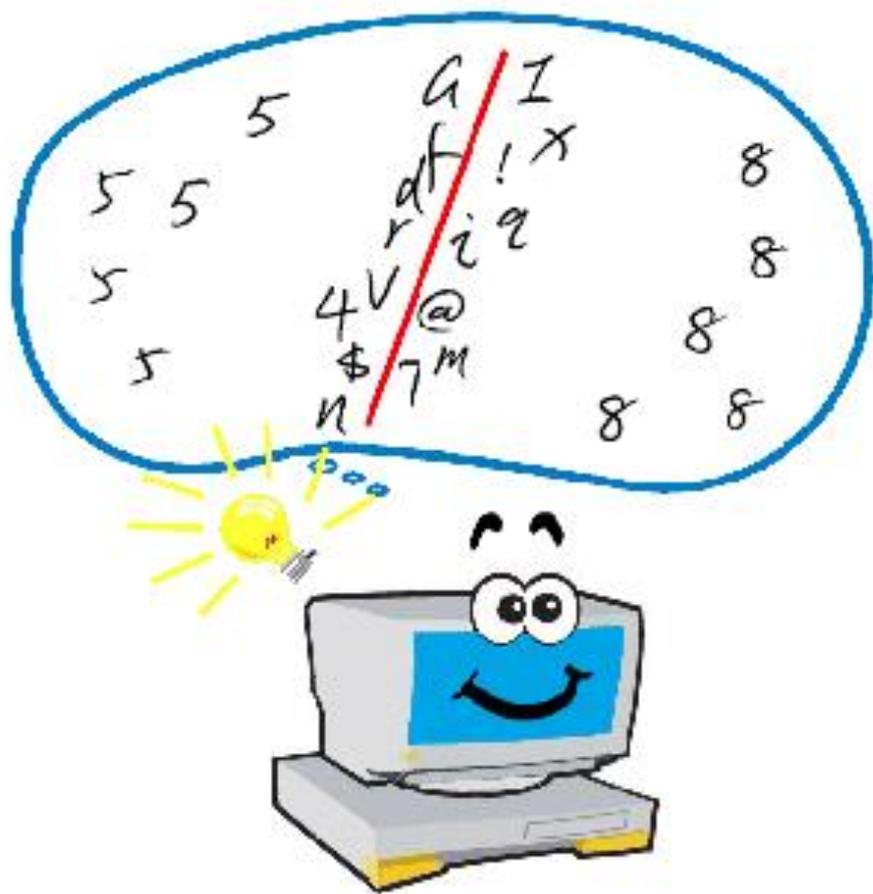


XIII Escuela de Probabilidad y Estadística

SVM's: Aplicaciones en Clasificación de Imágenes

17 de marzo de 2015, Domingo Iván Rodríguez González, CIMAT

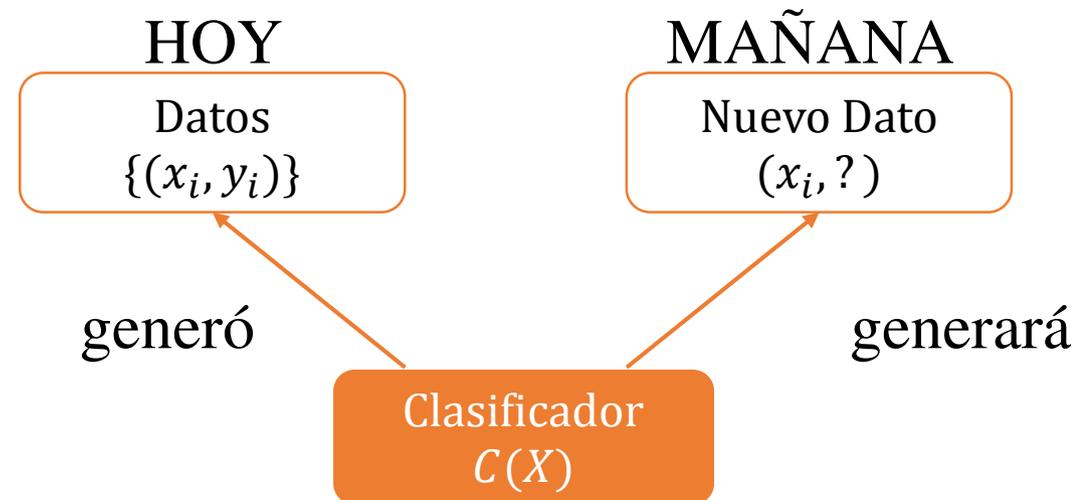
Introducción



5

Recordatorio: Problema General de Clasificación

- Dado un vector $X = (X_1, X_2, \dots, X_p)$ de características y Y una variable cualitativa de respuesta (categoría):



- Queremos construir un clasificador $C(X)$ que asigne una etiqueta del conjunto de categorías posibles a una nueva observación x_i .

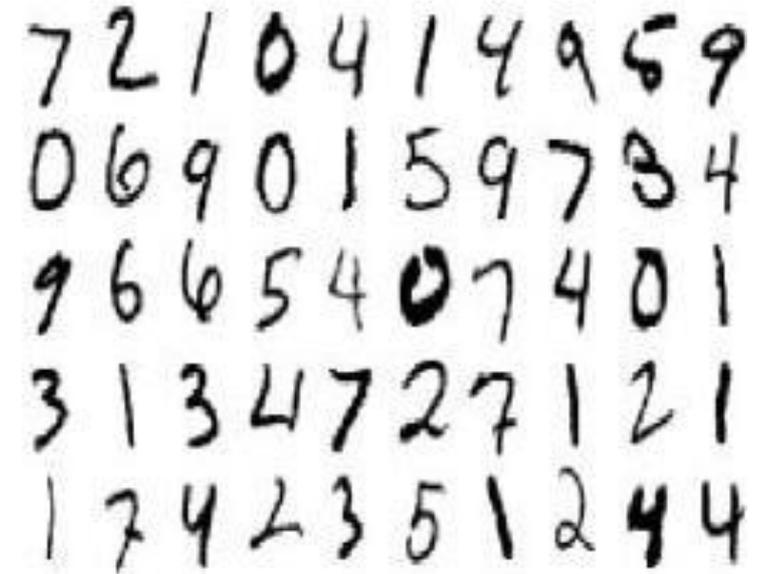
Problemas Típicos de Clasificación en Aprendizaje Máquina

- Determinar si un email es *spam* en base al contenido del texto.



Problemas Típicos de Clasificación en Aprendizaje Máquina

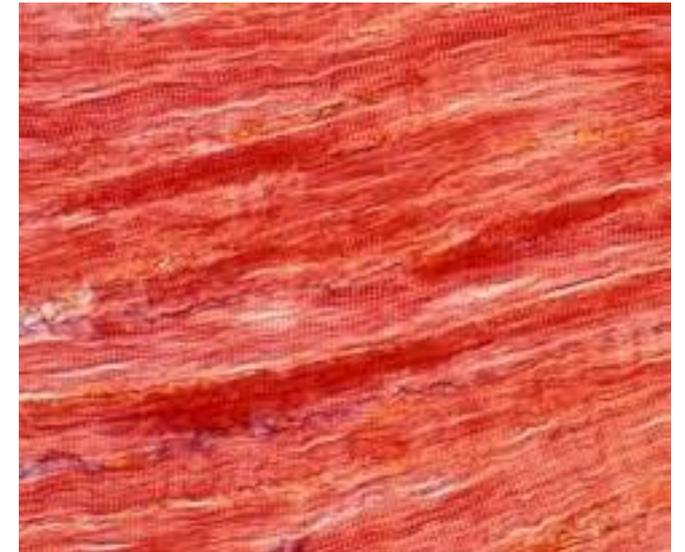
- Determinar si un email es *spam* en base al contenido del texto.
- Reconocer los números escritos manualmente de un código postal.



7 2 1 0 4 1 4 9 5 9
0 6 9 0 1 5 9 7 3 4
9 6 6 5 4 0 7 4 0 1
3 1 3 4 7 2 7 1 2 1
1 7 4 2 3 5 1 2 4 4

Problemas Típicos de Clasificación en Aprendizaje Máquina

- Determinar si un email es *spam* en base al contenido del texto.
- Reconocer los números escritos manualmente de un código postal.
- Clasificar un tipo de tejido humano en una de múltiples clases de cáncer.



Problemas Típicos de Clasificación en Aprendizaje Máquina

- Determinar si un email es *spam* en base al contenido del texto.
- Reconocer los números escritos manualmente de un código postal.
- Clasificar un tipo de tejido humano en una de múltiples clases de cáncer.
- Asignar a cada pixel de una imagen satelital una categoría de acuerdo al uso del suelo.



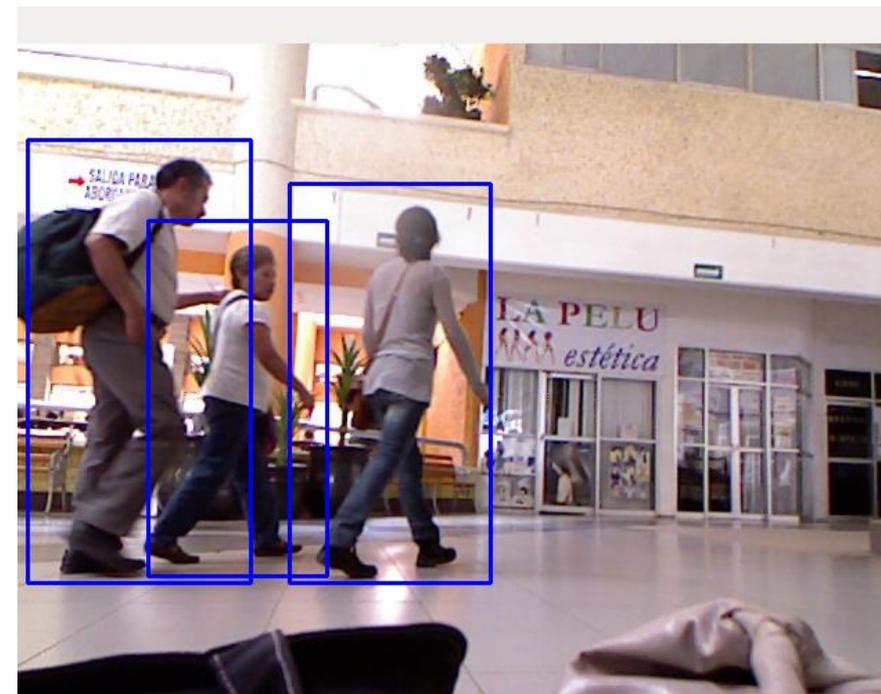
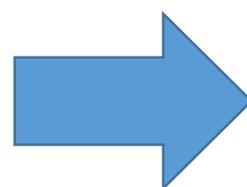
Clasificación de Imágenes

APLICACIÓN 1 → Detección de Peatones



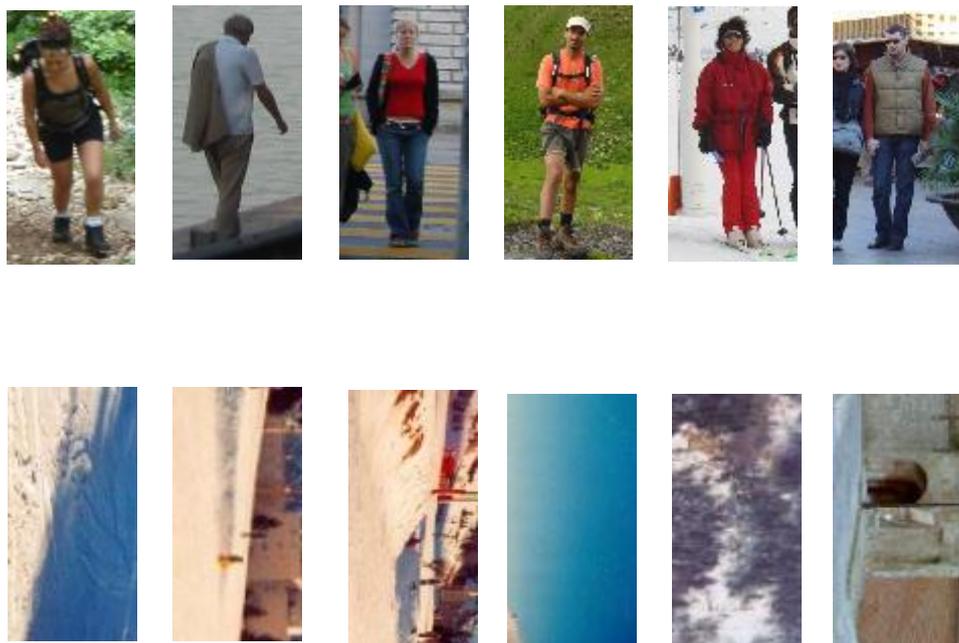
Clasificación de Imágenes

APLICACIÓN 1 → Detección de Peatones

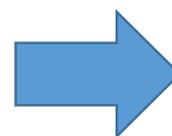


Clasificación de Imágenes

Conjunto de Datos

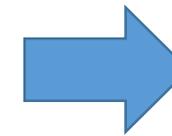


Procesamiento



Datos
 $\{(x_i, y_i)\}$

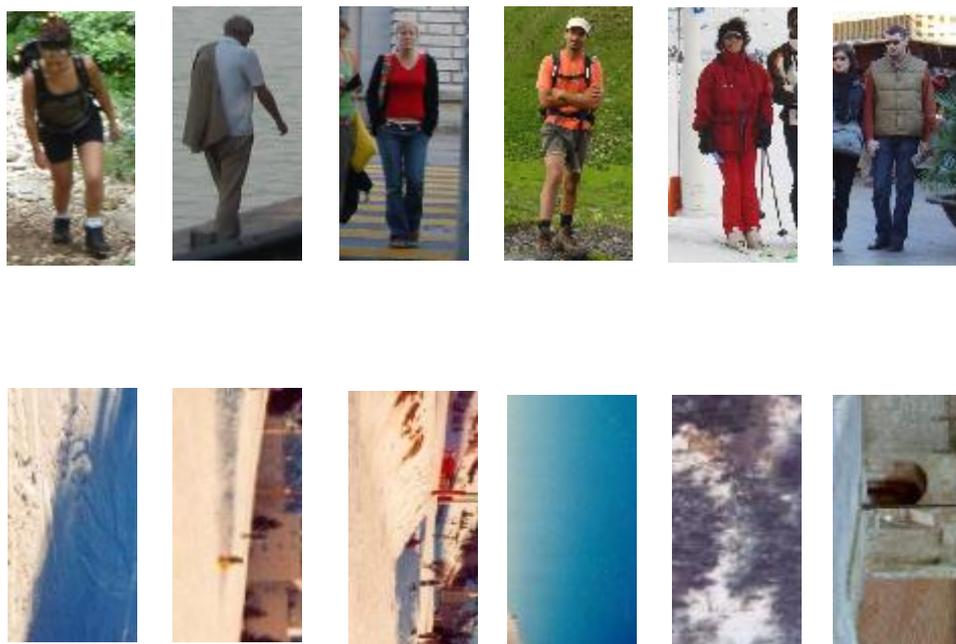
Entrenamiento



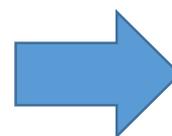
Clasificador
 $C(X)$

Clasificación de Imágenes

Conjunto de Datos

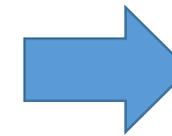


Procesamiento



Datos
 $\{(x_i, y_i)\}$

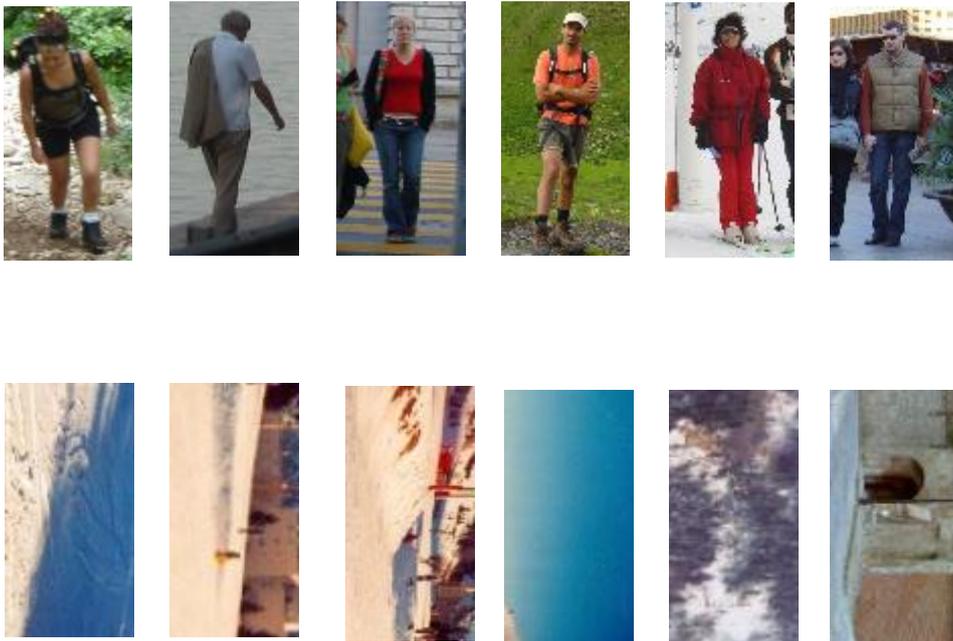
Entrenamiento



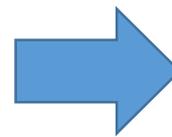
Clasificador
 $C(X)$

Clasificación de Imágenes

Conjunto de Datos

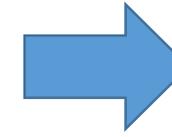


Procesamiento



Datos
 $\{(x_i, y_i)\}$

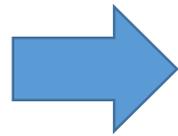
Entrenamiento



Clasificador
 $C(X)$

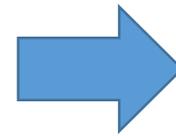
Clasificación de Imágenes

Clasificación de nuevas observaciones:



Procesamiento

Dato
($x_i, y_i = ?$)

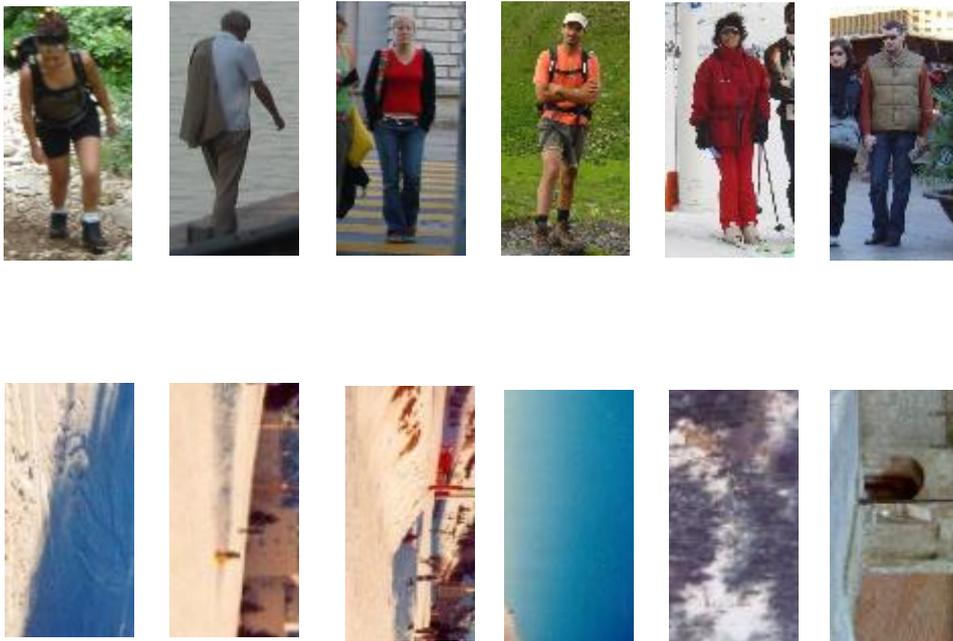


Clasificación

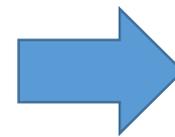
$y_i = \textit{persona}$

Clasificación de Imágenes

Conjunto de Datos

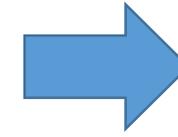


Procesamiento

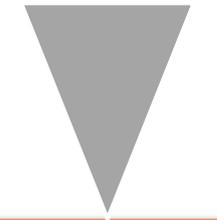


Datos
 $\{(x_i, y_i)\}$

Entrenamiento



Clasificador
 $C(X)$



SVM

Máquina de Soporte Vectorial (SVM)

An orange rounded rectangular button with the letters 'SVM' in white, bold, serif font.

SUPPORT VECTOR MACHINE

- Introducida en la década de los 90's por Vapnik y Cortes.
- No define un modelo probabilístico.

Máquina de Soporte Vectorial (SVM)

The logo consists of the letters 'SVM' in a white, serif font, centered within a rounded rectangular button with a gradient from light orange to dark orange.

SUPPORT VECTOR MACHINE

- Introducida en la década de los 90's por Vapnik y Cortes.
- No define un modelo probabilístico.

Encontrar un **HIPERPLANO** (en el espacio de las características)
que separe las dos clases de manera directa.

¿Qué es un Hiperplano de separación?

Definimos un HIPERPLANO en la dimensión p como un subespacio de dimensión $p - 1$ que cumple con la siguiente ecuación lineal:

$$\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p = 0 ,$$

donde $\beta = (\beta_1, \beta_2, \dots, \beta_p)$ es un vector ortogonal a la superficie del hiperplano, y β_0 es 0 si el hiperplano pasa por el origen.

¿Qué es un Hiperplano de separación?

Definimos un HIPERPLANO en la dimensión p como un subespacio de dimensión $p - 1$ que cumple con la siguiente ecuación lineal:

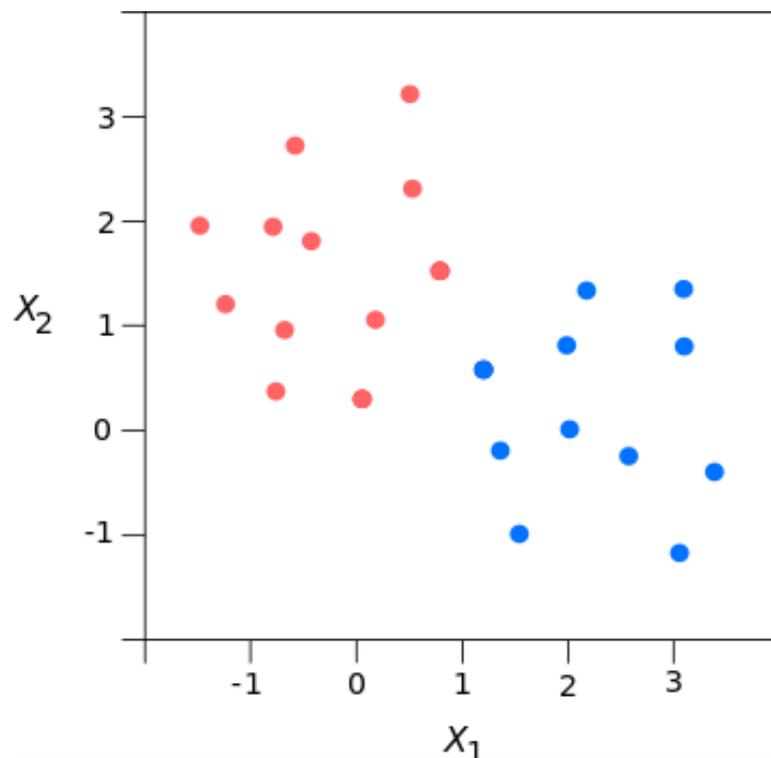
$$\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p = 0 ,$$

donde $\beta = (\beta_1, \beta_2, \dots, \beta_p)$ es un vector ortogonal a la superficie del hiperplano, y β_0 es 0 si el hiperplano pasa por el origen.

- En 2D el hiperplano es una recta.
- En 3D el hiperplano es un plano en dos dimensiones.

¿Qué es un Hiperplano de separación?

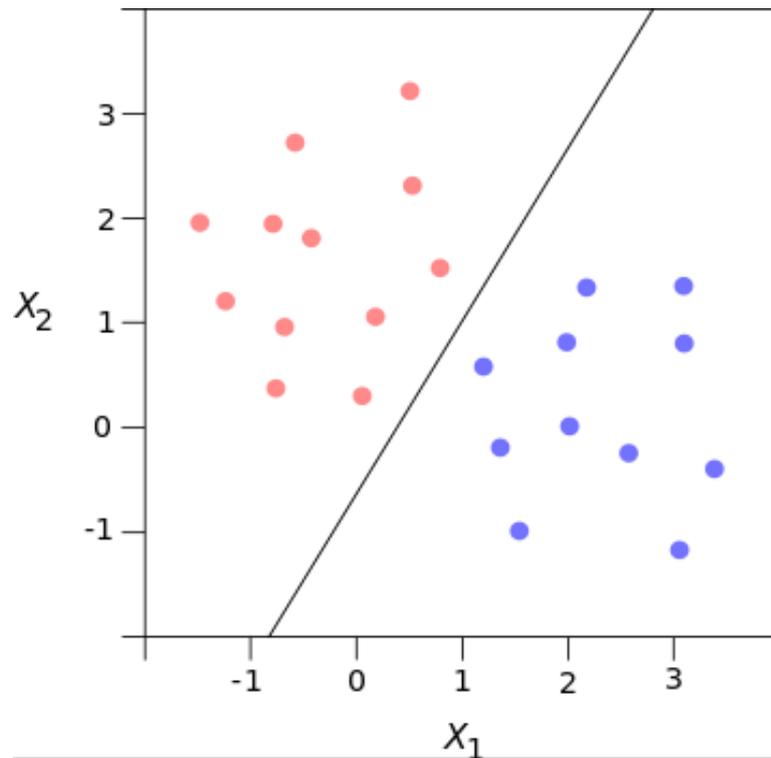
Revisemos el caso 2D:



Tenemos un conjunto de datos 2D separados en dos clases (rojo, azul).

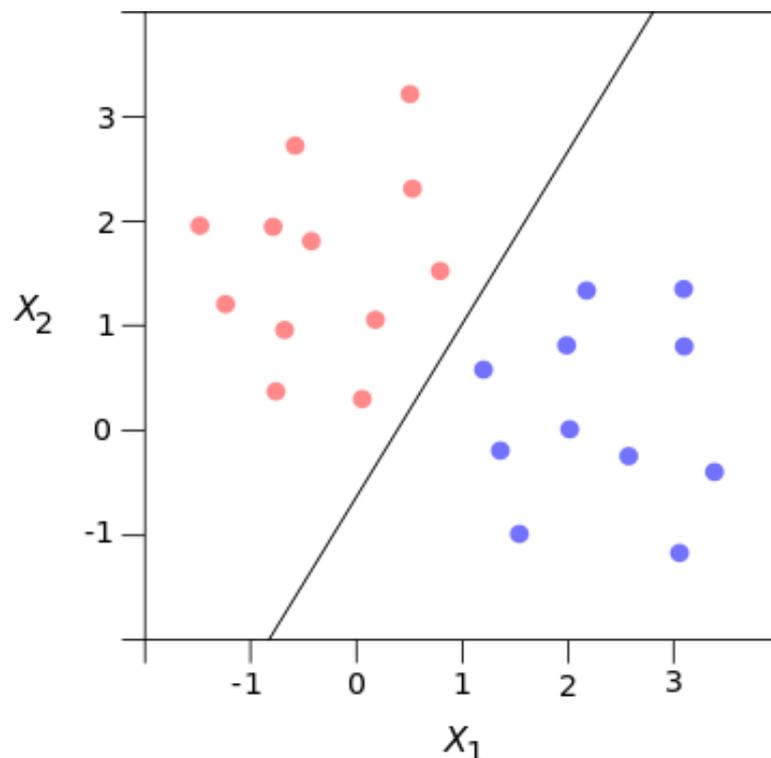
¿Qué es un Hiperplano de separación?

Seleccionamos una línea que separa las dos clases:



¿Qué es un Hiperplano de separación?

Seleccionamos una línea que separa las dos clases:



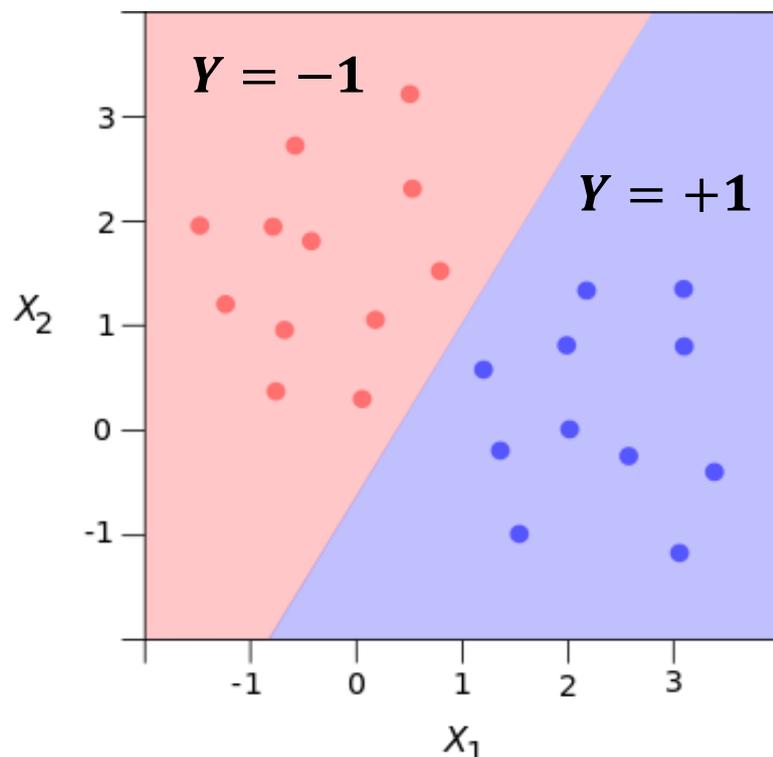
$$f(x) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p$$

Vemos que:

- si $f(x_i) = 0$, entonces x_i está en el hiperplano.
- si $f(x_i) > 0$, entonces x_i está de un lado del hiperplano.
- si $f(x_i) < 0$, entonces x_i está del otro lado del hiperplano.

¿Qué es un Hiperplano de separación?

Podemos separar las clases mediante el hiperplano:

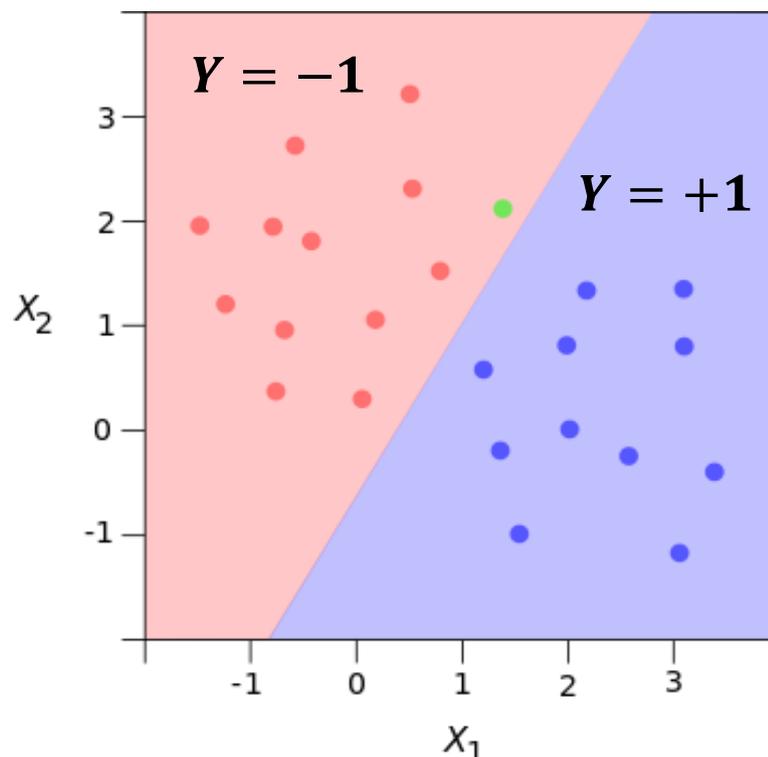


Recordemos que para cada x_i hay una y_i asociada.

Si $y_i f(x_i) > 0$ para todo i entonces $f(x) = 0$ define un **HIPERPLANO DE SEPARACIÓN**.

¿Qué es un Hiperplano de separación?

Clasificación de nuevas observaciones:



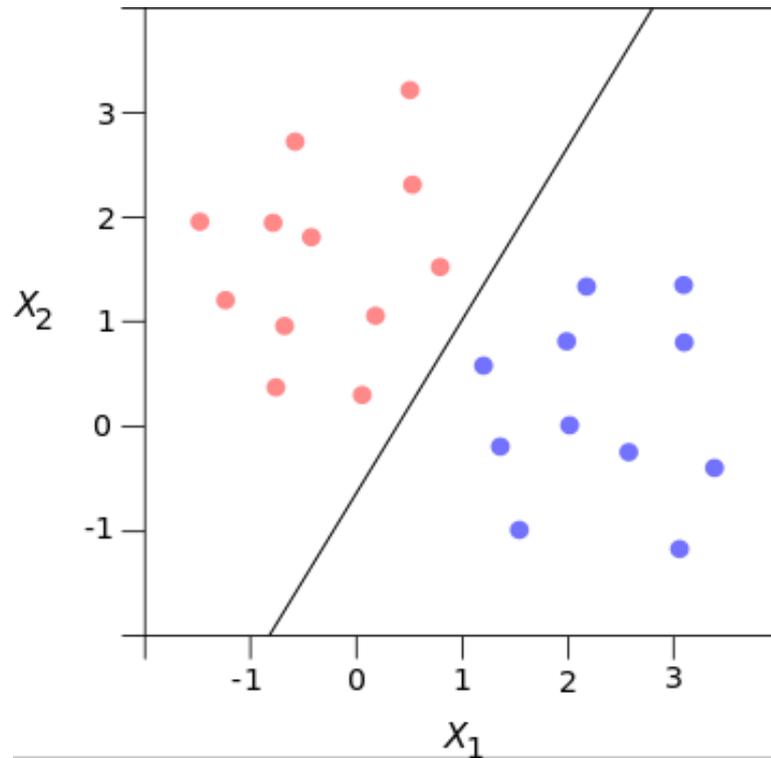
Para una nueva observación x_i determinamos la clase de acuerdo a:

Rojo si $y_i f(x_i) \leq 0$

Azul si $y_i f(x_i) > 0$

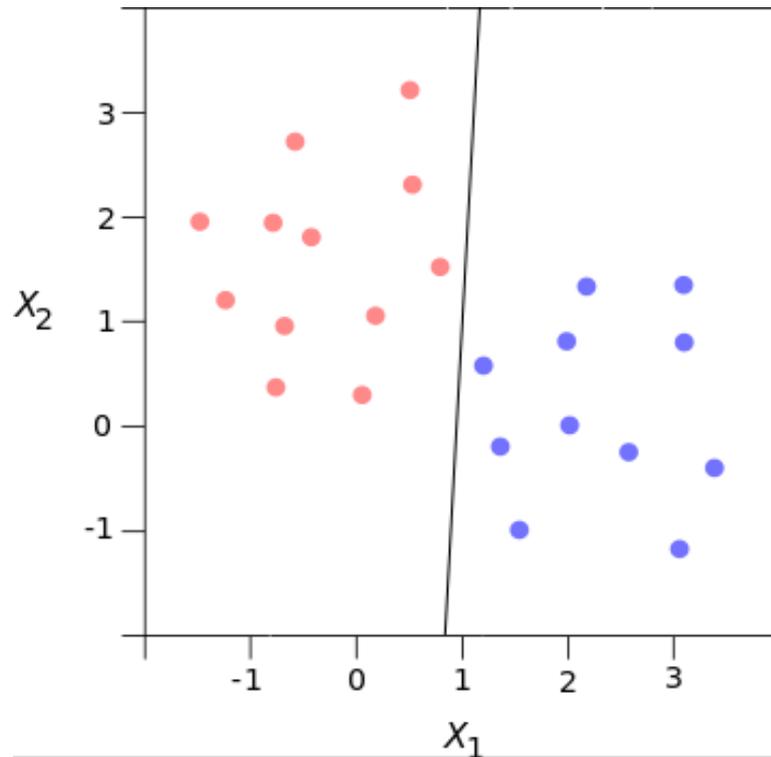
¿Cuál hiperplano seleccionar?

Muchas posibles opciones:



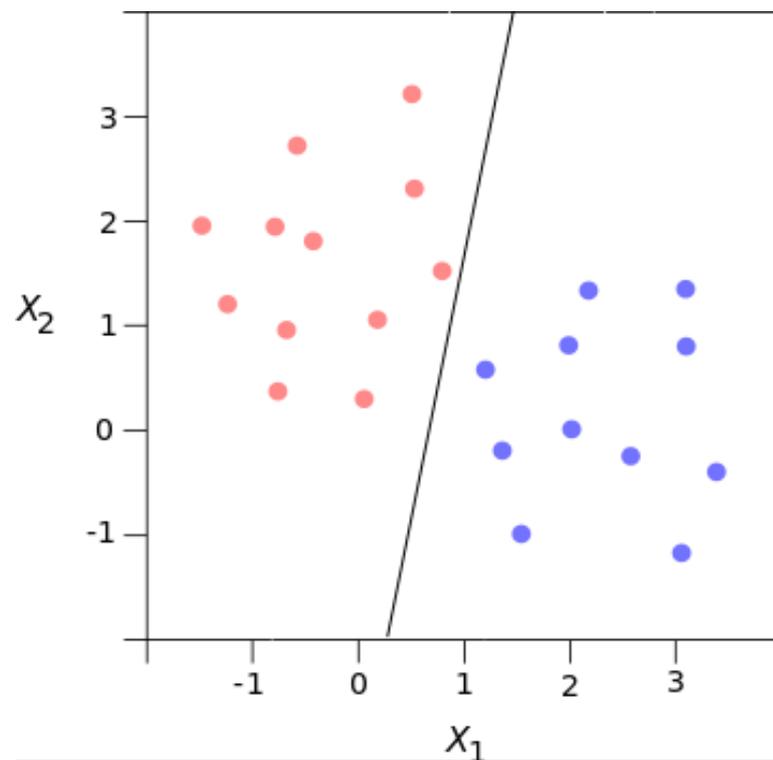
¿Cuál hiperplano seleccionar?

Muchas posibles opciones:



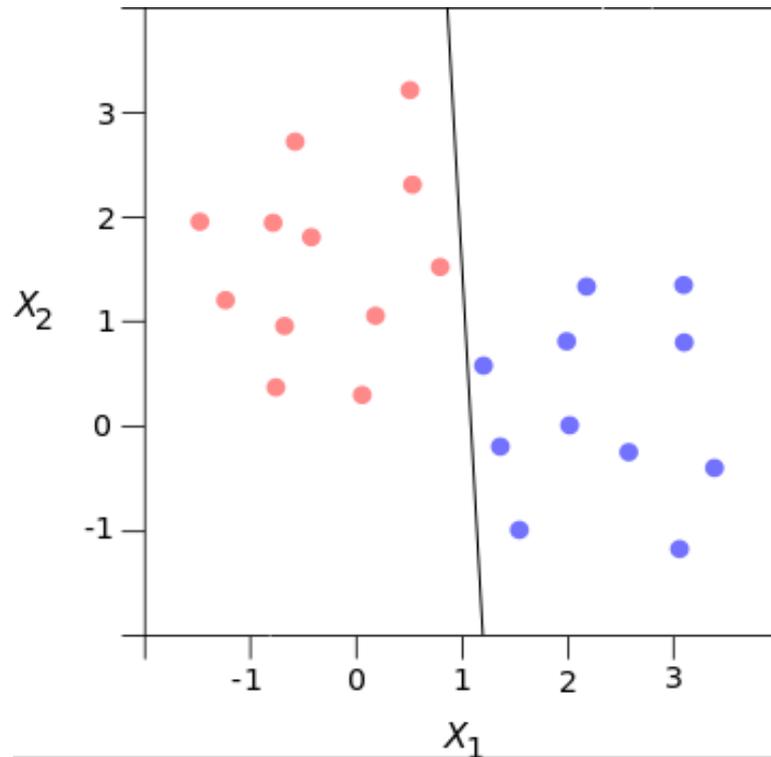
¿Cuál hiperplano seleccionar?

Muchas posibles opciones:



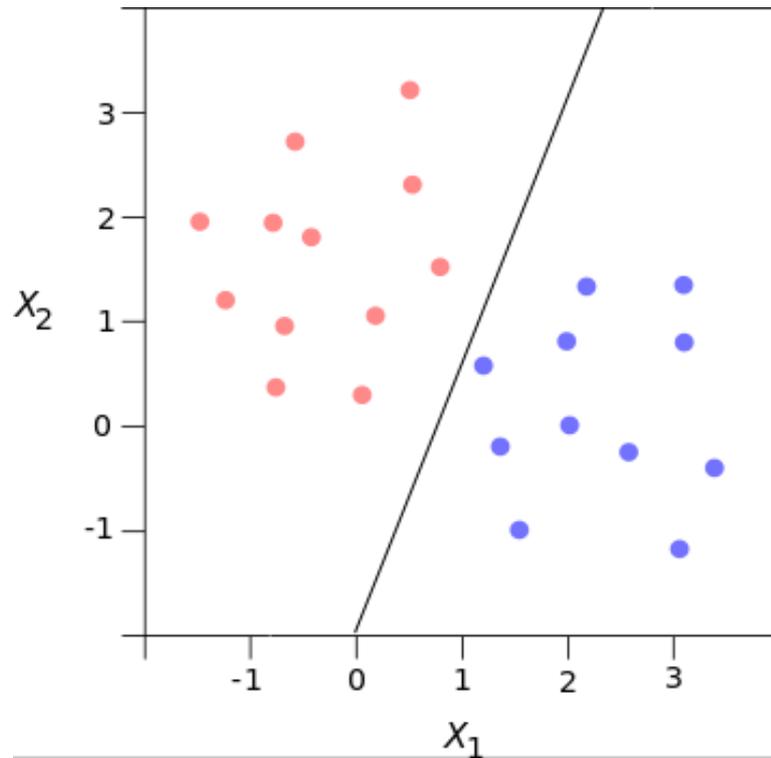
¿Cuál hiperplano seleccionar?

Muchas posibles opciones:



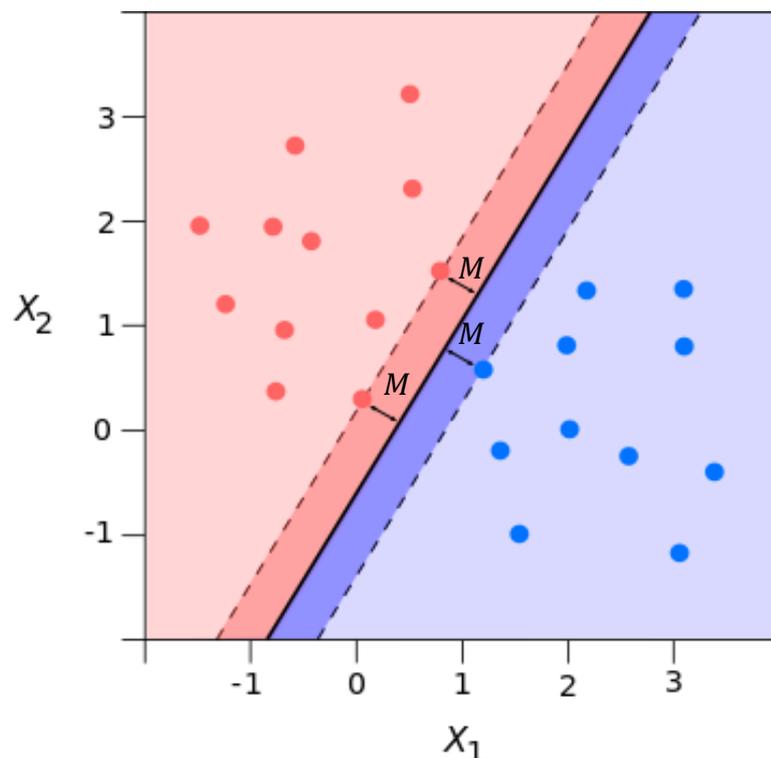
¿Cuál hiperplano seleccionar?

Muchas posibles opciones:



¿Cuál hiperplano seleccionar?

Encontrar la solución óptima para separar las clases:

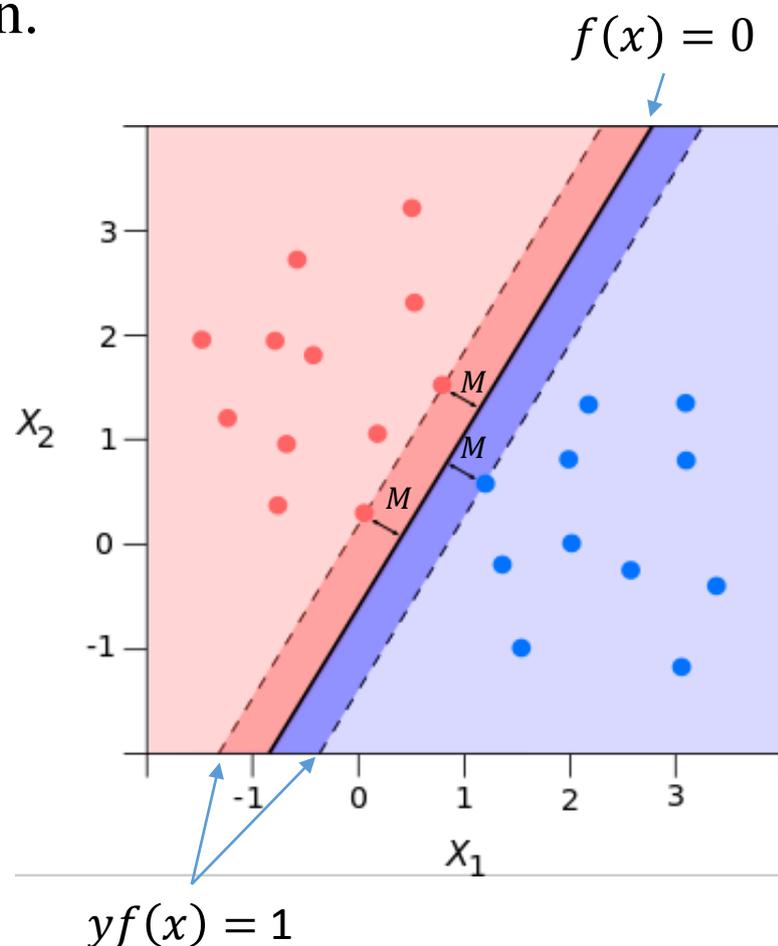


Seleccionar el hiperplano que maximiza el margen entre las dos clases.

- Todos los puntos están en el lado correcto del margen.
- Para todo x_i existe por lo menos una distancia M al hiperplano.

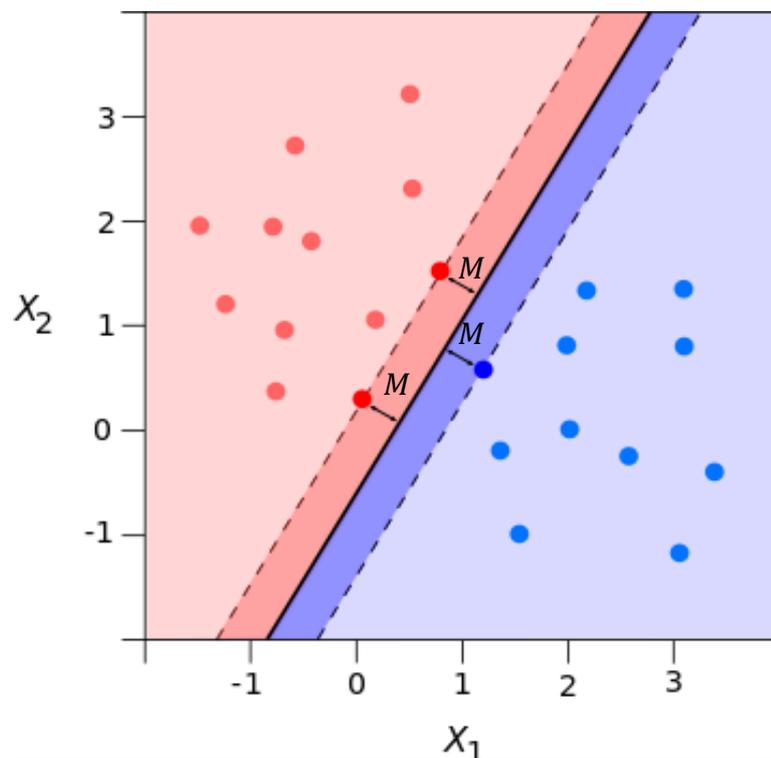
¿Cuál hiperplano seleccionar?

Queremos que $f(x)$ es igual a $+1$ o -1 en los puntos más cercanos a la frontera del margen.



De esta manera sabemos que si $yf(x) < 1$, entonces el punto está dentro del margen de separación.

Vectores de Soporte



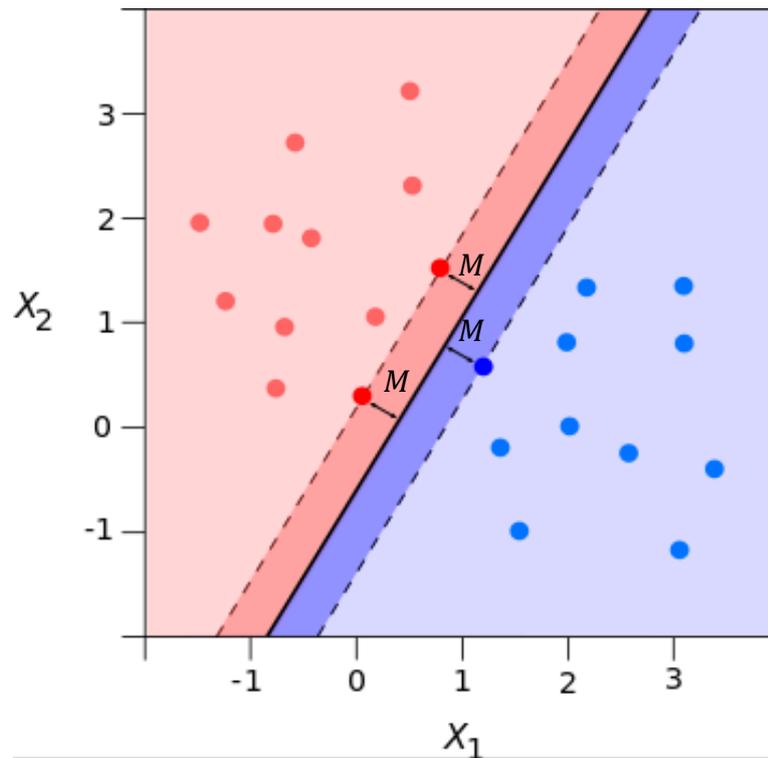
Vectores de Soporte

Son aquellos x_i involucrados en la determinación del margen, y por lo tanto determinan la solución del problema de clasificación.

Vemos que otros puntos no afectan en absoluto.

Vectores de Soporte

¿Cómo encontrar los parámetros $\beta_0, \beta_1, \dots, \beta_p$?



Problema de optimización;

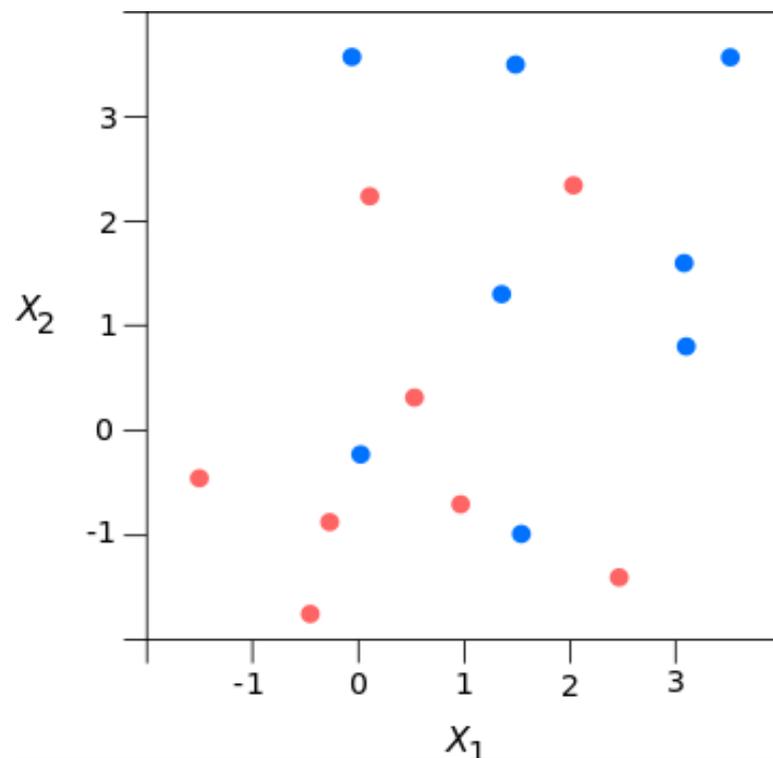
$$\max_{\beta_0, \beta_1, \dots, \beta_p} M$$

$$y_i(\beta_0 + \beta_1 X_1 + \dots + \beta_p X_p) \geq 1$$

para todo $i = 1, \dots, N$

Problemas de este enfoque

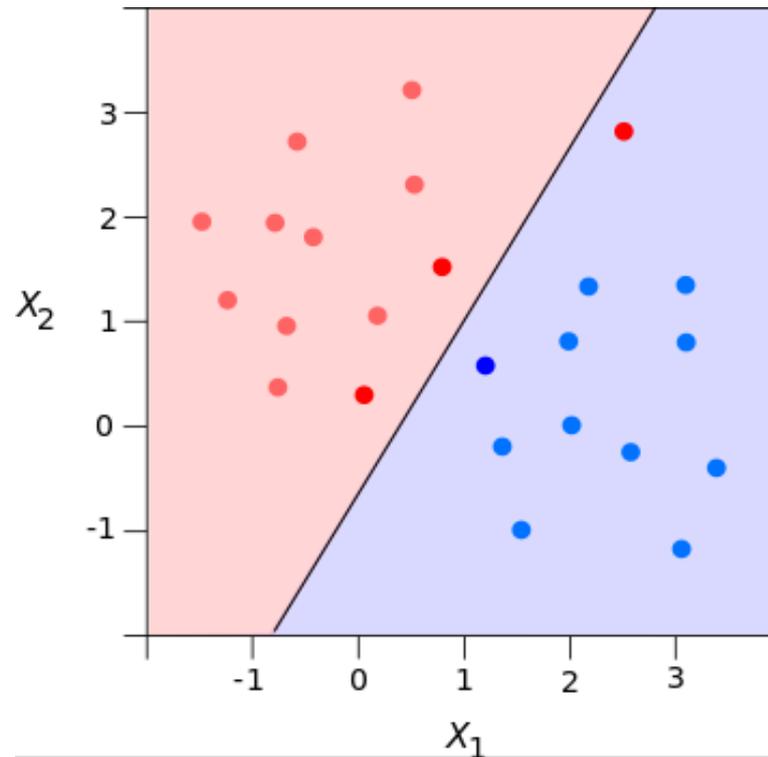
¿Qué pasa si los datos no son linealmente separables?



No puedo encontrar un hiperplano que cumpla las restricciones.

Problemas de este enfoque

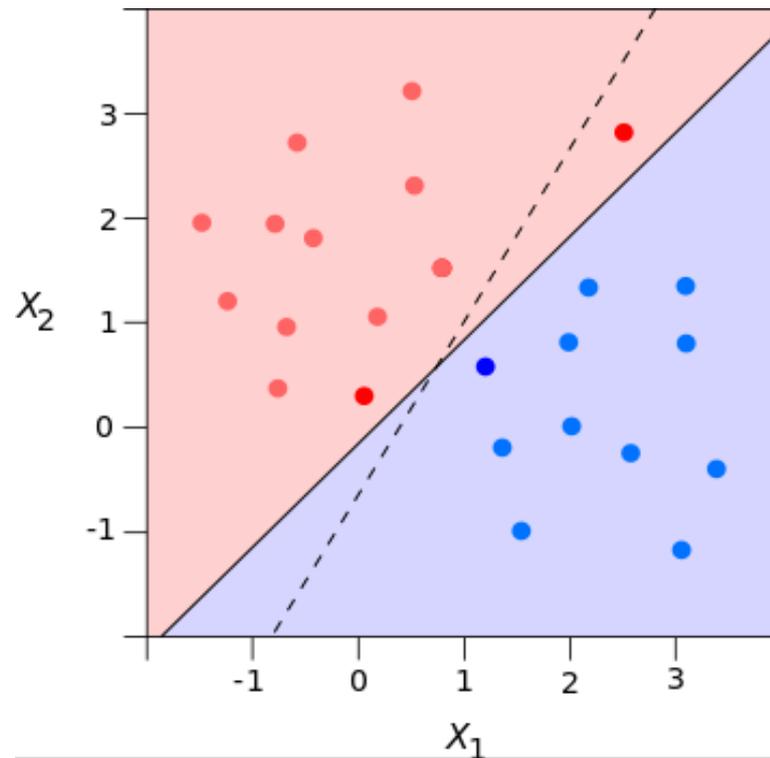
¿Qué pasa si los datos son ruidosos?



Si añadimos un nuevo dato cerca del margen.

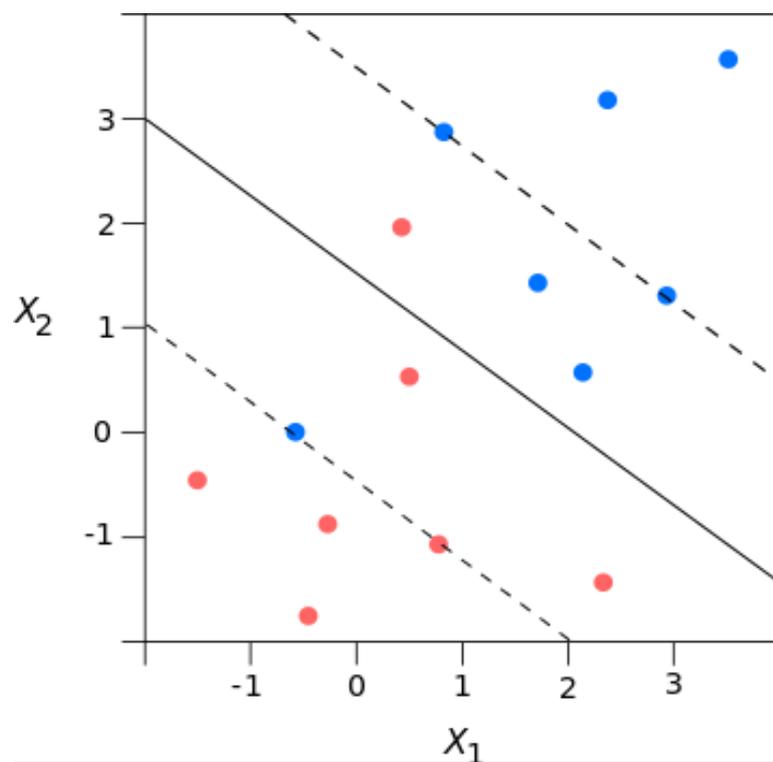
Problemas de este enfoque

¿Qué pasa si los datos son ruidosos?



Si añadimos un nuevo dato cerca del margen.

Solución: Margen Suave

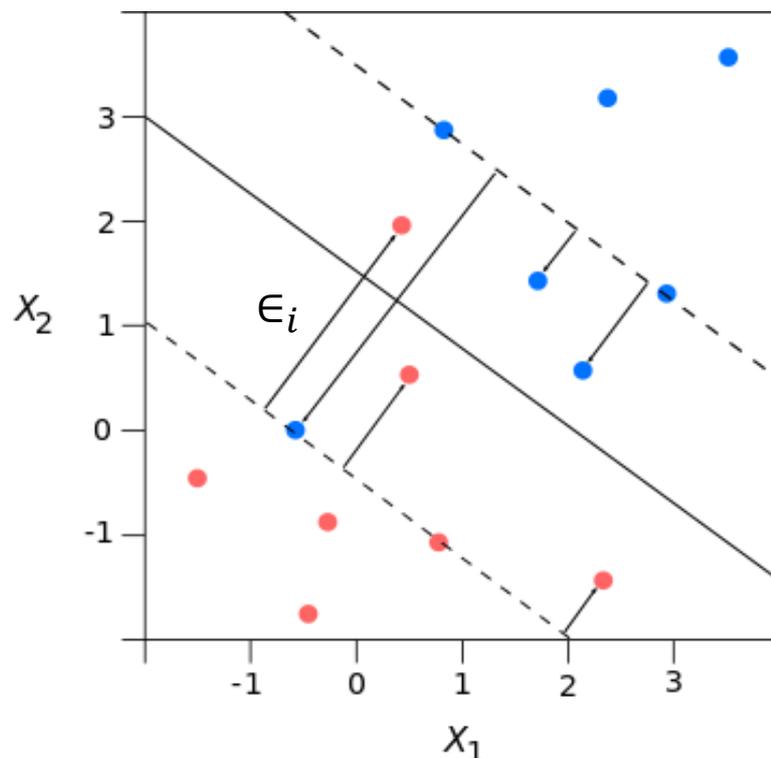


Aceptamos dos casos:

- Datos en el lado incorrecto del margen.
- Datos en el lado incorrecto de la frontera de decisión.

Solución: Margen Suave

¿Cómo cuantificar y regular la suavidad del margen?

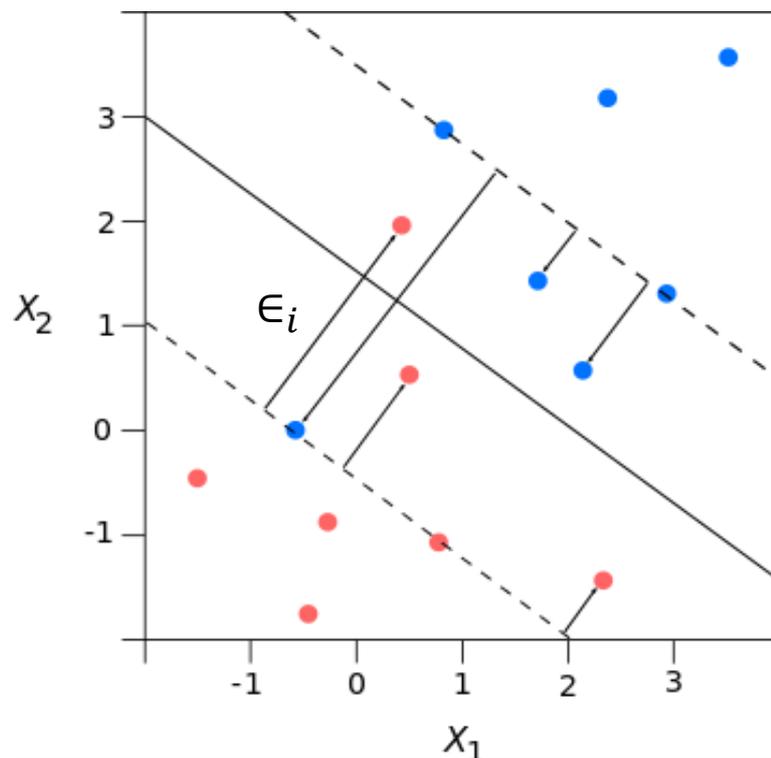


Introducimos dos nuevas restricciones:

- $\epsilon_i \geq 0$
- $\sum_i^N \epsilon_i \leq C$

Solución: Margen Suave

El problema de optimización se convierte en:



$$\max_{\beta_0, \beta_1, \dots, \beta_p} M$$

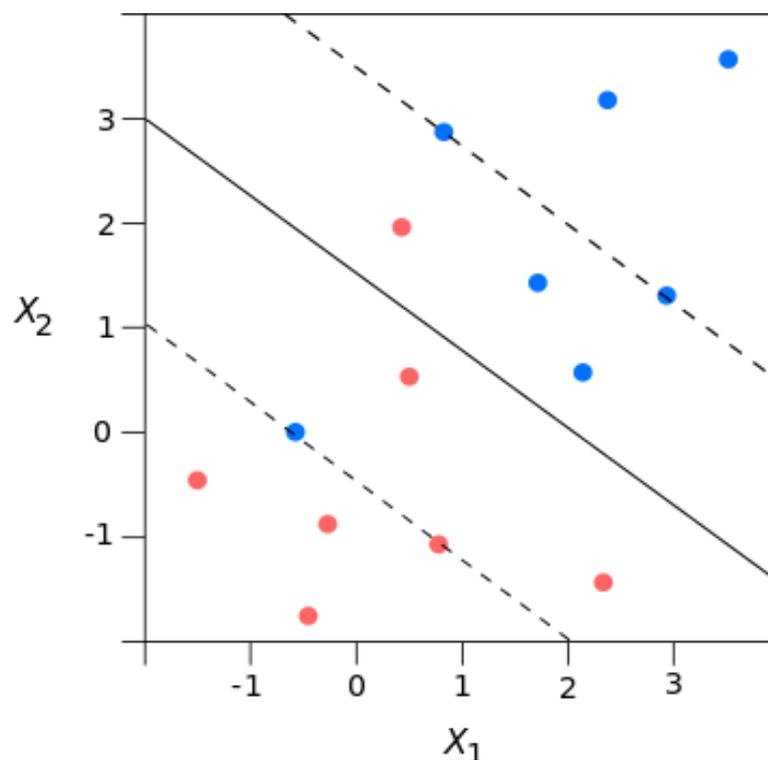
$$y_i(\beta_0 + \beta_1 X_1 + \dots + \beta_p X_p) \geq 1 - \epsilon_i$$

para todo $i = 1, \dots, N$,

$$\epsilon_i \geq 0, \quad \sum_i^N \epsilon_i \leq C$$

Solución: Margen Suave

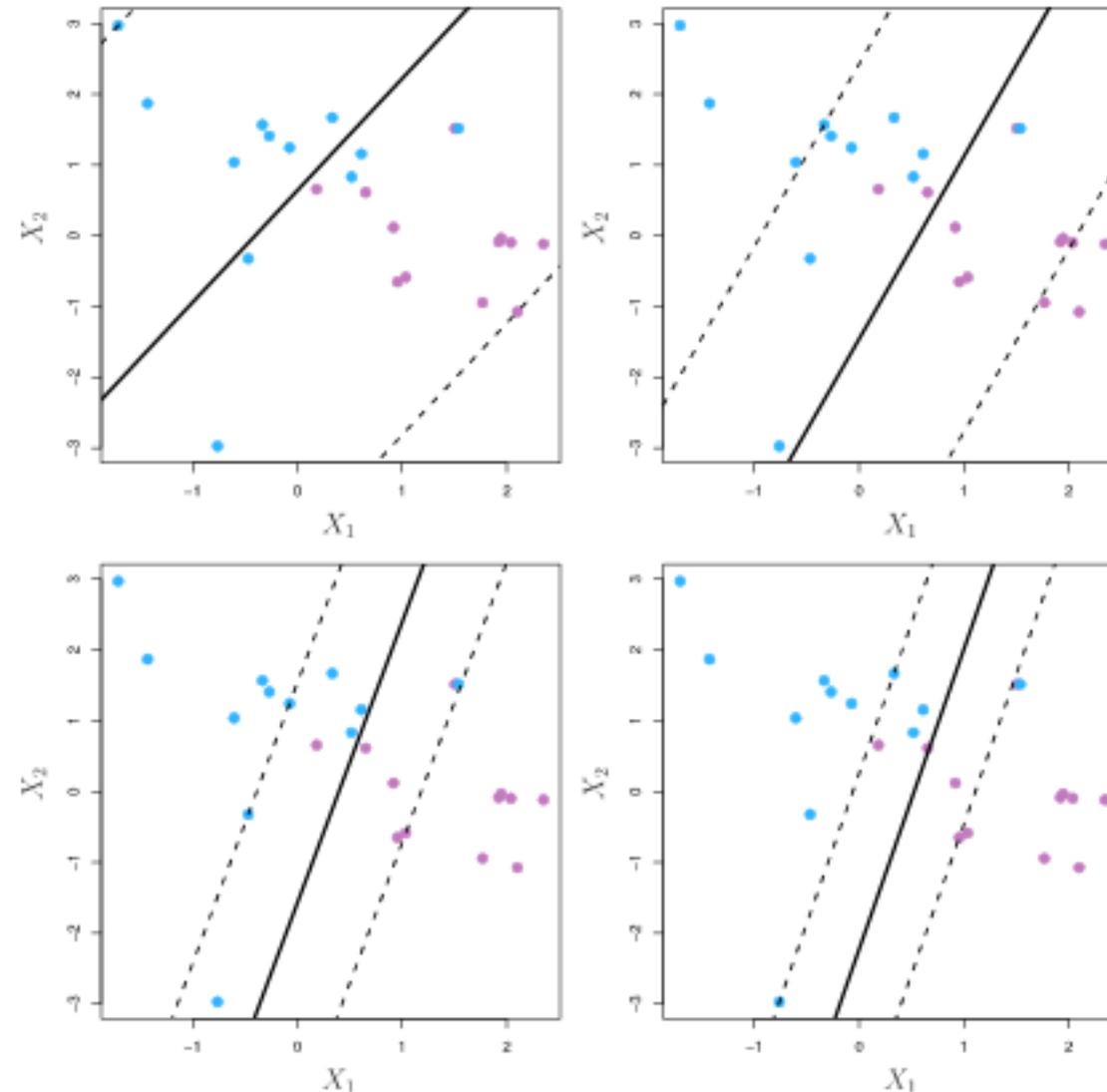
¿Qué pasa con los vectores de soporte?



Se tienen más vectores de soporte que están involucrados en la solución del problema de clasificación.

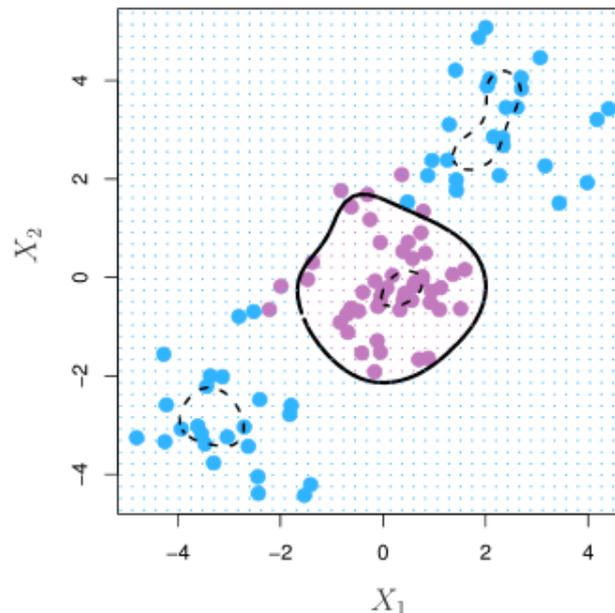
El número de vectores de soporte está determinado por el parámetro C .

Solución: Margen Suave



Observaciones sobre SVM

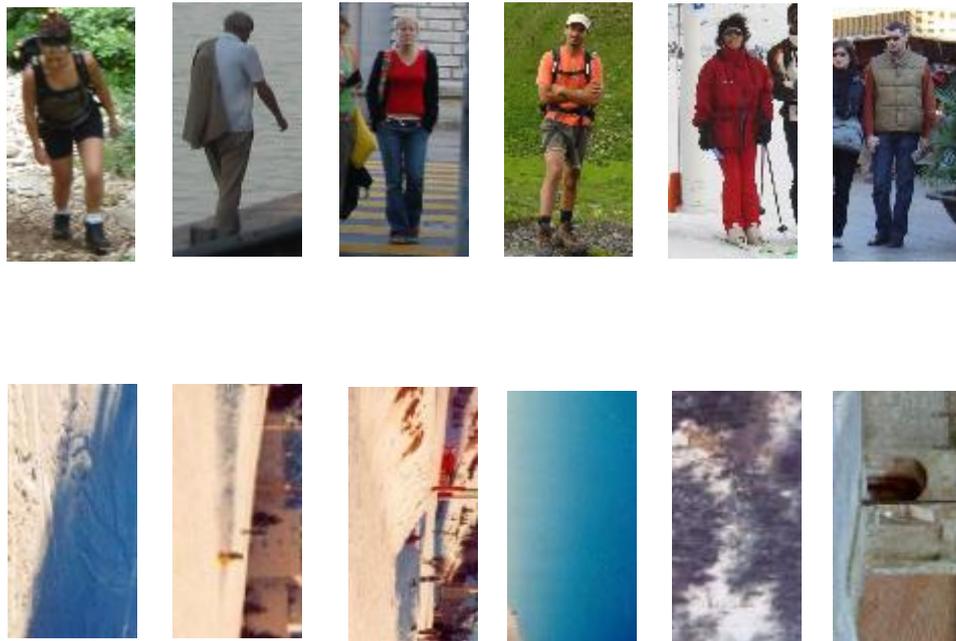
- No usa un modelo de probabilidad.
- Caja negra.
- Sólo depende de los vectores de soporte.
- Se puede extender a casos donde los datos son no linealmente separables.



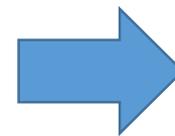
Procesamiento de las Imágenes

Regresando a nuestro diagrama de clasificación de imágenes:

Conjunto de Datos

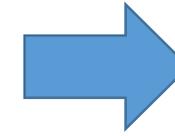


Procesamiento



Datos
 $\{(x_i, y_i)\}$

Entrenamiento

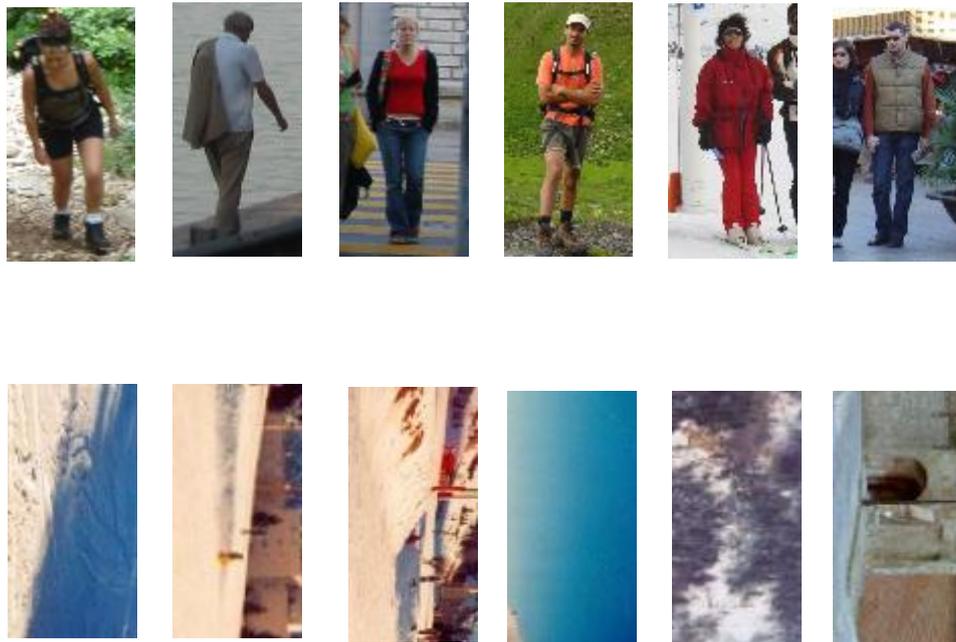


Clasificador
 $C(X)$

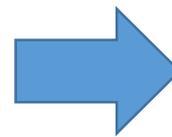
Procesamiento de las Imágenes

Regresando a nuestro diagrama de clasificación de imágenes:

Conjunto de Datos

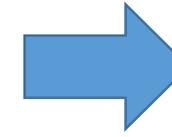


Procesamiento



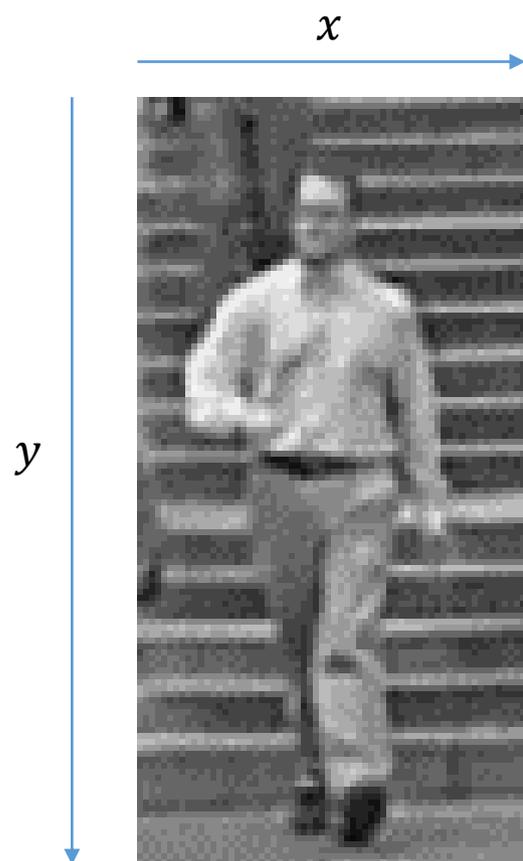
Datos
 $\{(x_i, y_i)\}$

Entrenamiento



Clasificador
 $C(X)$

Procesamiento de las Imágenes



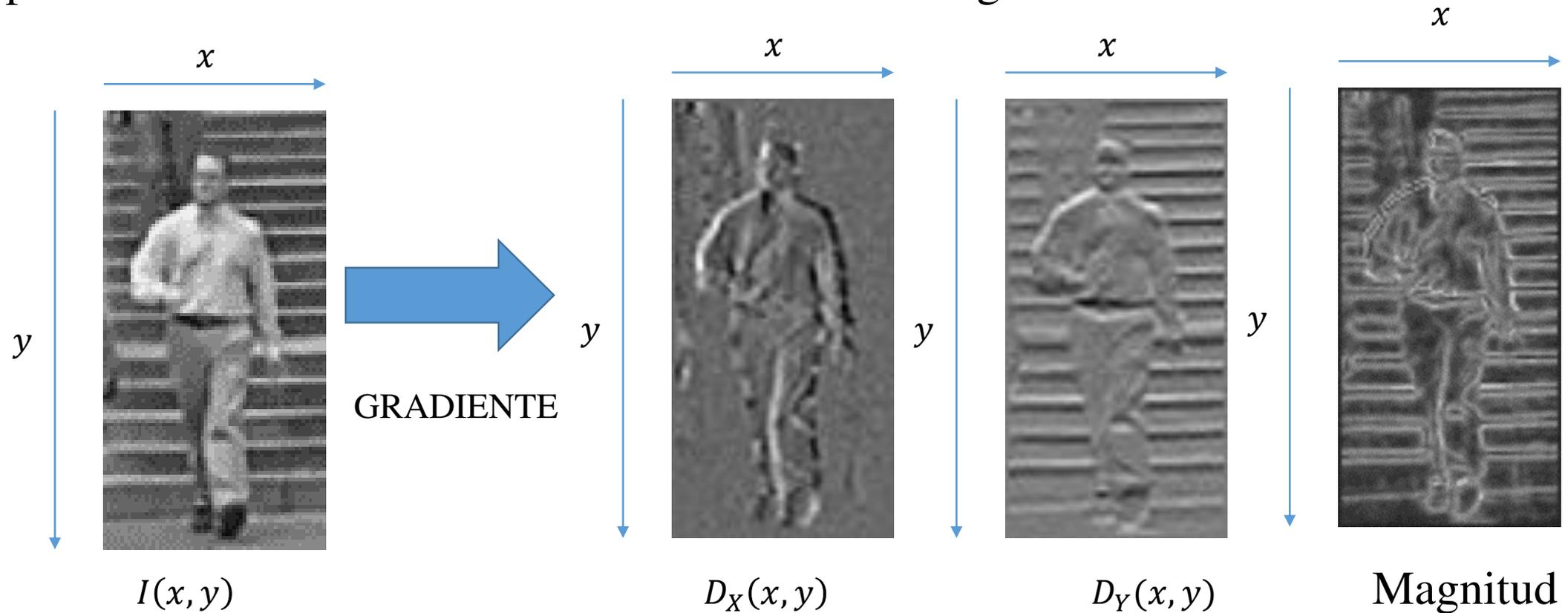
Una imagen puede ser vista como una función:

$$z = I(x, y) ,$$

donde (x, y) representa la posición de un pixel
y z representa la intensidad luminosa (escala de grises
de 0 a 255) para el pixel (x, y) .

Procesamiento de las Imágenes

En este sentido, a una imagen se le puede aplicar cualquier transformación que se aplicaría a una función. El resultado son nuevas imágenes:



¿Cómo Procesar Imágenes para Clasificación?

Objetivo:



$(X, Y = ?)$



$(X, Y = 1)$

¿Cómo Procesar Imágenes para Clasificación?

Objetivo:



$(X, Y = ?)$

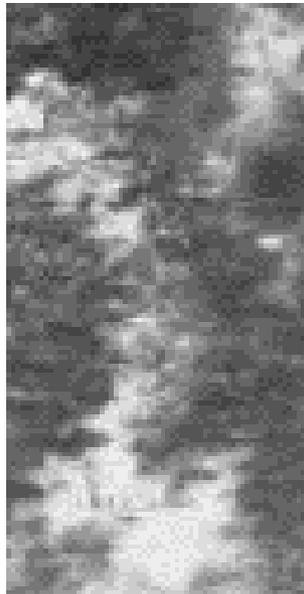


$(X, Y = -1)$

¿Cómo Procesar Imágenes para Clasificación?

Objetivo:

¿QUIEN ES X?



$(X, Y = ?)$

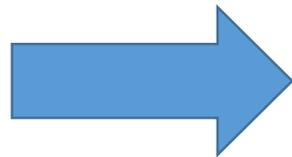


$(X, Y = -1)$

Histograma de Gradientes Orientados

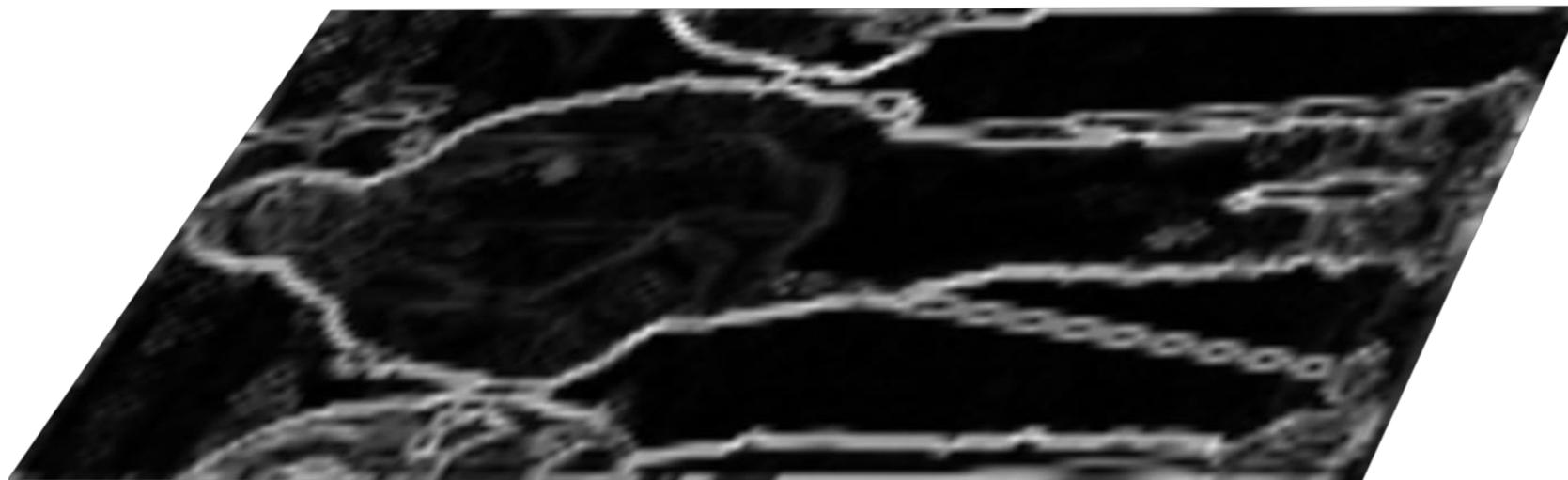
Dalal & Triggs (2001)

La información más discriminante está en el contorno/silueta de una persona:



Histograma de Gradientes Orientados

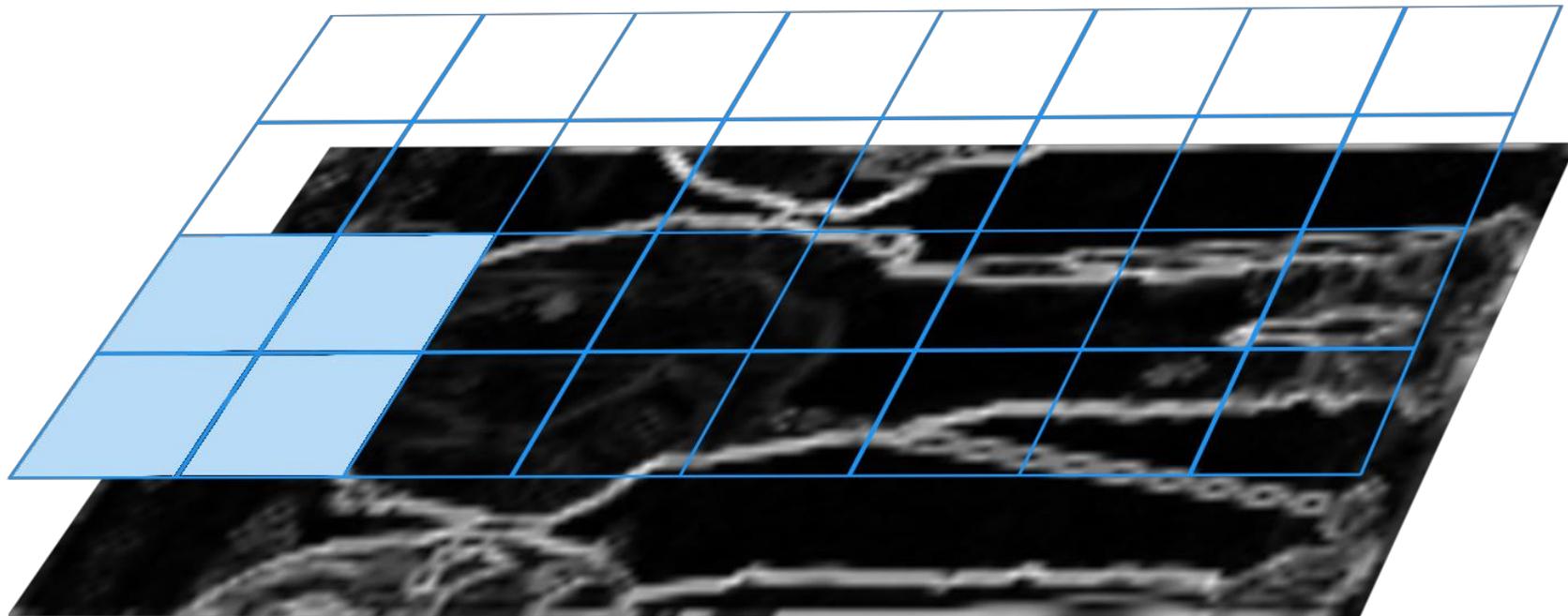
Dalal & Triggs (2001)



Histograma de Gradientes Orientados

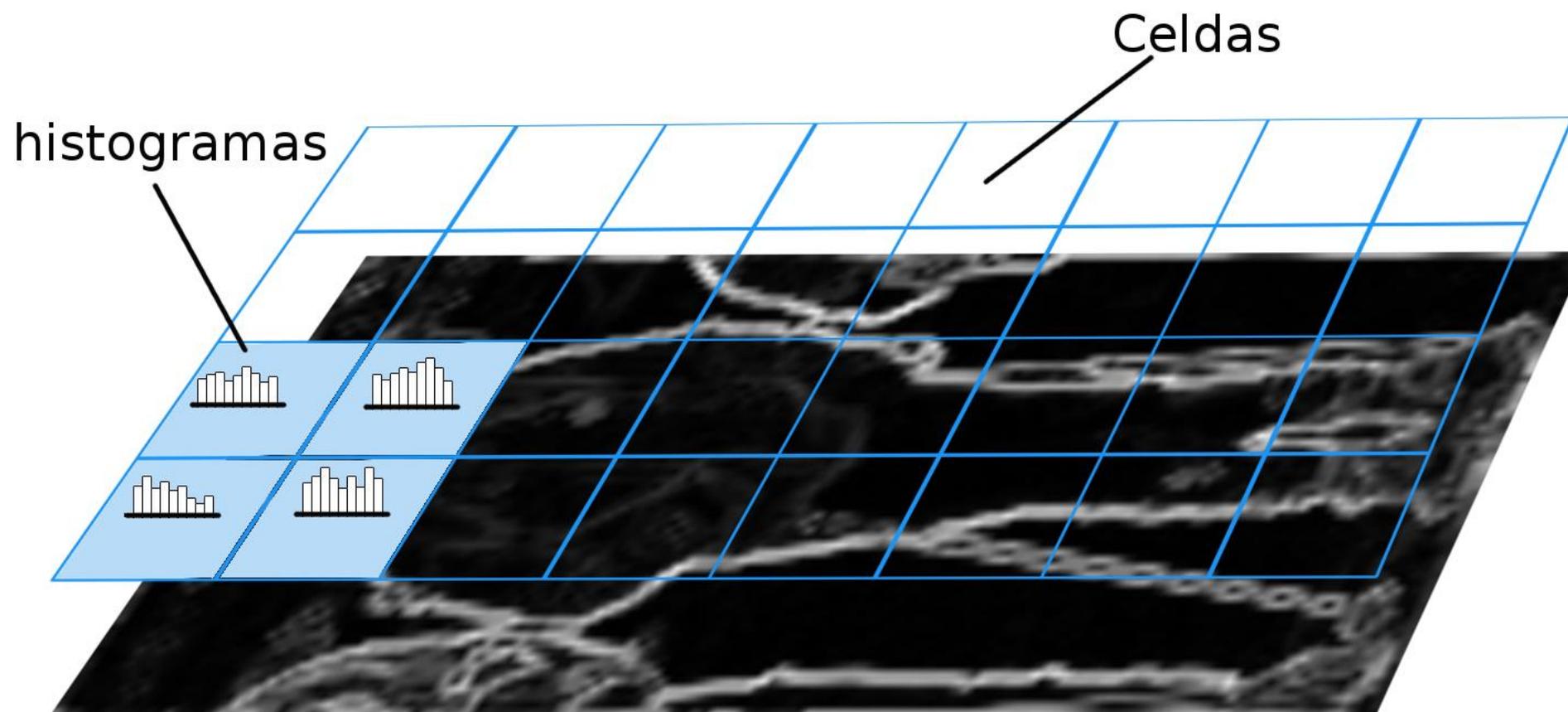
Dalal & Triggs (2001)

Celdas



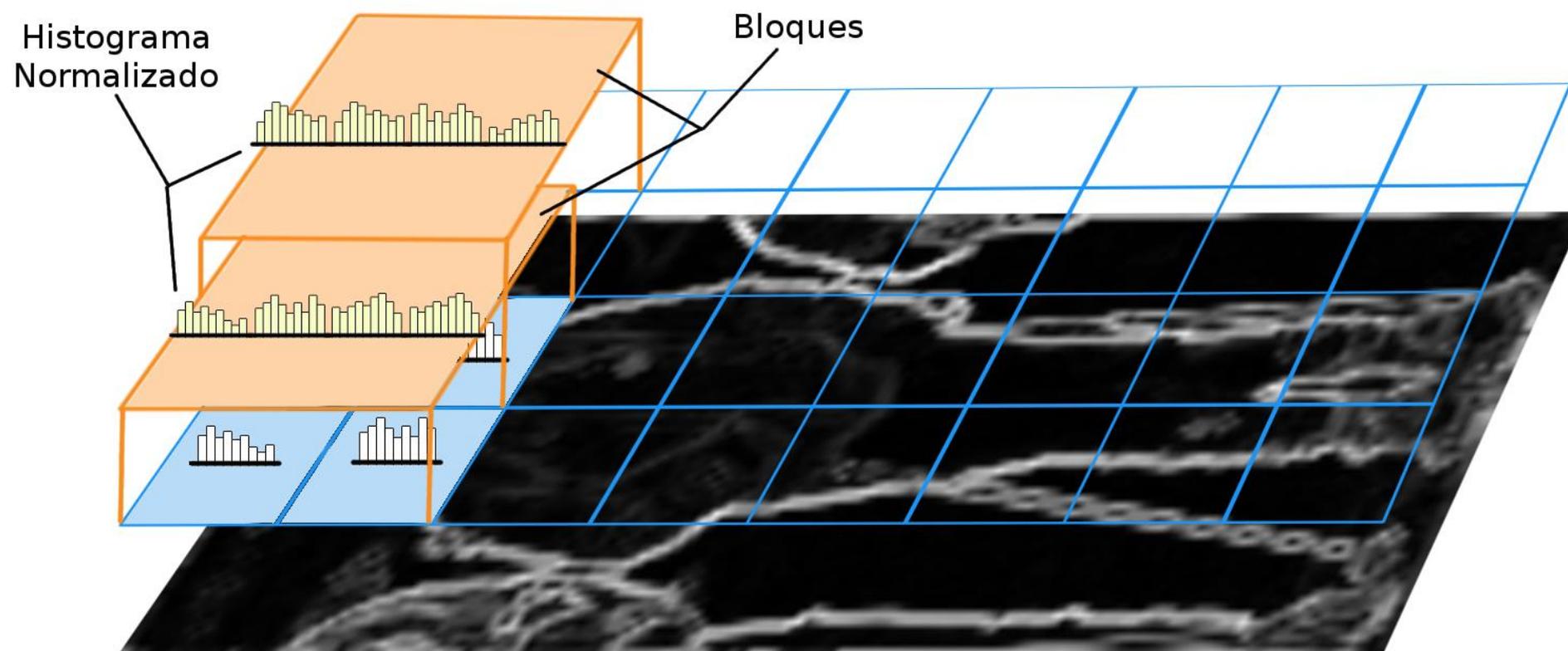
Histograma de Gradientes Orientados

Dalal & Triggs (2001)



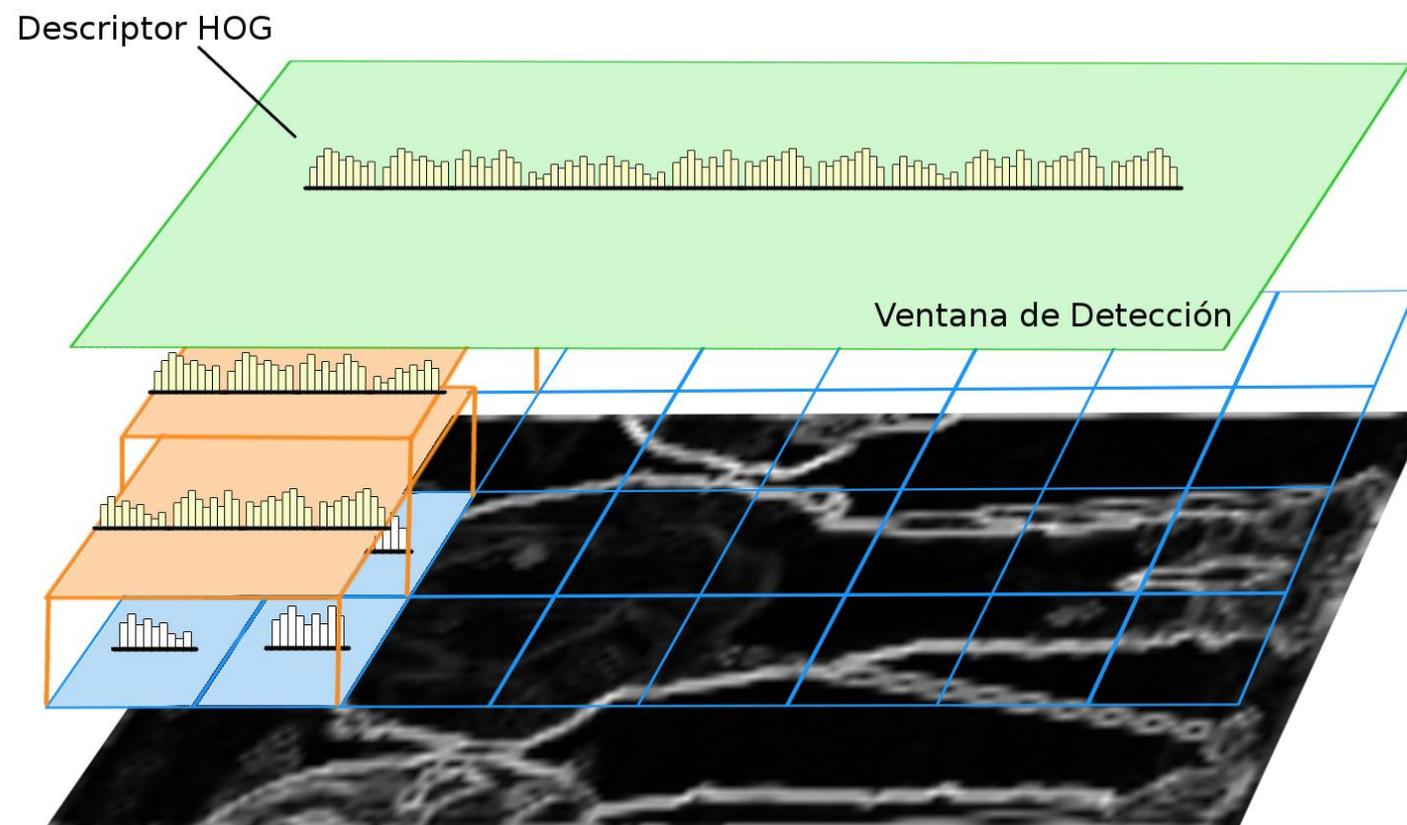
Histograma de Gradientes Orientados

Dalal & Triggs (2001)



Histograma de Gradientes Orientados

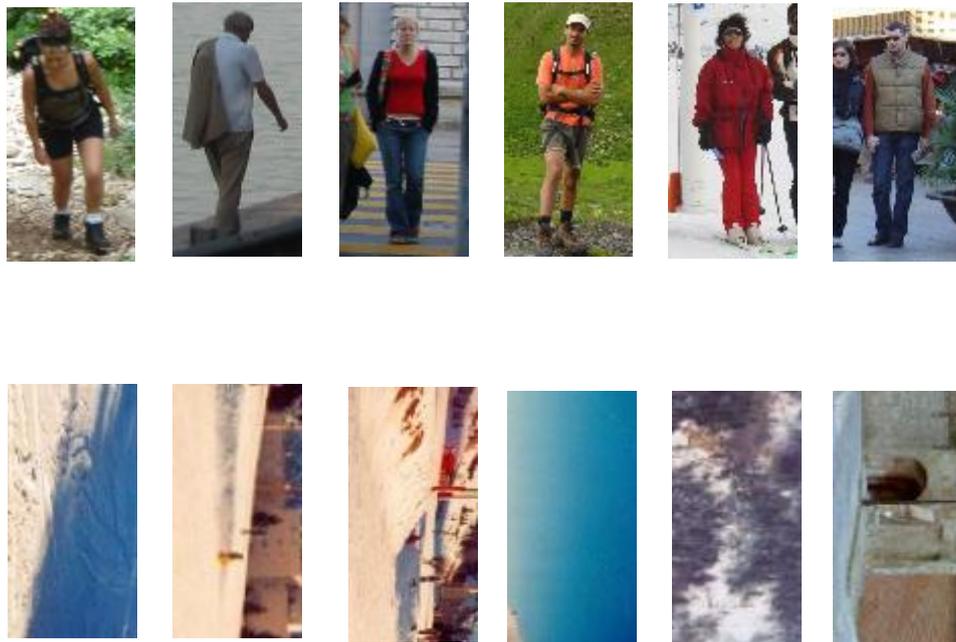
Dalal & Triggs (2001)



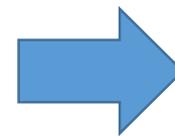
Clasificación de Imágenes

Ya tenemos definido nuestro vector característico X :

Conjunto de Datos

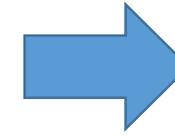


Procesamiento



Datos
 $\{(x_i, y_i)\}$

Entrenamiento



Clasificador
 $C(X)$

Clasificación de Imágenes

Clasificación de un nuevo dato:



$$f(x) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_{3780} X_{3780}$$

persona si $f(x) \geq 0$

no persona si $f(x) < 0$

Aplicación 1 – Detección de Peatones



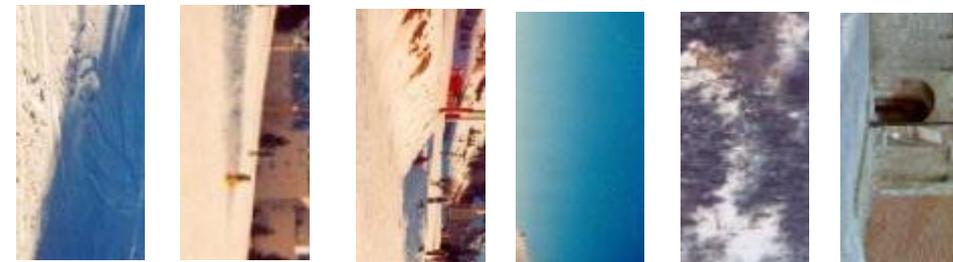
Entrenamiento

Utilizamos un conjunto de imágenes normalizadas que contienen personas (datos positivos) y que no contienen personas (datos negativos).

Conjunto de Datos



1500



10000

Detección

La metodología presentada sólo es capaz de detectar personas en imágenes normalizadas (68x124 píxeles y el peatón centrado en la imagen). ¿Qué pasa con las imágenes objetivo?

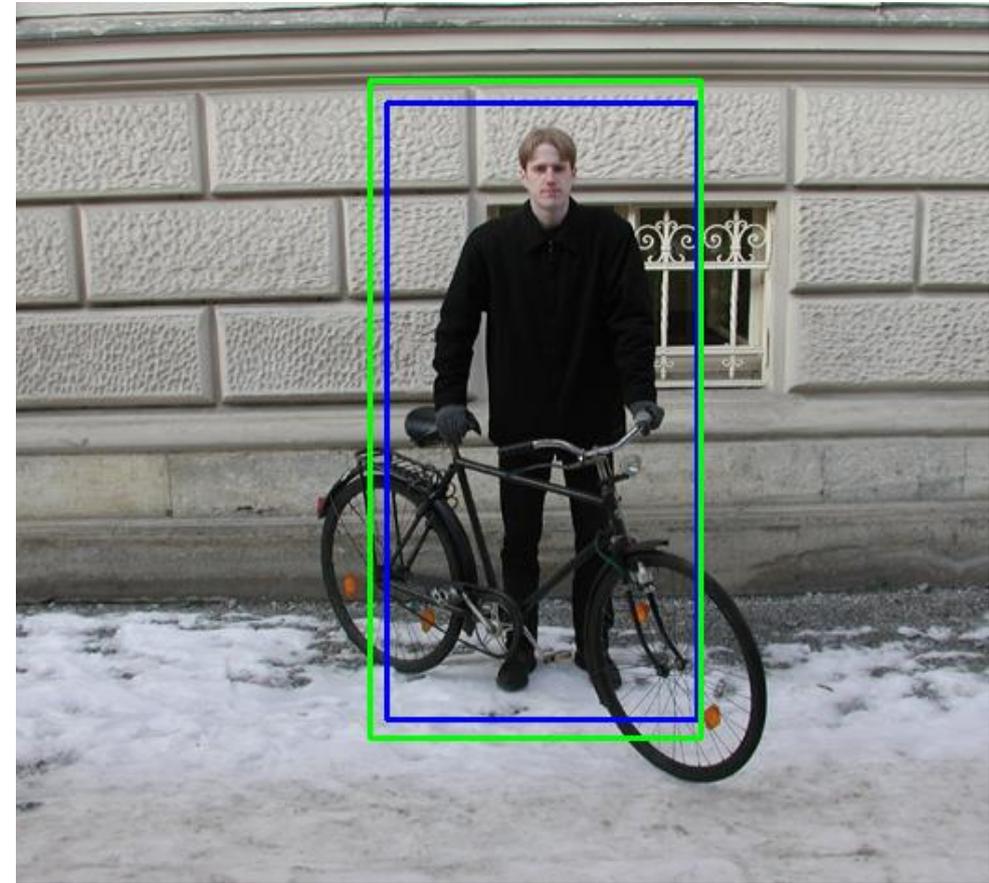
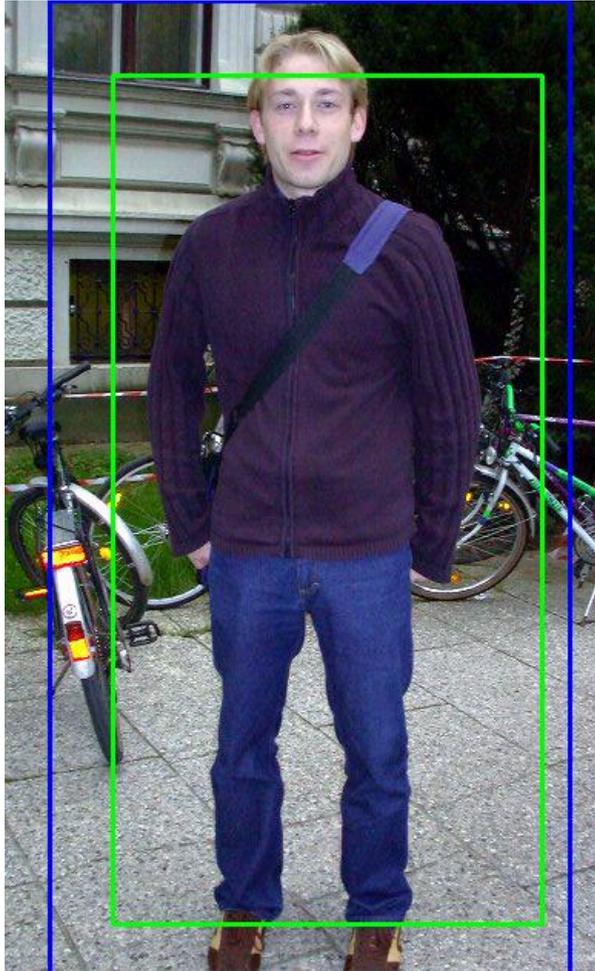


Detección

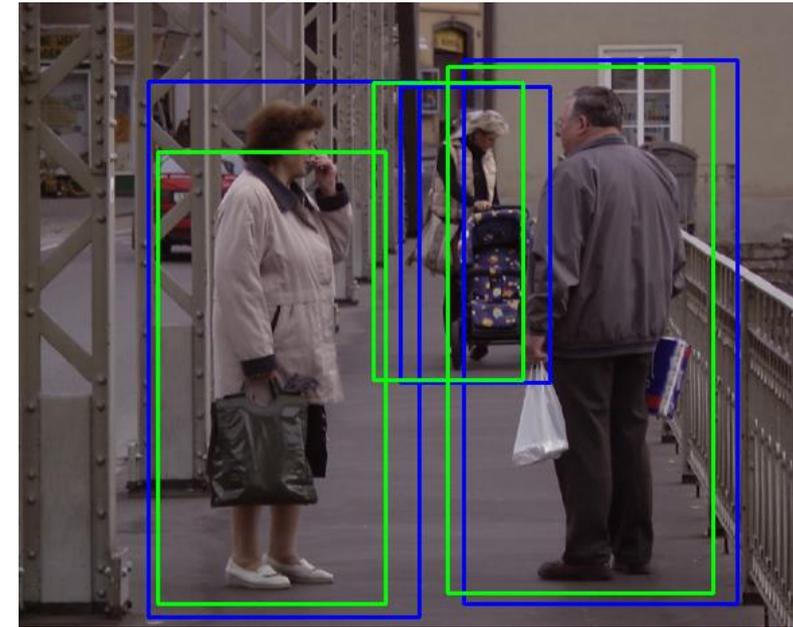
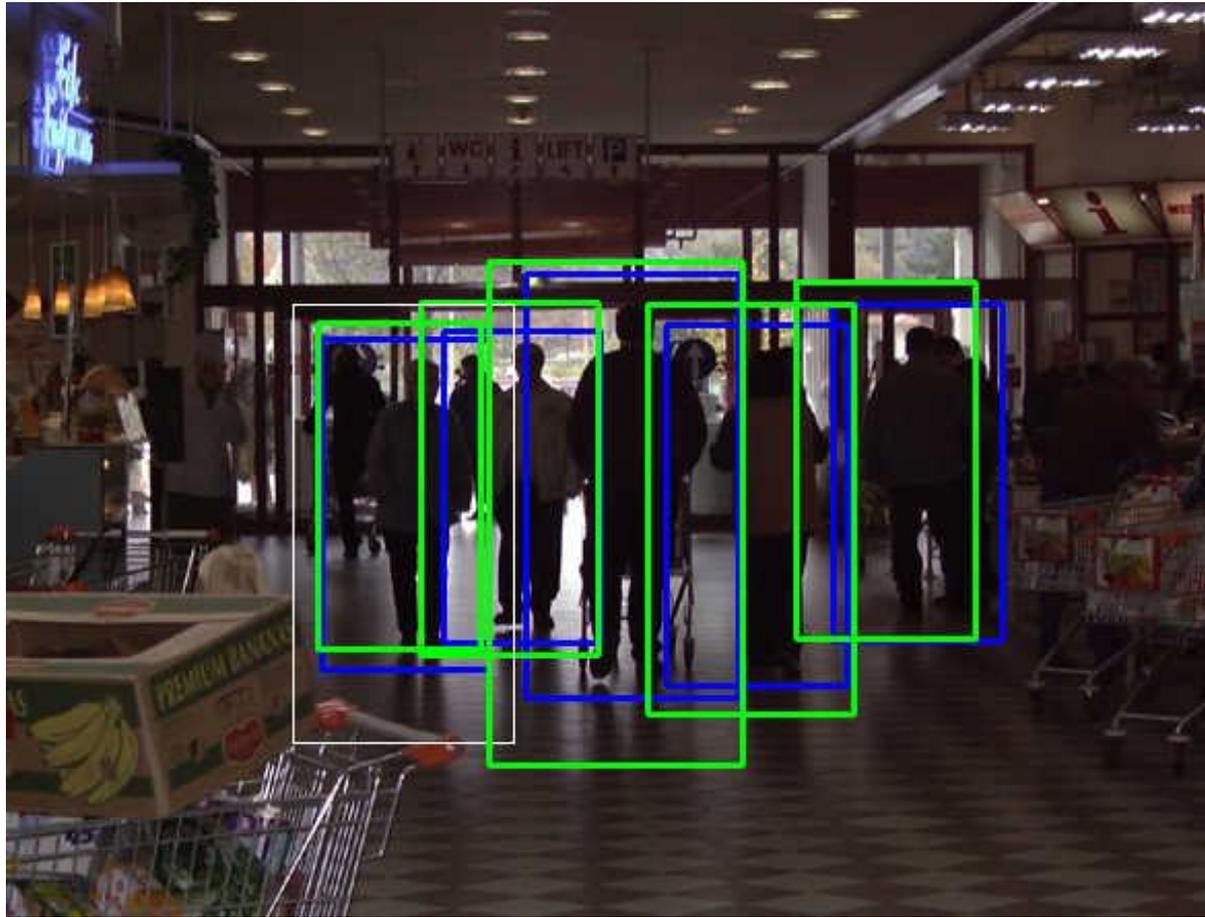
ESTRATEGIA: Búsqueda en posición y escala.



Resultados



Resultados



Resultados

People detection in RGB-D Data

Luciano Spinello, Kai O. Arras

Social Robotics Laboratory, University of Freiburg, Germany

Setup: three vertically mounted Kinect rgb-d cameras
in a busy university lobby



Social Robotics Laboratory



¡Muchas Gracias por su Atención!

