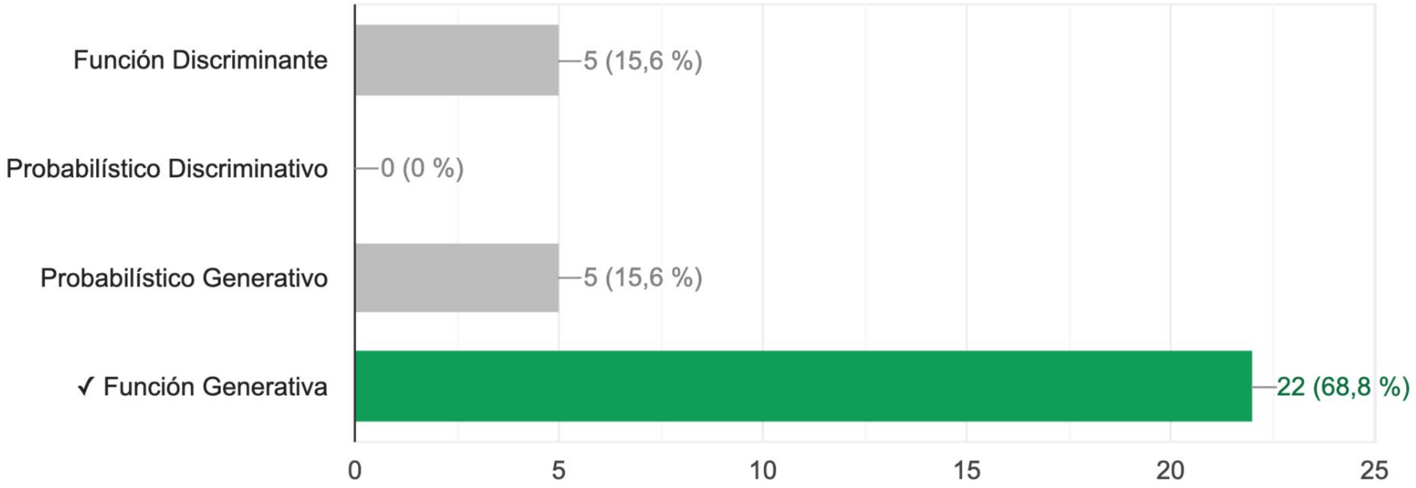
Form Clasificación

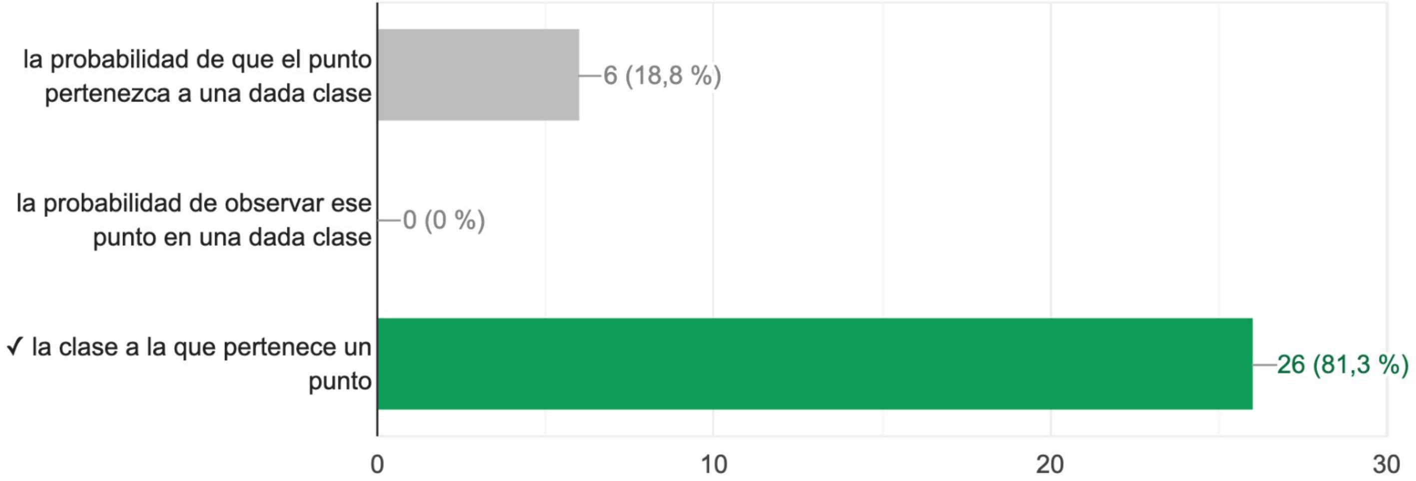
¿Cual de los siguientes no es un tipo de algoritmo de clasificación?

22 de 32 respuestas correctas



En un método de función discriminante, lo que se busca predecir es...

26 de 32 respuestas correctas



IAA-2023c1
Clase 4: Clasificación y Regresión Polinómica

UNSAM
UNIVERSIDAD
NACIONAL DE
SAN MARTÍN

Repaso: Modelos Lineales

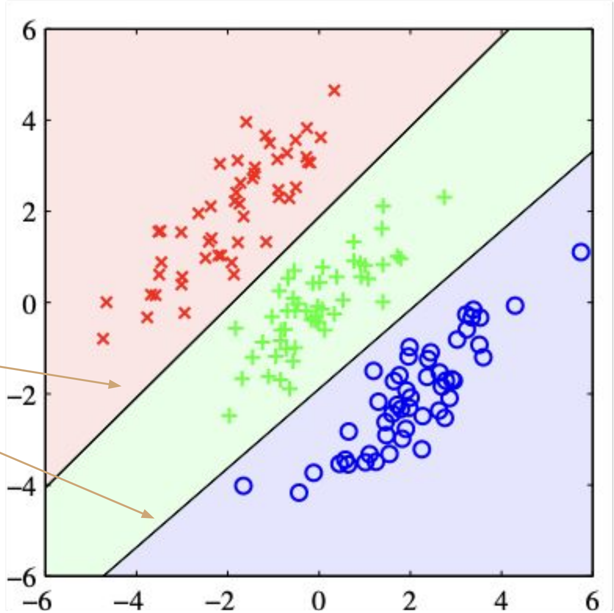
	Regresión Lineal	**Regresión Logística**
Modelo:	$y = \vec{w} \cdot \vec{x} + w_0$	$y = \sigma(\vec{w} \cdot \vec{x} + w_0)$
Función de Pérdida:	$L(y, t) = (y - t)^2$	$L(y; t) = -[t \log(y)(1 - t) \log(1 - y)]$
Target:	$y,\ t \in \mathbb{R}$	$t = k$ $\quad t = (0, \ldots, 0, 1, 0, \ldots, 0)$
Optimizador	Solución algebraica Descenso por Gradiente	Descenso por Gradiente

Repaso: Modelos Lineales Generalizados

El output es constante cuando $\vec{w} \cdot \vec{x} + w_0 = cte$

En particular la zona donde cambio mi predicción de clase también sigue la ecuación de una recta:

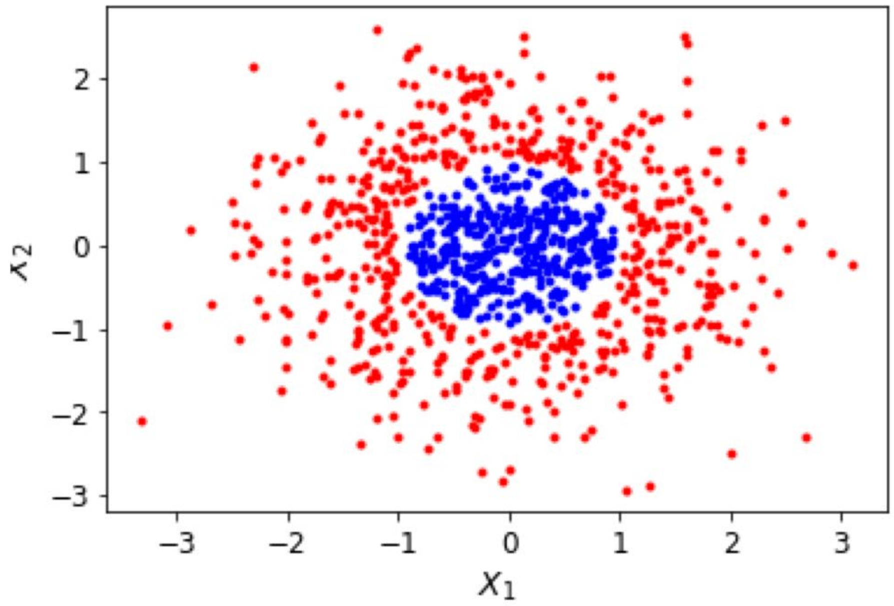
Frontera de Decisión Lineal

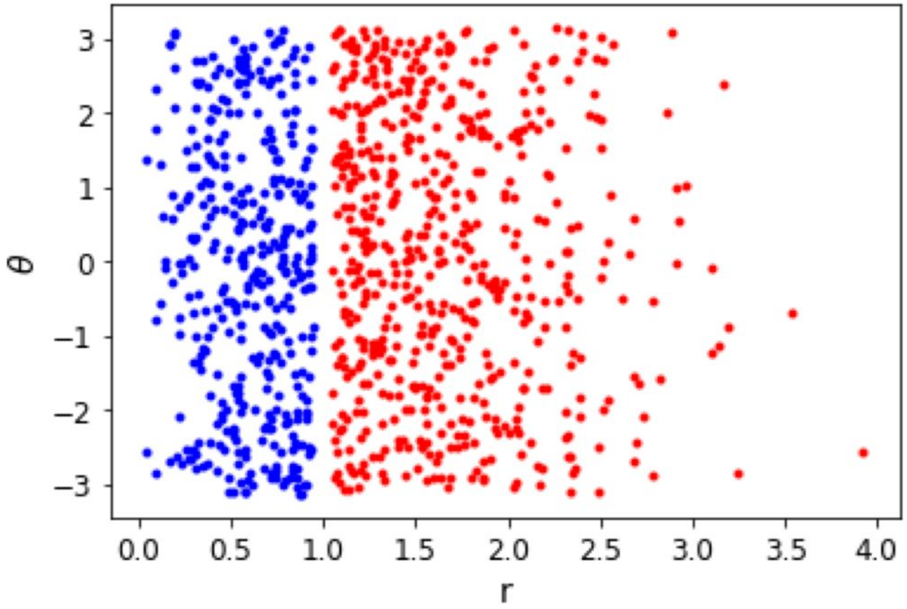


Repaso: Separabilidad Lineal

Podemos separar las clases trazando una linea recta? (i.e. con un modelo lineal)

Depende del las variables/coordenadas que usemos





Coordenadas polares

Modelo Polinomial Univariado

“activación”

Lineal

$$y = z$$

$$y = \sigma(z)$$

$$z = w_0 + w_1 x$$

Modelo Polinomial Univariado

“activación”

Lineal

$$y = z$$
$$y = \sigma(z)$$

$$z = w_0 + w_1 x$$

Cuadrático

$$z = w_0 + w_1 x + w_2 x^2$$

Modelo Polinomial Univariado

“activación”

Lineal

$$y = z$$

$$y = \sigma(z)$$

$$z = w_0 + w_1 x$$

Cuadrático

$$z = w_0 + w_1 x + w_2 x^2$$

Polinómico grado M

$$z = w_0 + w_1 x + w_2 x^2 + \cdots + w_M x^M$$

Modelo Polinomial Univariado

"activación"

Lineal

$$y = z$$
$$y = \sigma(z)$$

$$z = w_0 + w_1 x$$

Cuadrático

$$z = w_0 + w_1 x + w_2 x^2$$

Polinómico grado M

$$z = w_0 + w_1 x + w_2 x^2 + \cdots + w_M x^M$$

Es como un modelo lineal pero cambiando el input a $x \rightarrow \vec{x} = (x, x^2, \ldots, x^M)$

"Preprocesado Polinómico"

Modelo Polinomial Univariado

En términos de la Matriz de diseño X: $\vec{z} = X \cdot \vec{w}$

$$x \to \vec{x} = (x, x^2, \dots, x^M)$$

$$X = \begin{pmatrix} 1 & x^{(1)} \\ 1 & x^{(2)} \\ \vdots & \vdots \\ 1 & x^{(N)} \end{pmatrix} \longrightarrow X = \begin{pmatrix} 1 & x^{(1)} & (x^{(1)})^2 & \dots & (x^{(1)})^M \\ 1 & x^{(2)} & (x^{(2)})^2 & \dots & (x^{(2)})^M \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x^{(N)} & (x^{(N)})^2 & \dots & (x^{(N)})^M \end{pmatrix}$$

Modelo Polinomial Multivariado

$$x \to \vec{x} = (x, x^2, \ldots, x^M)$$

Y si tengo muchos features?

Modelo Polinomial Multivariado

$$x \to \vec{x} = (x, x^2, \ldots, x^M)$$

Y si tengo muchos features?

$$\vec{x} = (x_1, x_2) \to \vec{x} = (x_1, x_2, x_1^2, x_1 x_2, x_2^2, \ldots, x_1^M, x_1^{M-1} x_2, \ldots, x_1 x_2^{M-1}, x_2^M)$$

Ejemplo grado 3:

$$\begin{aligned} z = \quad & w_0 + w_{1,0} x_1 + w_{0,1} x_2 + w_{2,0} x_1^2 + w_{1,1} x_1 x_2 + w_{0,2} x_2^2 \\ & + w_{3,0} x_1^3 + w_{2,1} x_1^2 x_2 + w_{1,2} x_1 x_2^2 + w_{0,3} x_2^3 \end{aligned}$$

Modelo Polinomial Multivariado

$$x \rightarrow \vec{x} = (x, x^2, \ldots, x^M)$$

Y si tengo muchos features?

$$\vec{x} = (x_1, x_2) \rightarrow \vec{x} = (x_1, x_2, x_1^2, x_1x_2, x_2^2, \ldots, x_1^M, x_1^{M-1}x_2, \ldots, x_1x_2^{M-1}, x_2^M)$$

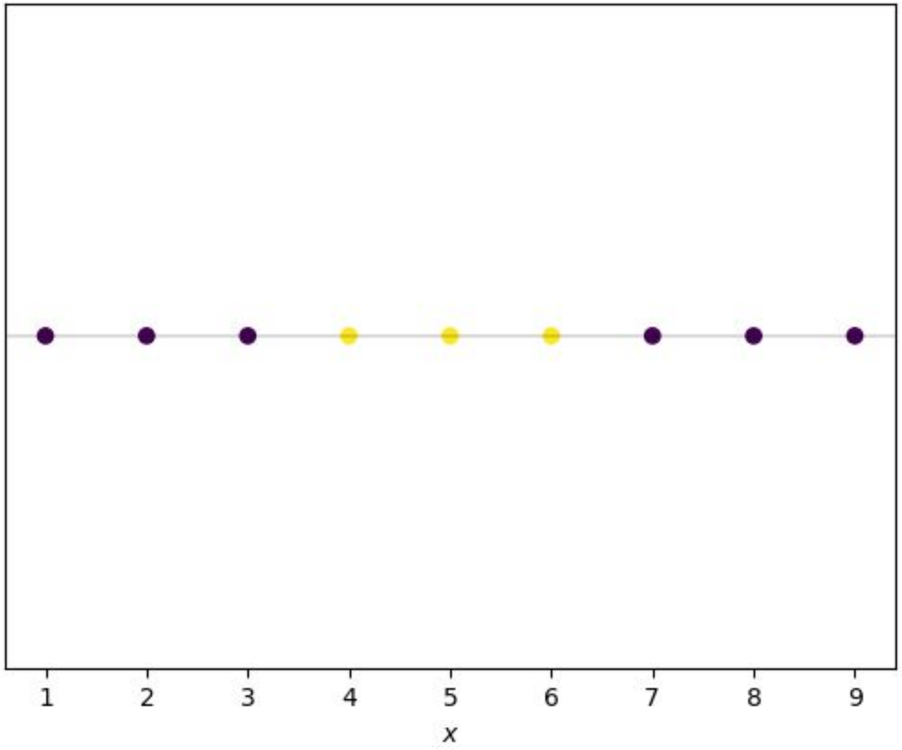
Ejemplo grado 3:

$$\begin{aligned} z = \quad & w_0 + w_{1,0}x_1 + w_{0,1}x_2 + w_{2,0}x_1^2 + w_{1,1}x_1x_2 + w_{0,2}x_2^2 \\ & + w_{3,0}x_1^3 + w_{2,1}x_1^2x_2 + w_{1,2}x_1x_2^2 + w_{0,3}x_2^3 \end{aligned}$$

Pasamos de 2 features a 9... **rápido crecimiento de features con el grado M!**

Ejemplo: Clasificación Unidimensional

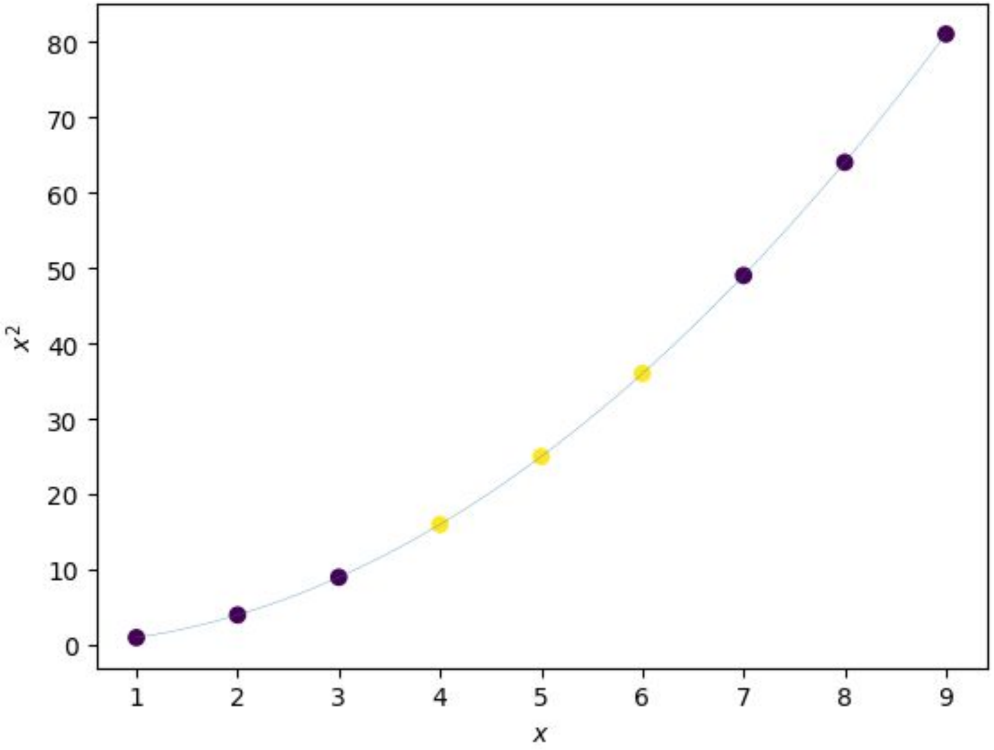
Es linealmente separable?



Ejemplo: Clasificación Unidimensional (cuadrático)

Es linealmente separable?

y ahora?

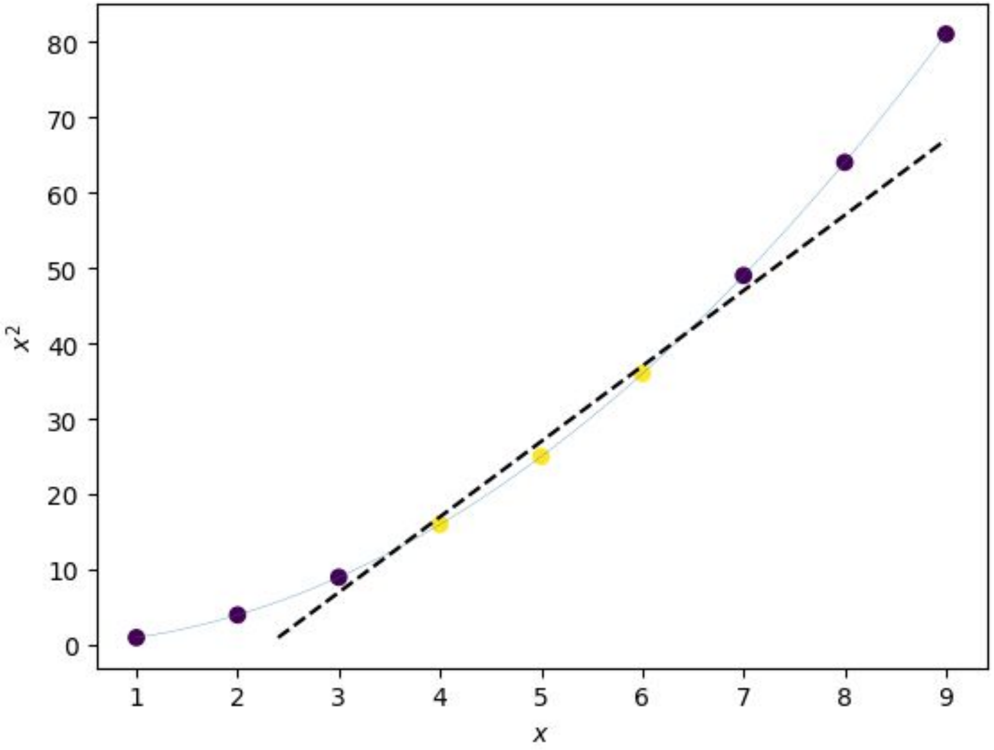


Ejemplo: Clasificación Unidimensional (cuadrático)

Es linealmente separable?

y ahora?

que significa esa frontera de decisión en términos del espacio original?

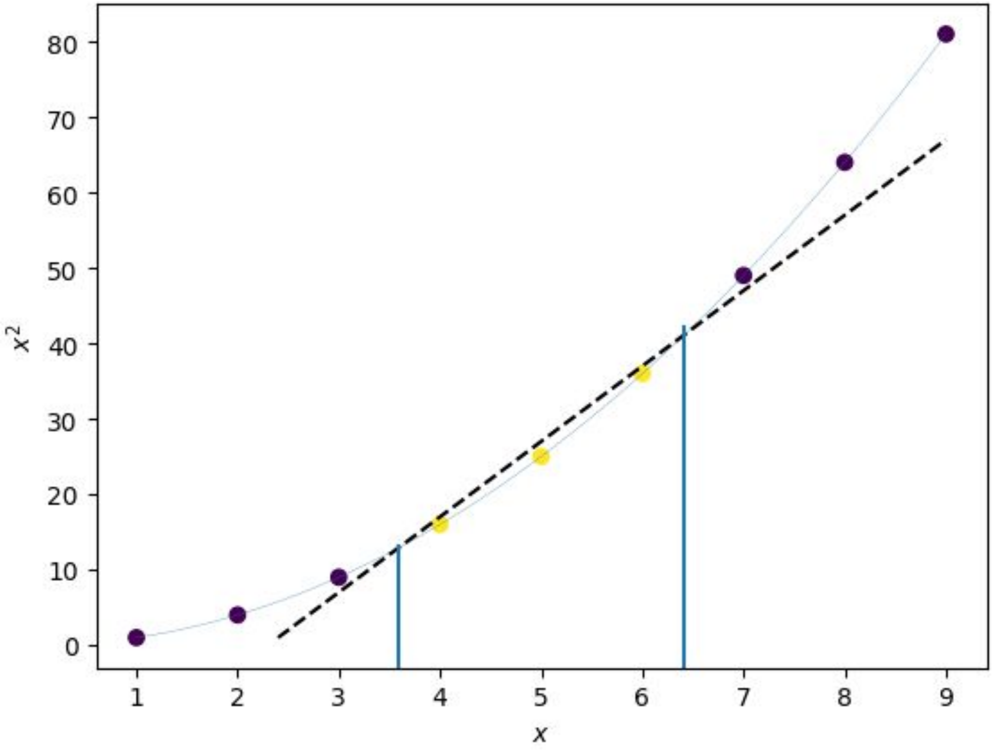


Ejemplo: Clasificación Unidimensional (cuadrático)

Es linealmente separable?

y ahora?

que significa esa frontera de decisión en términos del espacio original?



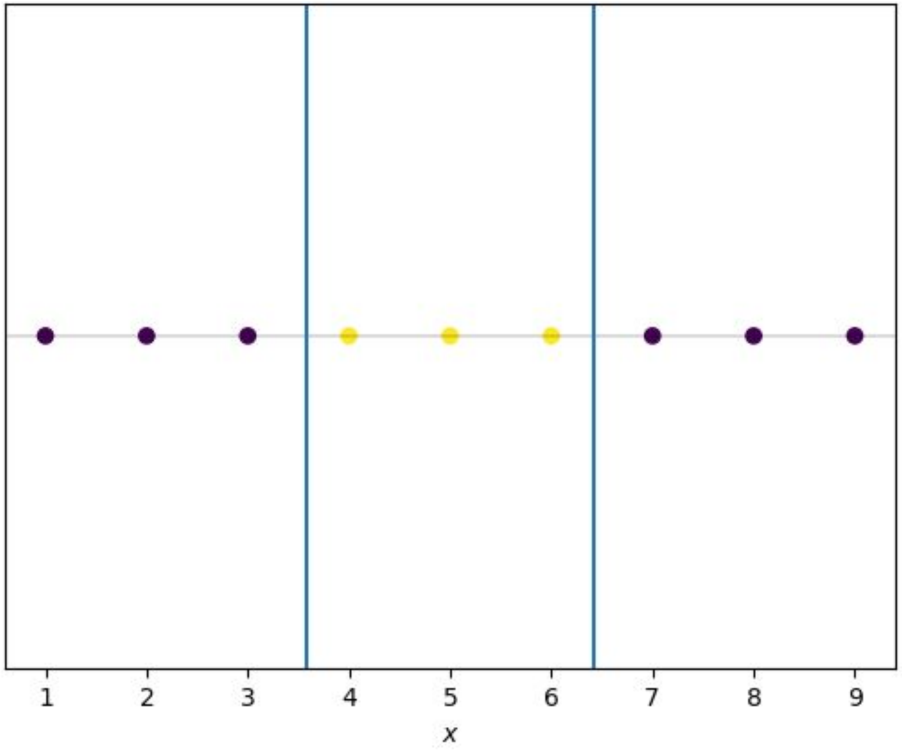
Ejemplo: Clasificación Unidimensional (cuadrático)

Es linealmente separable?

y ahora?

que significa esa frontera de decisión en términos del espacio original?

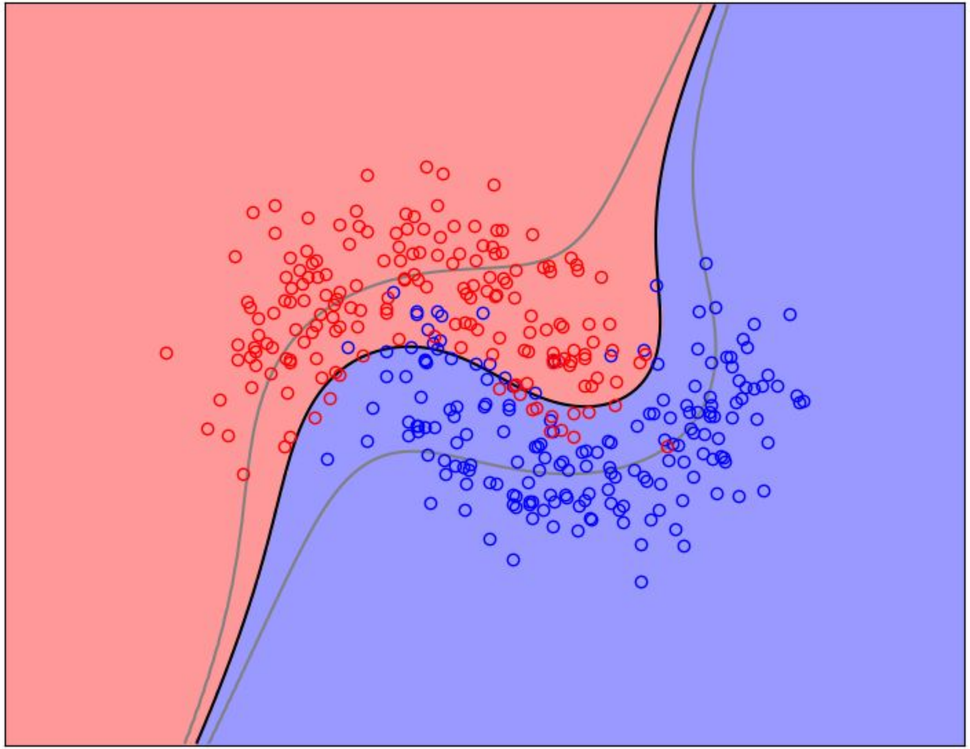
Wow... Esa frontera si que es *no lineal*



Ejemplo: Clasificación Bidimensional (cúbico)

En el caso genérico, conseguimos fronteras de decisión que ***proyectadas al espacio original*** se ven **no lineales**.

(Aunque en el espacio luego del procesado polinómico, sigue siendo un hiperplano **lineal**)



Ejemplo: Regresión Unidimensional

$$z = \sum_{i=0}^{M} w_i x^i$$

En regresión suele ser más intuitivo ver que ventaja tiene.

Figura del Bishop

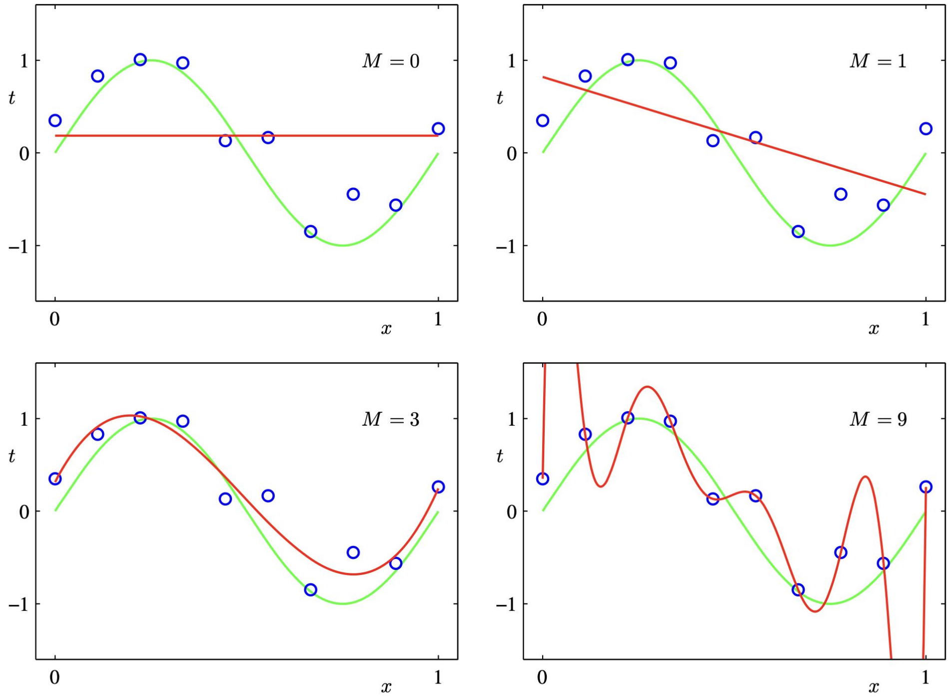


Figure 1.4 Plots of polynomials having various orders M, shown as red curves, fitted to the data set shown in Figure 1.2.

Ejemplo: Regresión Unidimensional

$$z = \sum_{i=0}^{M} w_i x^i$$

En regresión suele ser más intuitivo ver que ventaja tiene.

Pero ahora tengo que elegir el M **antes** de *entrenar.*

Hiperparámetro

Figura del Bishop

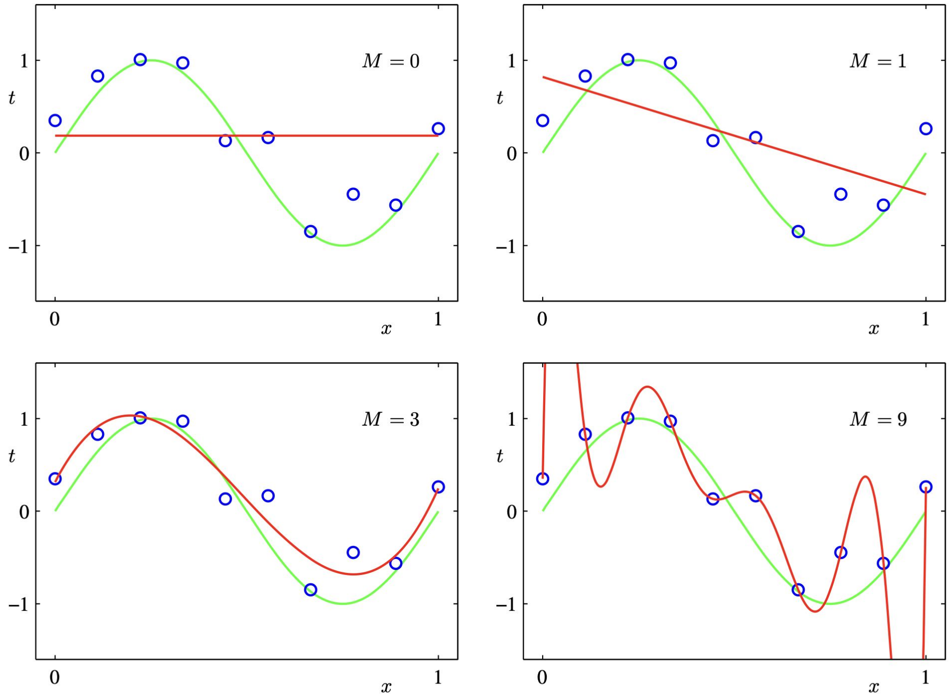


Figure 1.4 Plots of polynomials having various orders M, shown as red curves, fitted to the data set shown in Figure 1.2.

Parámetros, hiperparámetros y generalización

$$z = \sum_{i=0}^{M} w_i x^i$$

Hiperparámetro

parámetros

Parámetros:

- Son elegidos por el algoritmo de *optimización* para minimizar la *función de pérdida* medida sobre el *set de datos de* ***entrenamiento.***

Hiperparámetros:

- Son elegidos *antes* de entrenar el modelo. El criterio de elección es para conseguir un modelo que *generalice* mejor.

Parámetros, hiperparámetros y generalización

$$z = \sum_{i=0}^{M} w_i x^i$$

Hiperparámetro

parámetros

Conjunto de Datos

Entrenamiento	Evaluación

Parámetros:

- Son elegidos por el algoritmo de *optimización* para minimizar la *función de pérdida* medida sobre el *set de datos de* ***entrenamiento.***

Hiperparámetros:

- Son elegidos *antes* de entrenar el modelo. El criterio de elección es para conseguir un modelo que *generalice* mejor.

Poder de generalización:

- La performance esperada sobre un conjunto de datos *nuevo* (i.e. no visto durante el entrenamiento)