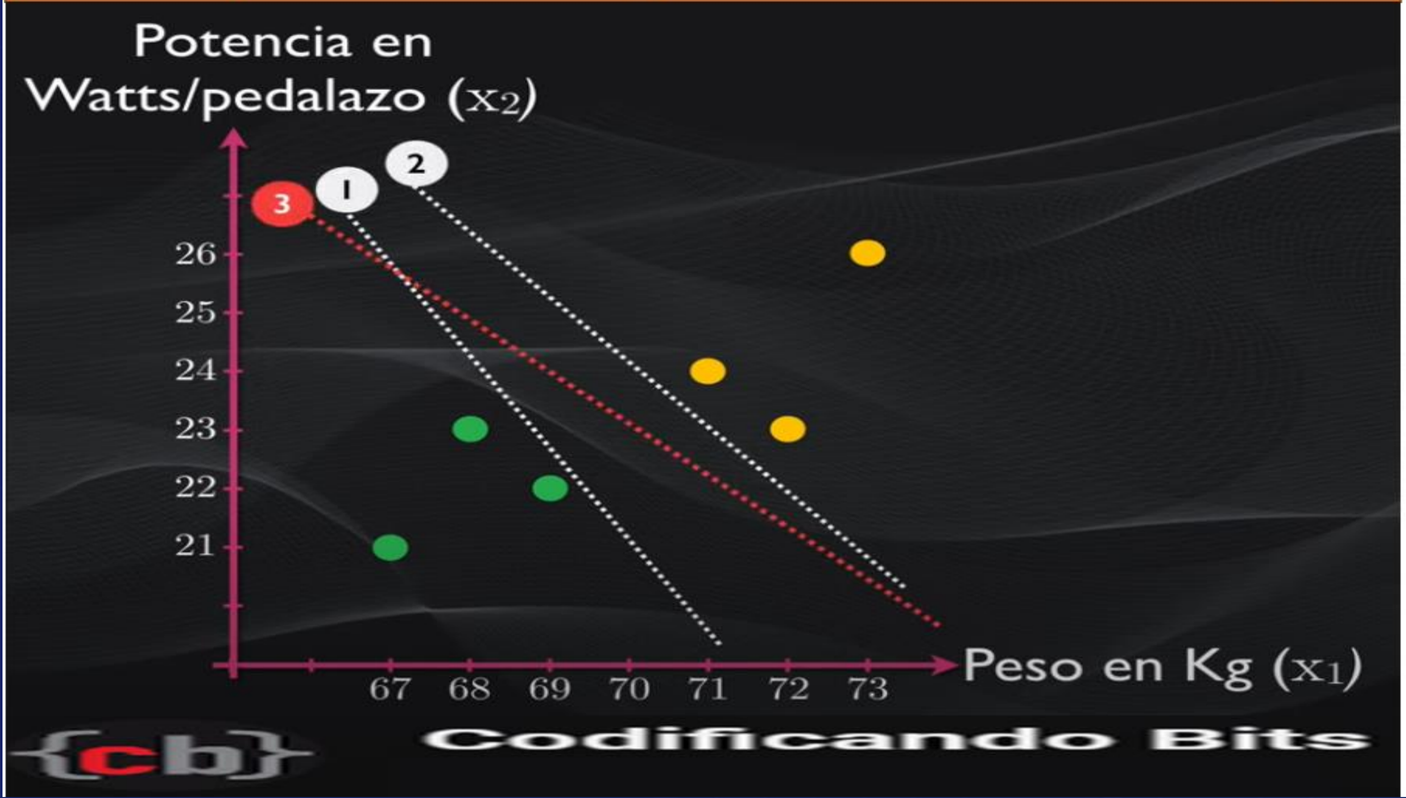
Máquinas de Soporte Vectorial (SVM)



Los puntos más cercanos al hiperplano de separación se conocen como vectores de soporte.

Se trata de maximizar la distancia desde la línea de separación hasta los vectores de soporte, necesitamos encontrar una manera de optimizar este problema.

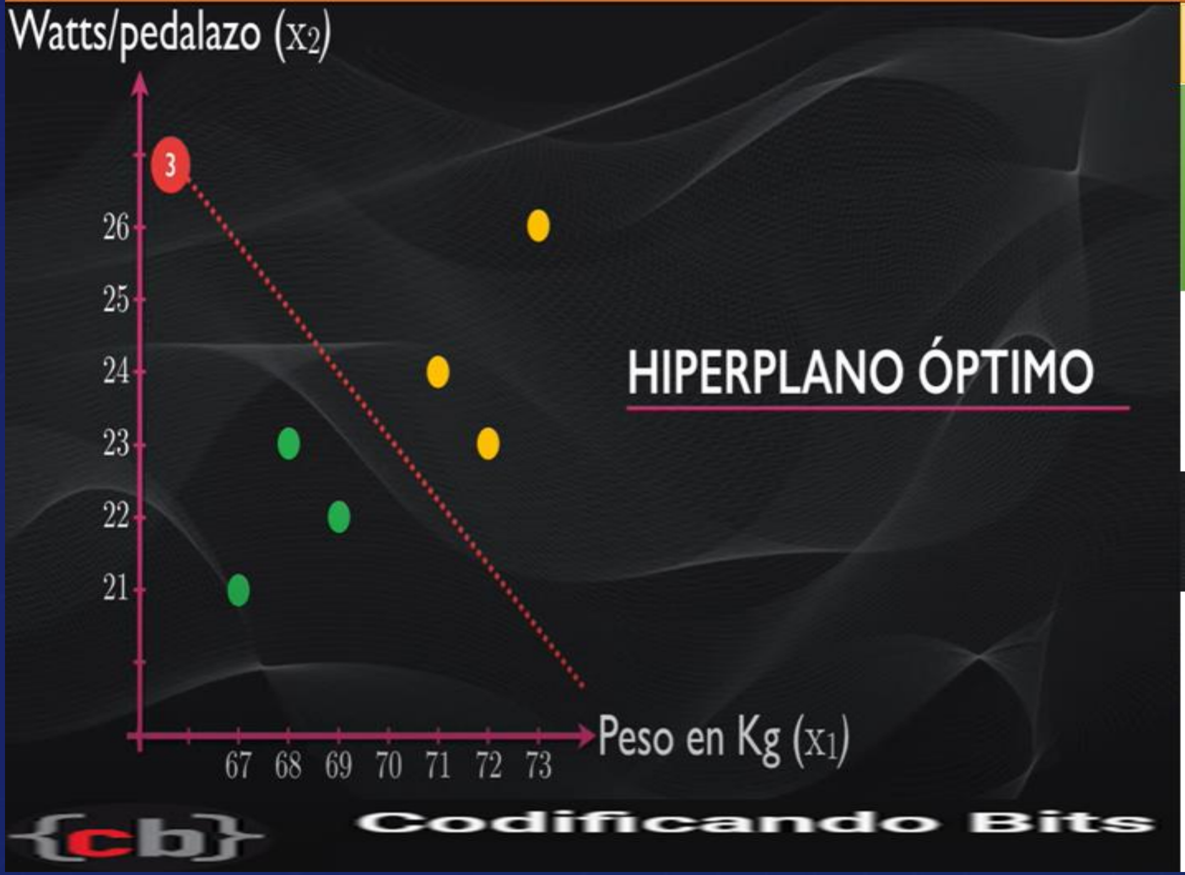
Máquinas de Soporte Vectorial VM)



Cuál es el mejor Hiperplano?

Profesor: Víctor Viera Balanta

Máquinas de Soporte Vectorial (SVM)



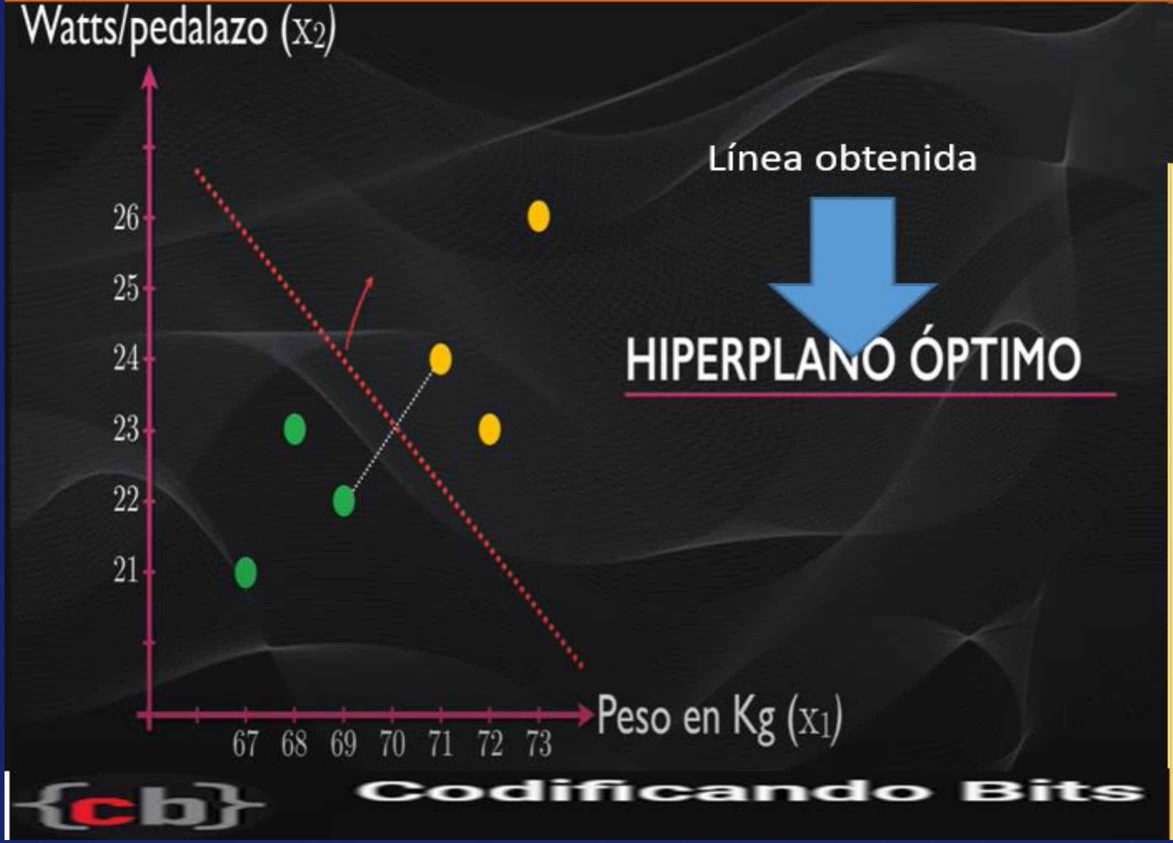
Cuál es el mejor Hiperplano?

La línea 3, está más alejada de las categorías y permite que no haya sesgo

HIPERPLANO ÓPTIMO

Profesor: Víctor Viera Balanta

Máquinas de Soporte Vectorial (SVM)



HARD MARGIN

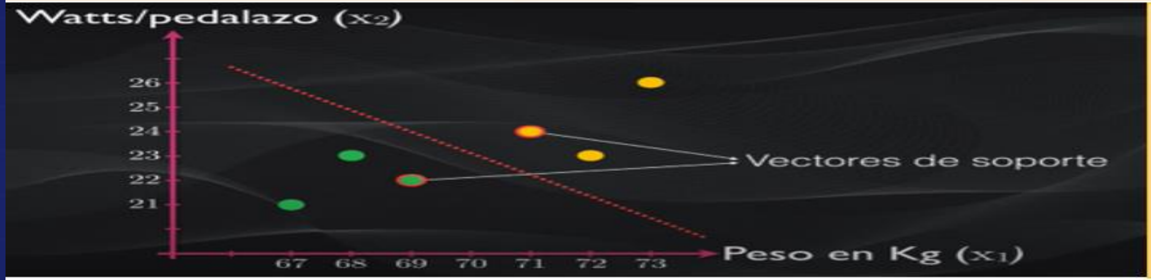
Se detectan los puntos más cercanos entre una clase y otra.

Encuentra la línea que los conecta

Traza una frontera perpendicular, que divide la línea en 2.

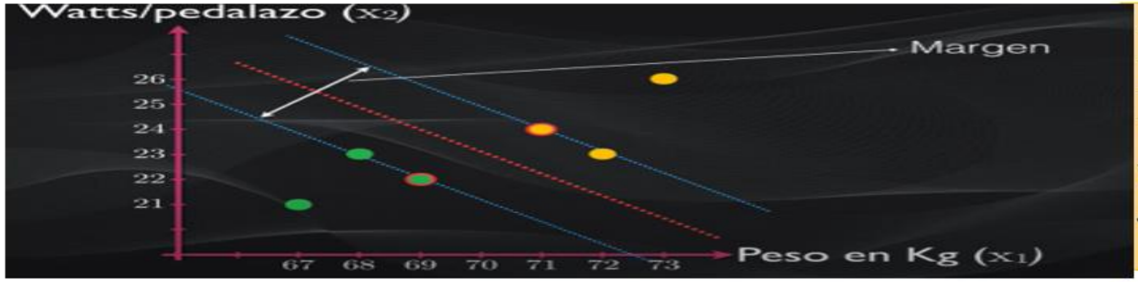
Profesor: Víctor Viera Balanta

Máquinas de Soporte Vectorial (SVM)



Vectores de soporte.

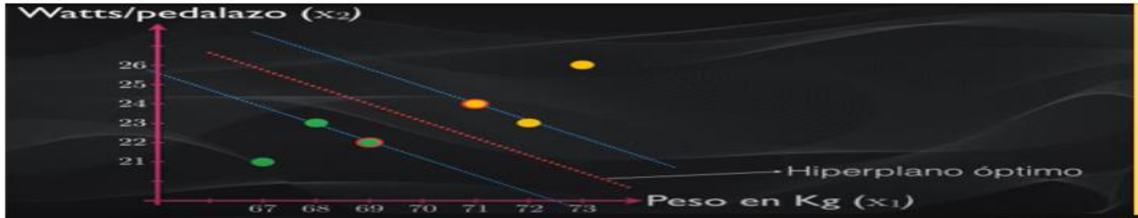
Puntos más cercanos entre una clase y otro.



Margen

Distancia entre el hiperplano y vectores de soporte.

Toca al menos un punto de cada clase.



Hiperplano óptimo

Frontera con mayor margen posible

COLOMBIA POTENCIA DE LA VIDA

TALENTO TECH

TIC

$$w^T x + b.$$

$$w_1x_1 + w_2x_2 + b = 0$$

coeficientes

Codificando Bits

Profesor: Víctor Viera Balanta

cymetria tecnalia colombia

Máquinas de Soporte Vectorial (SVM)

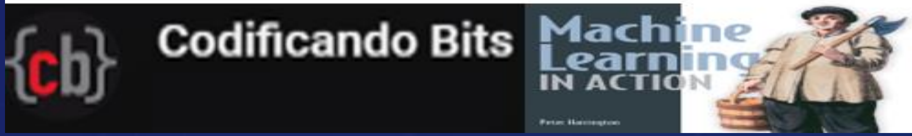
$$w^T x + b.$$

$$w_1 x_1 + w_2 x_2 + b = 0$$

coeficientes

Si queremos encontrar el distancia de A al plano de separación, debemos medir normal o perpendicularmente la línea. Esto está dado por

$$|w^T x + b| \,/\, ||w||$$



Máquinas de Soporte Vectorial (SVM)

El objetivo ahora es encontrar los valores de w y b que definirán el clasificador.

Se encontrar los puntos con menor margen, estos son los vectores de soporte.

cuando se encuentran los puntos con menor margen, se debe maximizar ese margen.

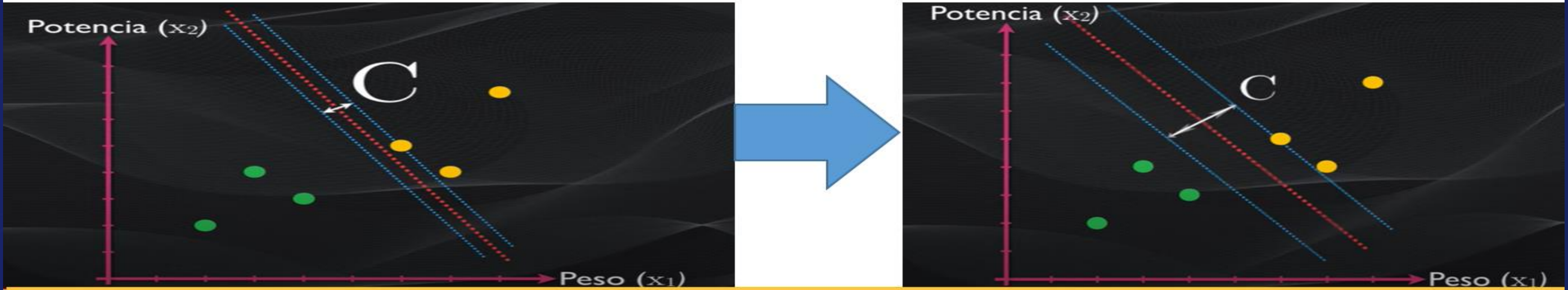
$$\arg\max_{w,b}\left\{\min_{n}\left(label\cdot\left(w^{T}x+b\right)\right)\cdot\frac{1}{\|w\|}\right\}$$

{cb} Codificando Bits

Máquinas de Soporte Vectorial (SVM)

Soft Margin

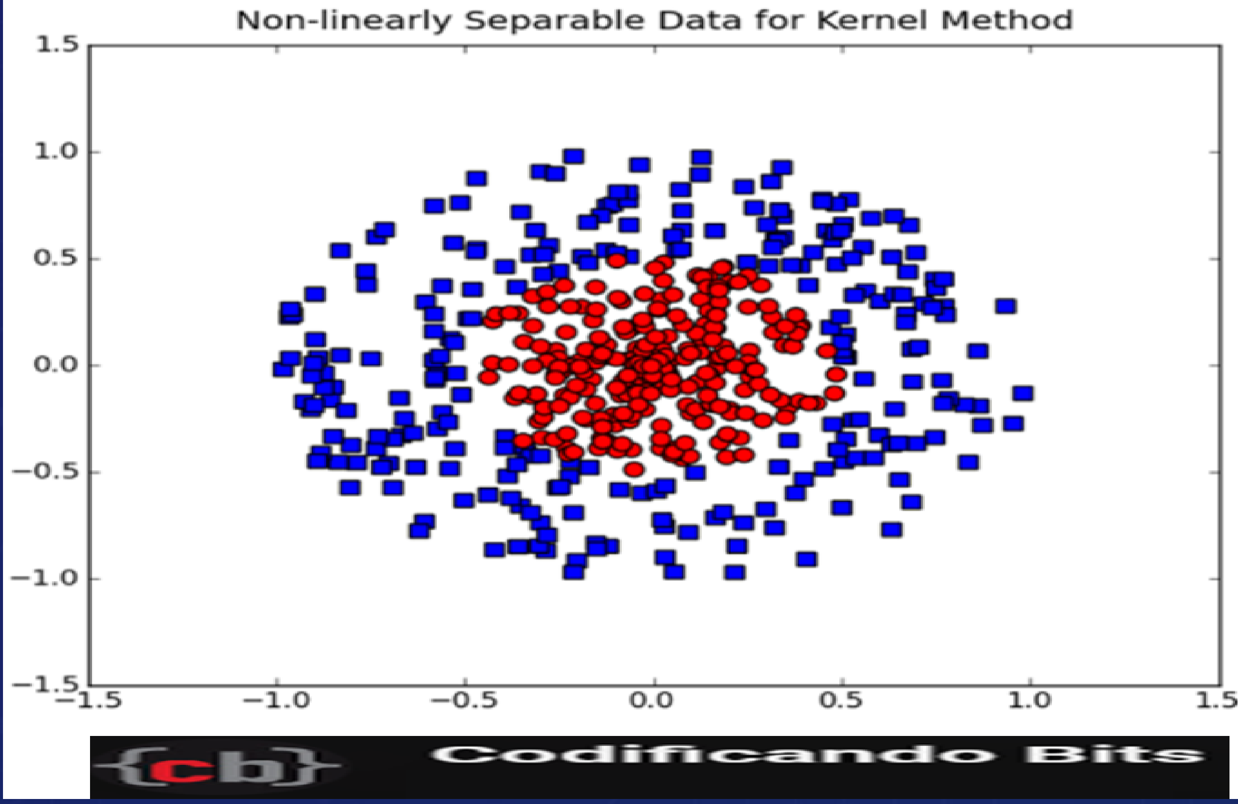
En ocasiones no es posible encontrar datos "perfectos", en estos casos se utiliza un Hiperparámetro C, con el que podemos aumentar o disminuir el margen.



Si el tamaño aumenta, el margen se reduce. Se escoge por prueba y error.. (error)

{cb} Codificando Bits

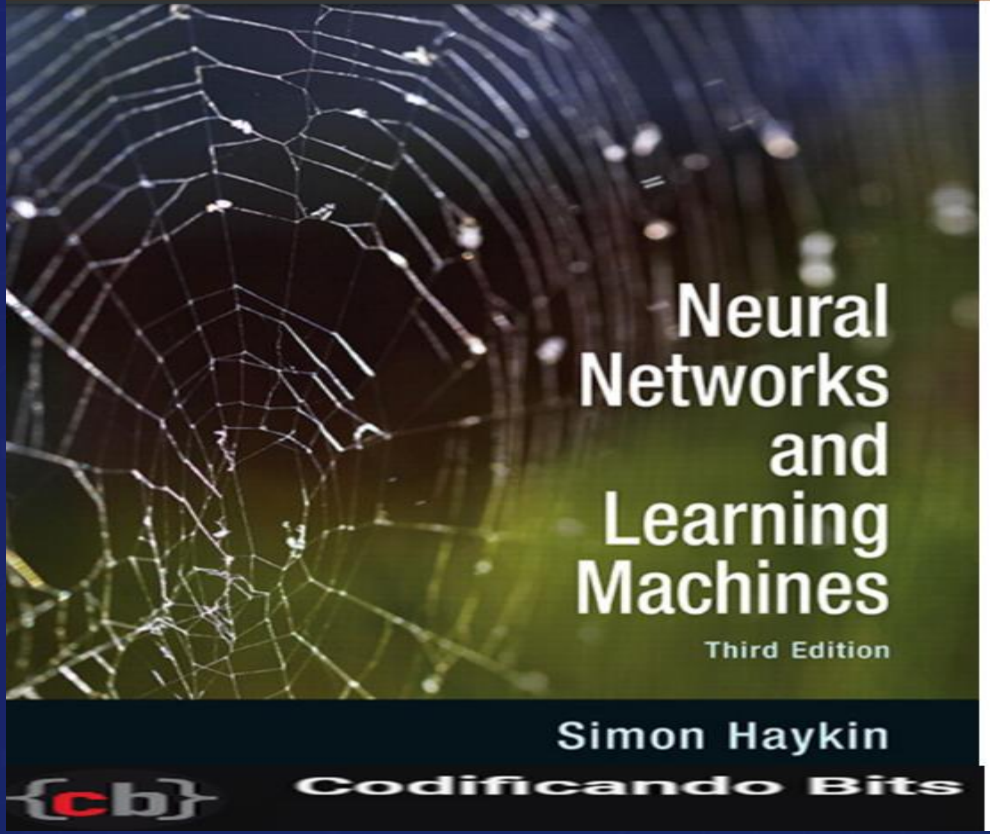
Máquinas de Soporte Vectorial (SVM)



Qué pasa si los datos tienen otro comportamiento?

Profesor: Víctor Viera Balanta

Máquinas de Soporte Vectorial (SVM)



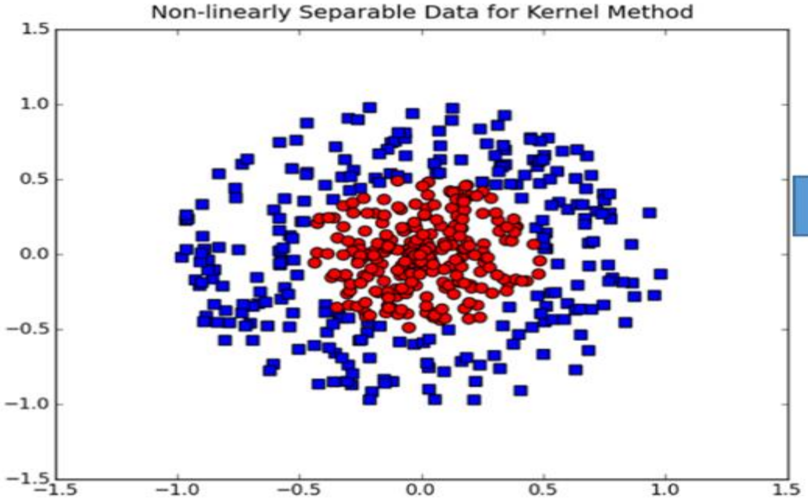
Kernel

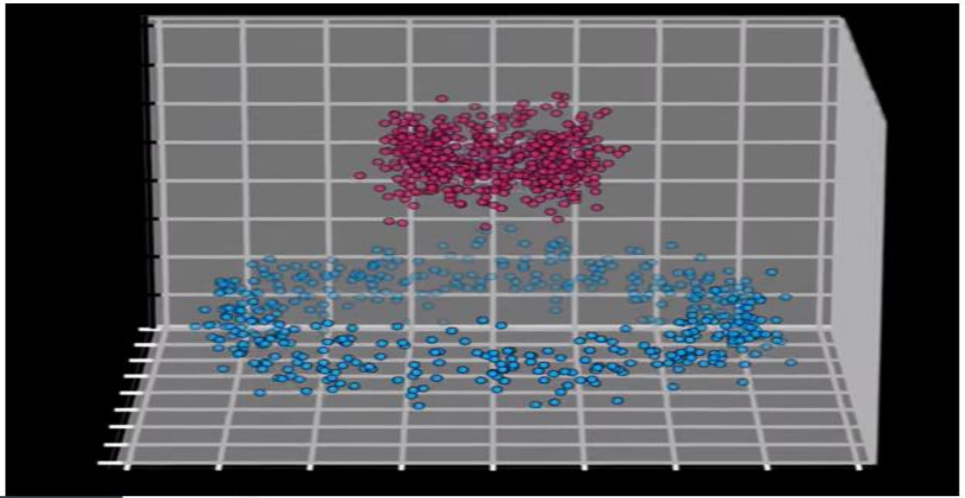
The Kernel Trick

Profesor: Víctor Viera Balanta

Máquinas de Soporte Vectorial (SVM)

Los datos se llevan a otra dimensión.

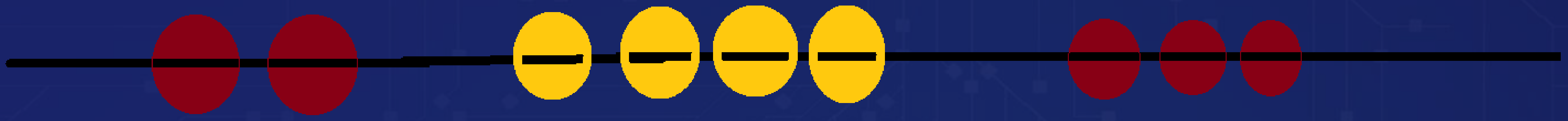




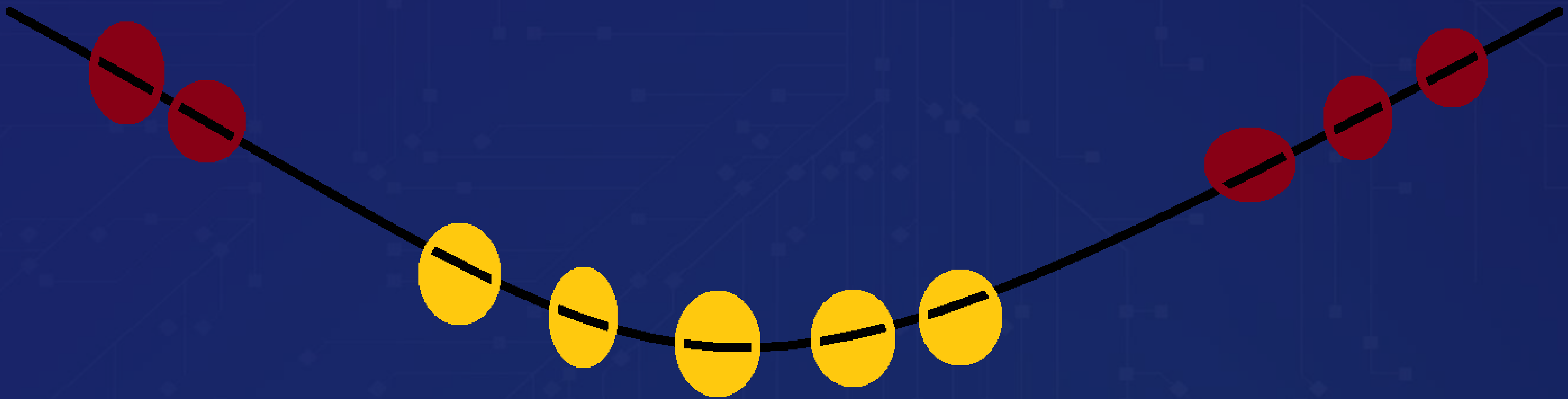
Se agregan más dimensiones a los datos.

{cb} Codificando Bits

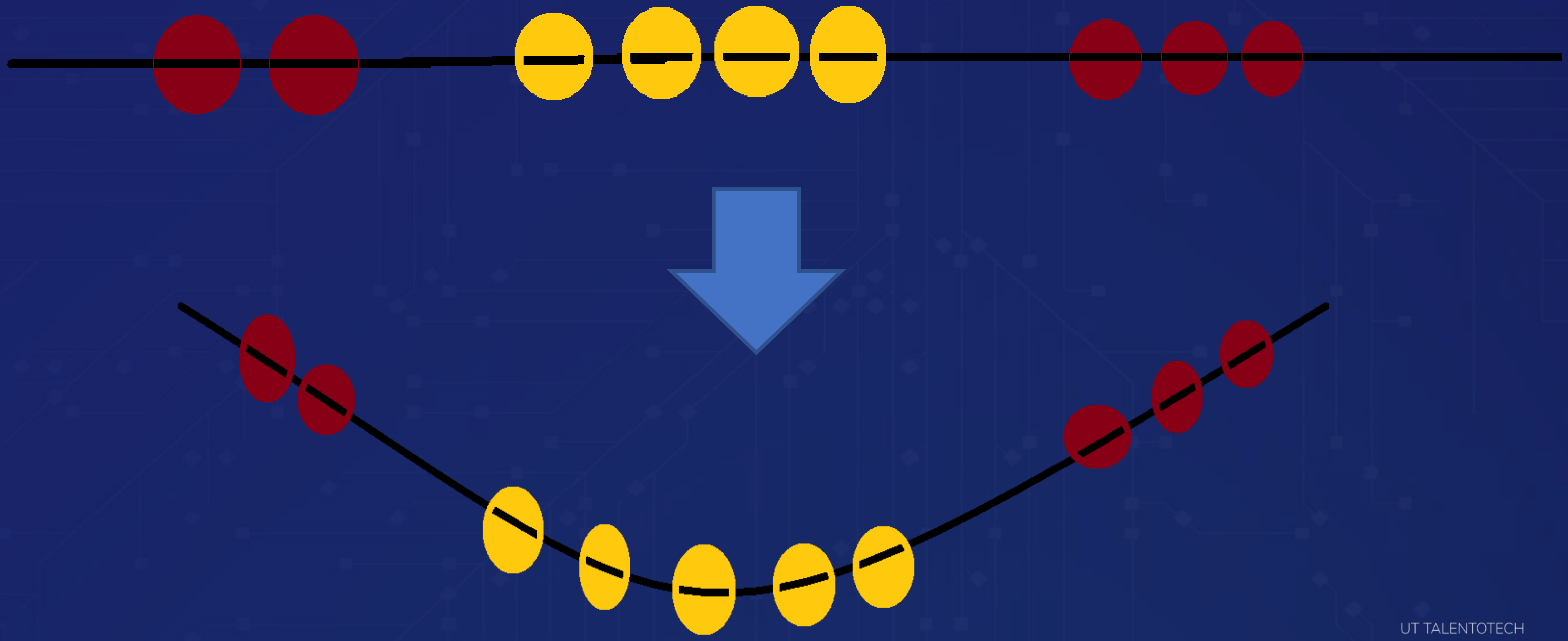
Profesor: Víctor Viera Balanta



Profesor: Víctor Viera Balanta



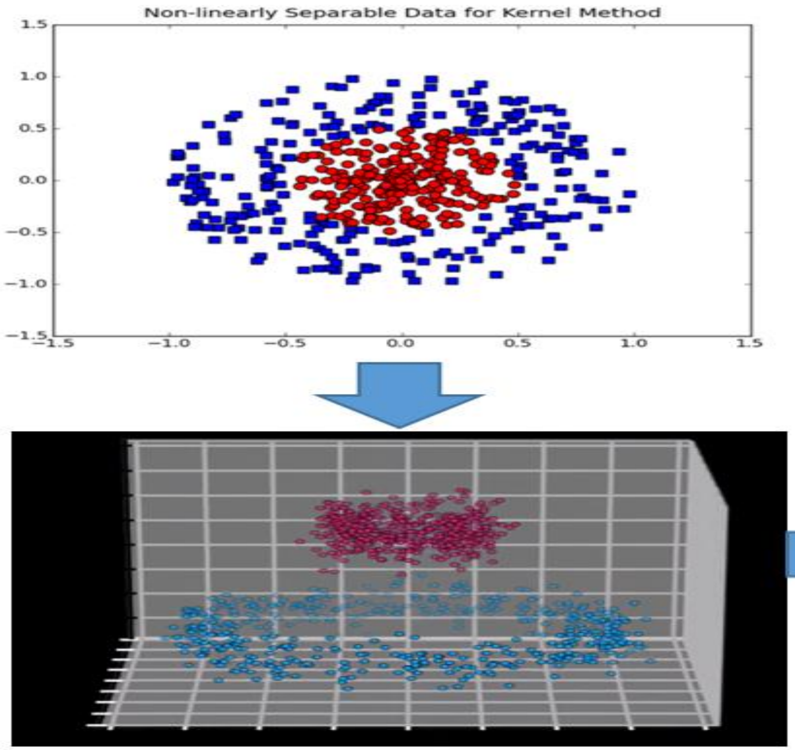
Profesor: Víctor Viera Balanta



Profesor: Víctor Viera Balanta

Profesor: Víctor Viera Balanta

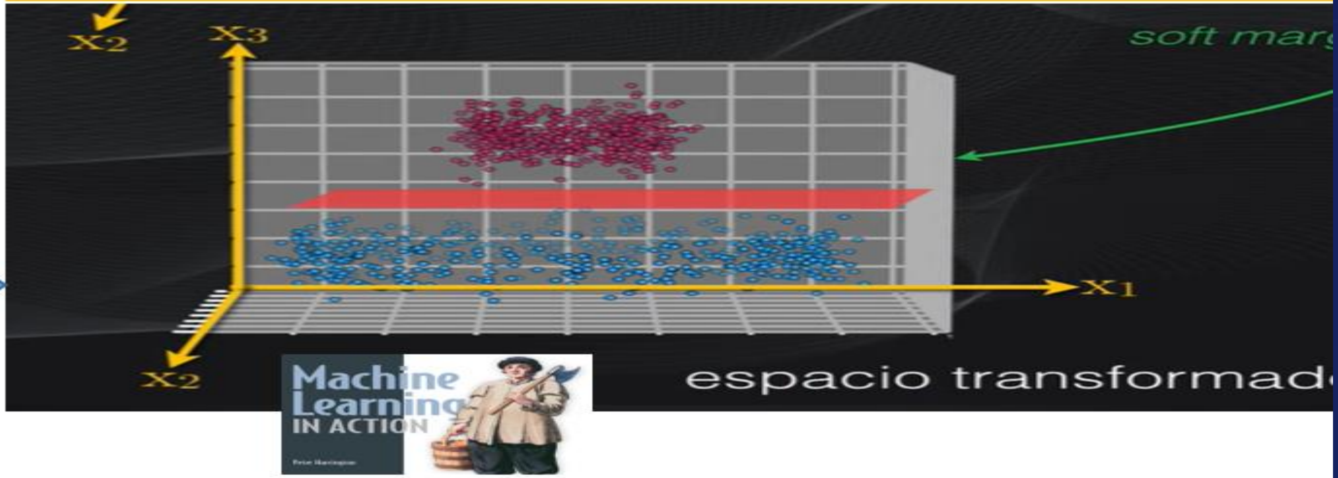
Máquinas de Soporte Vectorial (SVM)



Kernel

Agregar más dimensiones a los datos

Queda linealmente separable



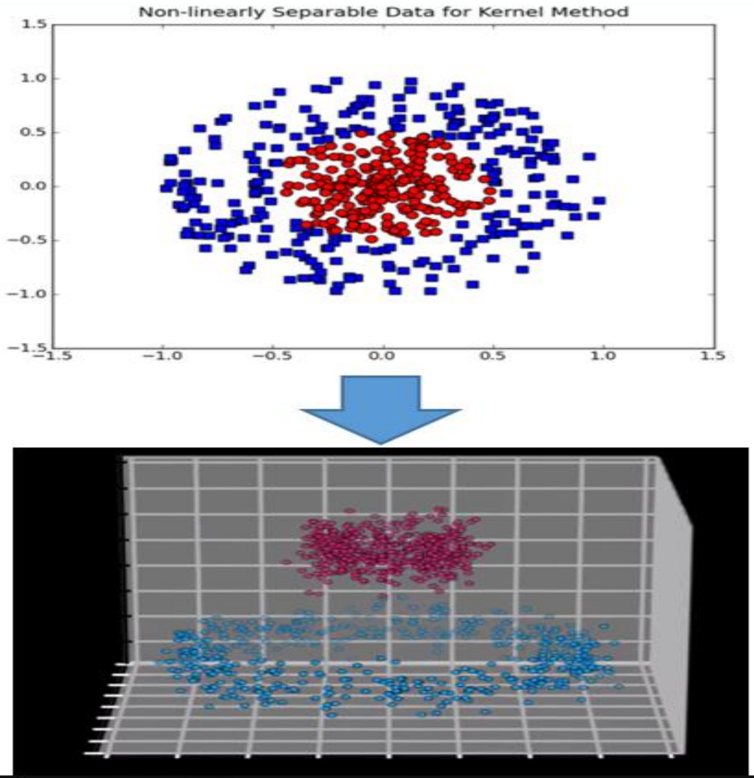
Profesor: Víctor Viera Balanta

COLOMBIA POTENCIA DE LA VIDA

TALENTO TECH

TIC

Máquinas de Soporte Vectorial (SVM)



Funciones Kernel

Polinomiales

Gaussianas

$$k(x,y) = exp\left(\frac{-\|x-y\|^2}{2\sigma^2}\right)$$

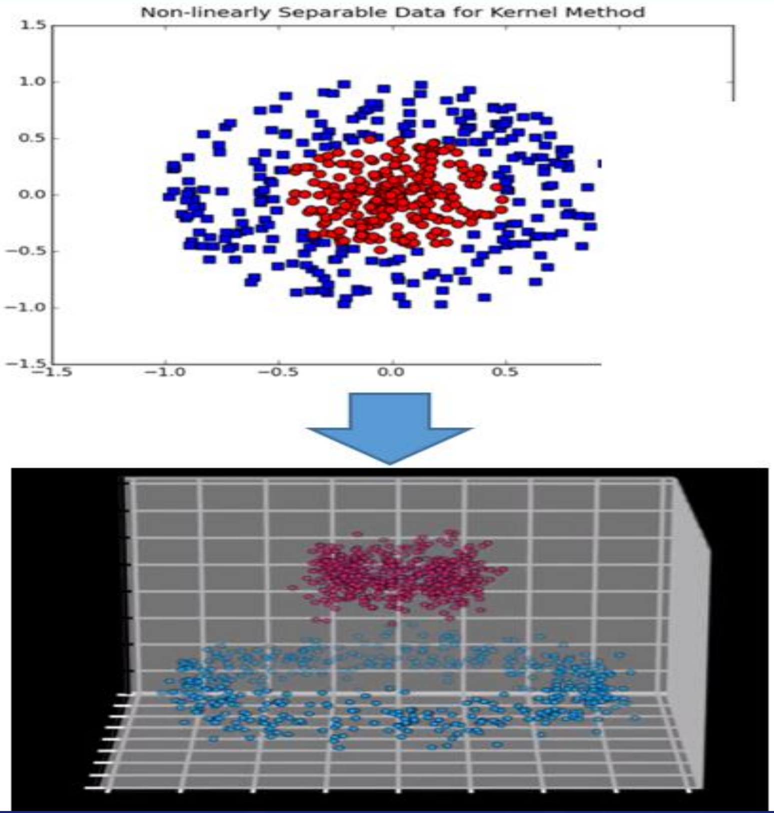
donde alpha es un parámetro definido por el usuario que determina el "alcance", o la rapidez con la que este cae a 0

{cb} Codificando Bits

Profesor: Víctor Viera Balanta

cymetria tecnalia colombia

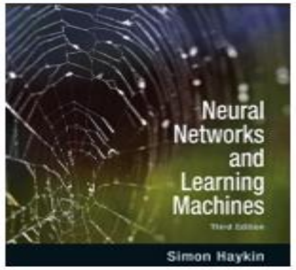
Máquinas de Soporte Vectorial (SVM)



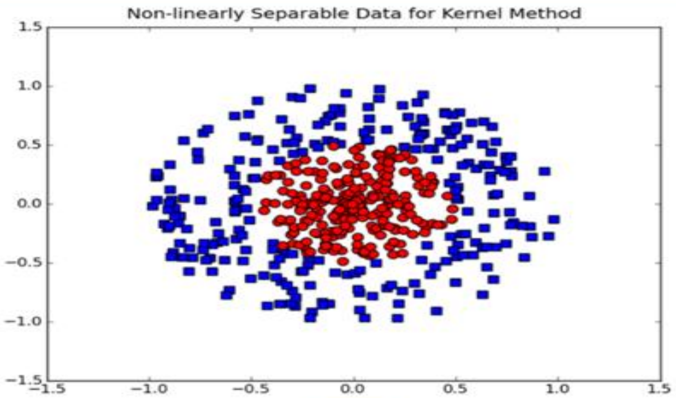
$$\mathcal{T} = \{\mathbf{x}_i, d_i\}_{i=1}^{N}$$

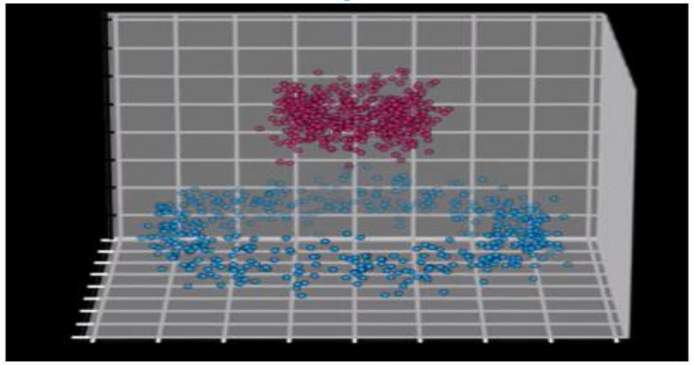
Etiquetas

Datos Independientes

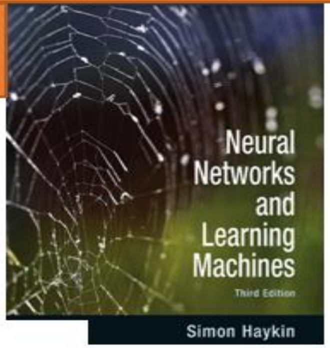


Máquinas de Soporte Vectorial (SVM)





Multiplicadores de Lagrange



$$\mathcal{T} = \{\mathbf{x}_i, d_i\}_{i=1}^{N}$$

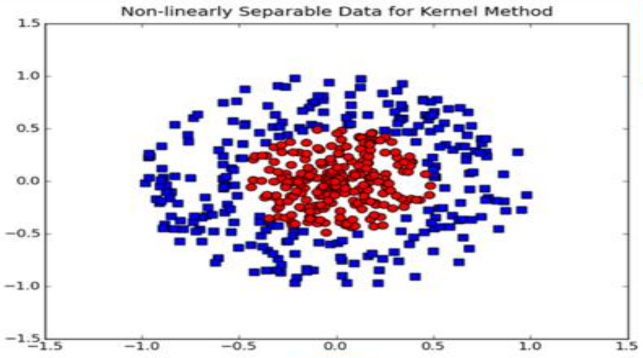
$$d_i(\mathbf{w}^T\mathbf{x}_i + b) \geq 1 \qquad \text{for } i = 1, 2, \ldots, N$$

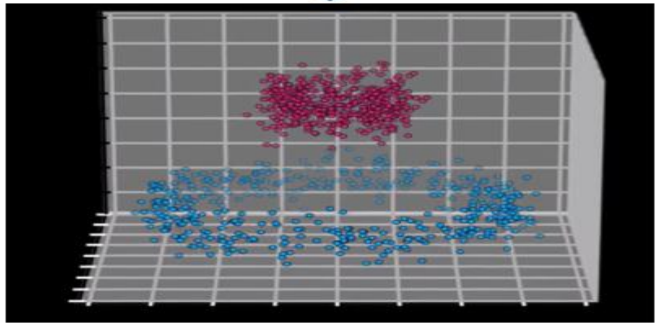
$$\sum_{i=1}^{N} \alpha_i d_i = 0$$

>=0

Restricción:
La sumatoria de los multiplicadores de Lagrange por las etiquetas debe ser igual a cero

Máquinas de Soporte Vectorial (SVM)





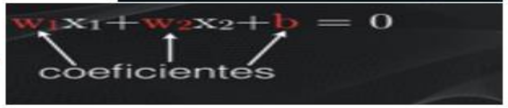
Multiplicadores de Lagrange

$$\mathscr{T} = \{\mathbf{x}_i, d_i\}_{i=1}^{N}$$

$$d_i(\mathbf{w}^T\mathbf{x}_i + b) \geq 1 \qquad \text{for } i = 1, 2, ..., N$$

$$\sum_{i=1}^{N} \alpha_i d_i = 0$$

$$\mathbf{w}^{\mathbf{T}}\mathbf{x}+\mathrm{b}.$$



$$y = w_o x + b_o$$

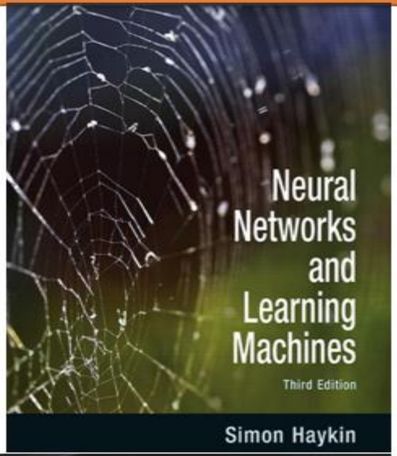
In this case, y can be $+1$ (class 1) or -1 (class 2).

$$w_o = \sum_{i=1}^{N_s} \alpha_{o,i} d_i x_i$$

and

$$b_o = 1 - \sum_{i=1}^{N_s} \alpha_{o,i} d_i x_i^T x^{(s)}$$

Producto punto



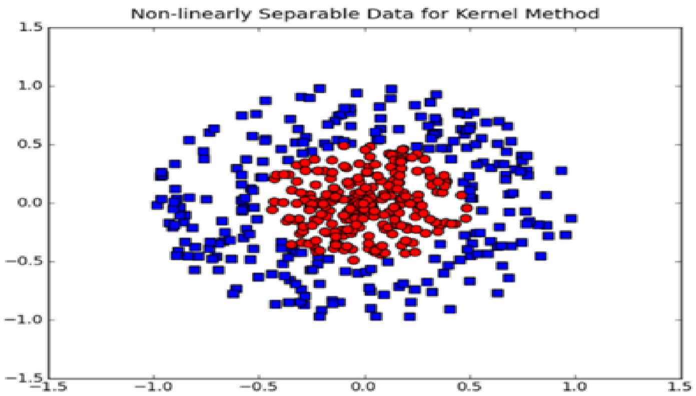
https://dai.fmph.uniba.sk/courses/NN/haykin.neural-networks.3ed.2009.pdf

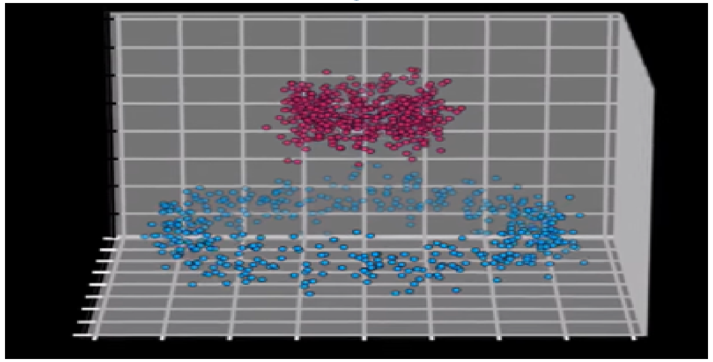
COLOMBIA POTENCIA DE LA VIDA

TALENTO TECH

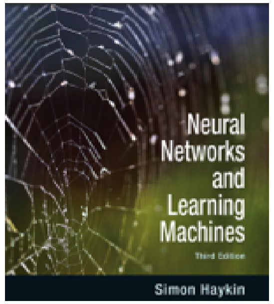
TIC

Máquinas de Soporte Vectorial (SVM)





Multiplicadores de Lagrange



$$\mathscr{T} = \{\mathbf{x}_i, d_i\}_{i=1}^{N}$$

$$d_i(\mathbf{w}^T\mathbf{x}_i + b) \geq 1 \quad \text{for } i = 1, 2, ..., N$$

$$(1) \quad \sum_{i=1}^{N} \alpha_i d_i = 0$$

$$(2) \quad \alpha_i \geq 0 \quad \text{for } i = 1, 2, ..., N$$

Restricciones

$$Q(\alpha) = \sum_{i=1}^{N} \alpha_i - \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} \alpha_i \alpha_j d_i d_j \mathbf{x}_i^T \mathbf{x}_j \qquad (6.16)$$

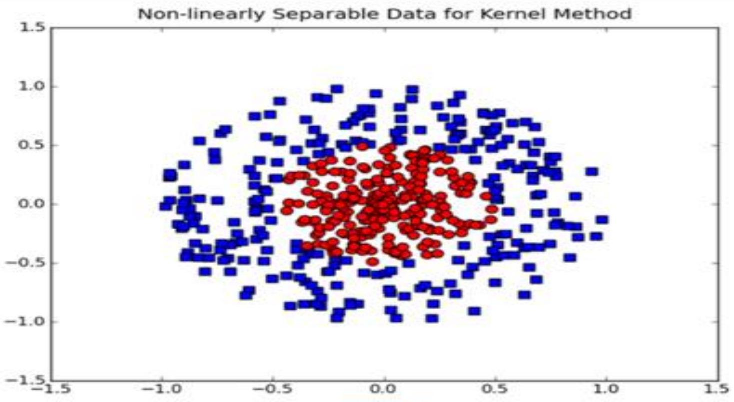
La función de costo depende: de los multiplicadores de Lagrange para cada punto, las etiquetas y los patrones (los Xs)

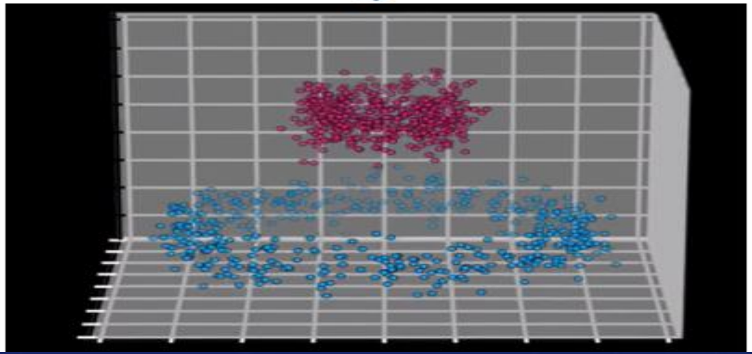
Maximizar la función de costo. Representación Dual

Profesor: Víctor Viera Balanta

cymetria tecnalia colombia

Máquinas de Soporte Vectorial (SVM)



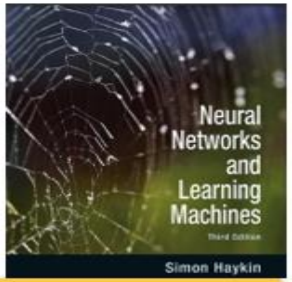


Multiplicadores de Lagrange

$$Q(\alpha) = \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j d_i d_j \mathbf{x}_i^T \mathbf{x}_j \qquad (6.16)$$

$$Q(\alpha) = \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j d_i d_j k(\mathbf{x}_i, \mathbf{x}_j) \qquad (6.39)$$

Kernel



COLOMBIA POTENCIA DE LA VIDA

TALENTO TECH

TIC

¡Gracias!

UT TALENTOTECH

cymetria

tecnalia colombia