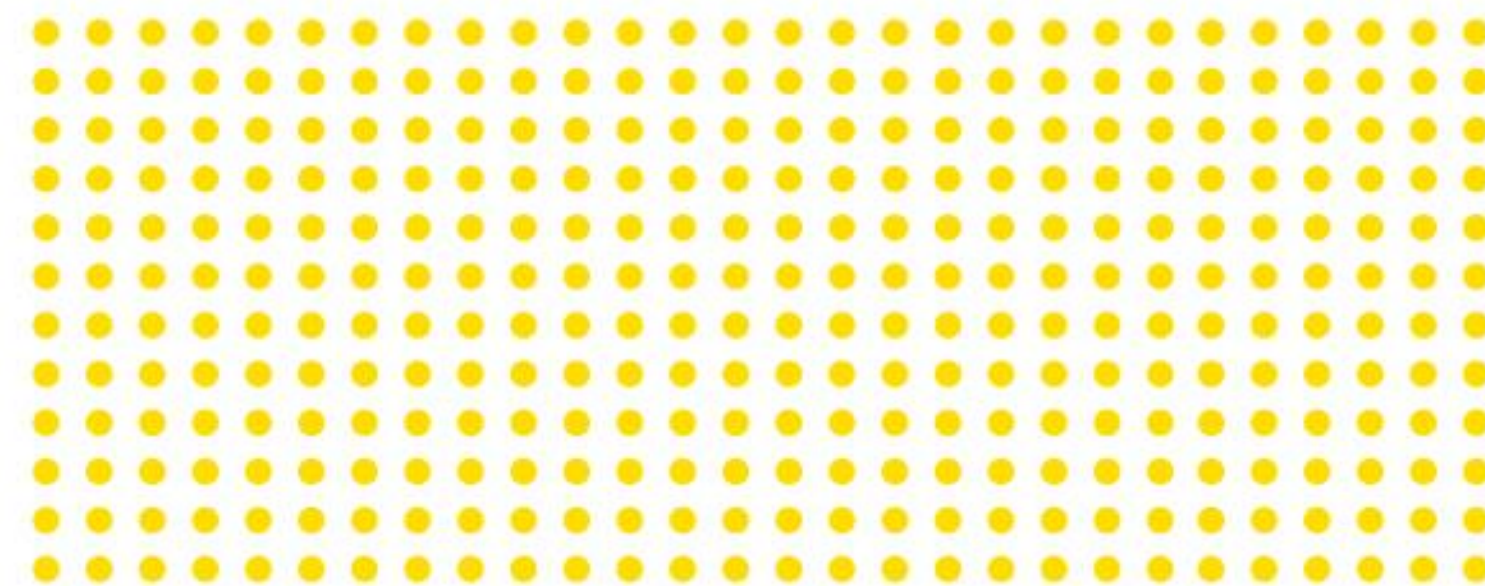


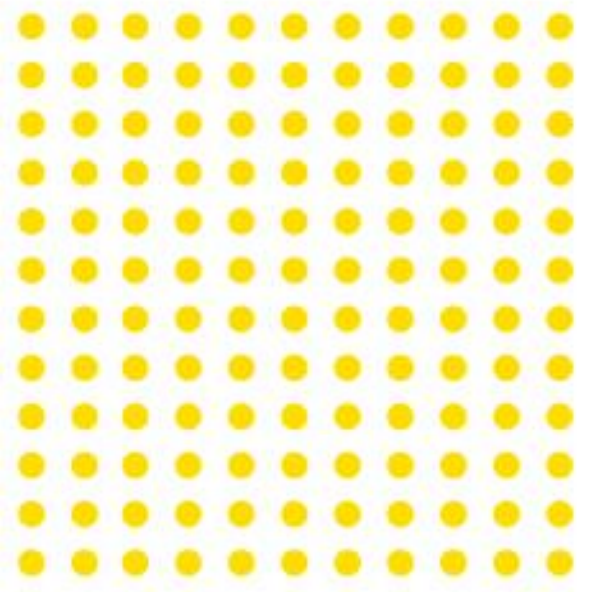


Universidad de
los Andes

Educación
Continua
Vicerrectoría Académica



Métricas de desempeño y selección de modelos

- 
- Modelo *Logit* y podemos comparar
 - Métricas de desempeño
 - Sobreajuste
 - Validación cruzada
 - Hiperparámetros

Parémonos en:



Recordemos

En problemas de clasificación queremos predecir columnas que nos indican si el dato tendrá o no alguna condición.

Los modelos que hemos visto y que vamos a ver hoy nos **predicen** la probabilidad de que el dato tenga la condición.

Definimos un **umbral**. Por encima de esa probabilidad decimos que se espera que un valor de 1 para la variable.

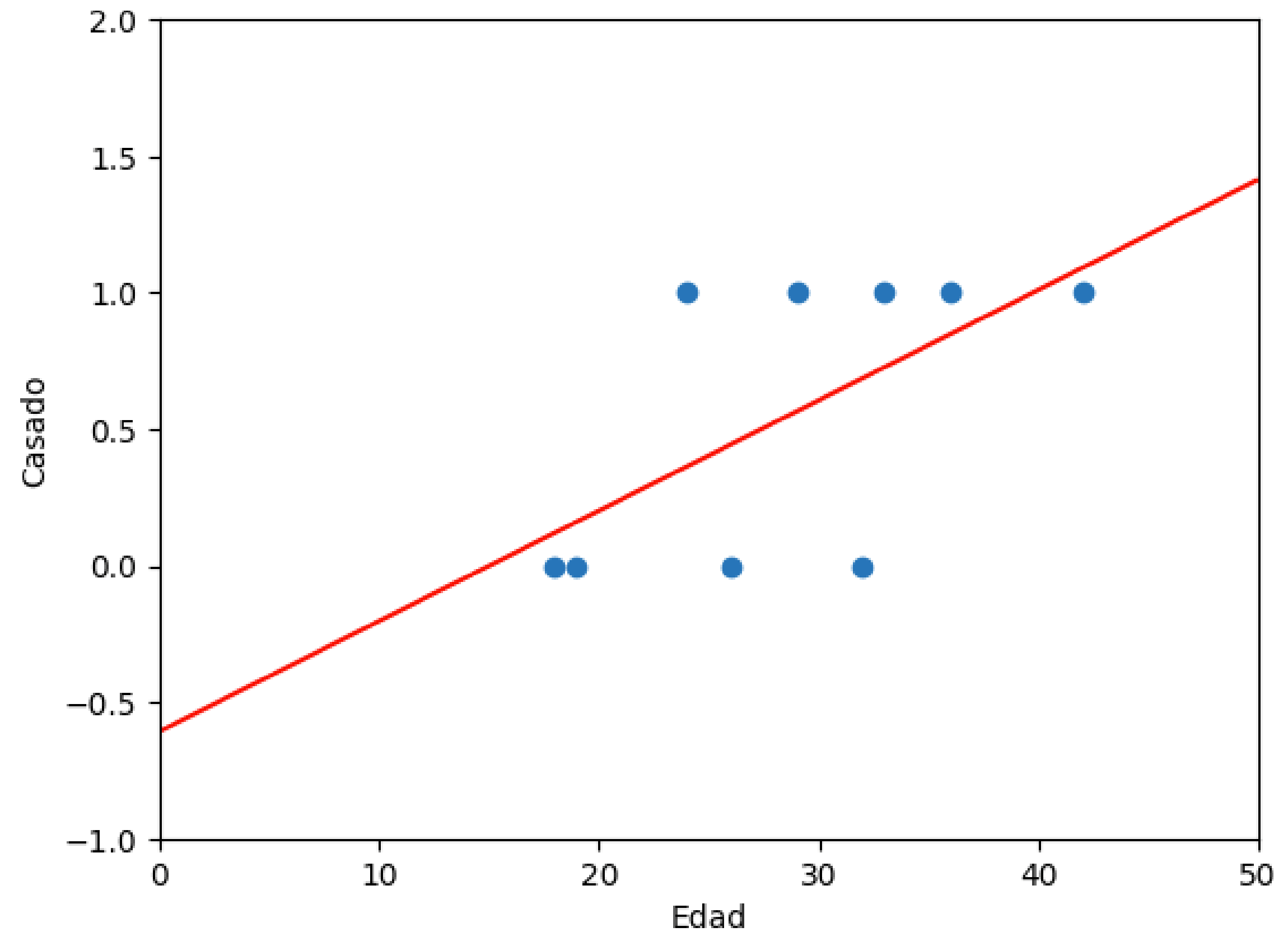
Casado	Edad	Hijos
1	24	0
0	18	0
0	26	1
0	19	0
1	36	1
1	29	0
1	42	0
0	32	0
1	33	1

Recordemos

¿Recuerdan el problema que tenía regresión lineal para clasificar?

¡Tenemos probabilidades inferiores a cero y superiores a 1!

Sabemos que toda probabilidad está entre cero y uno.



Una solución

Conocemos la función Sigmoid. Esta función siempre está por encima de 0 y por debajo de 1.

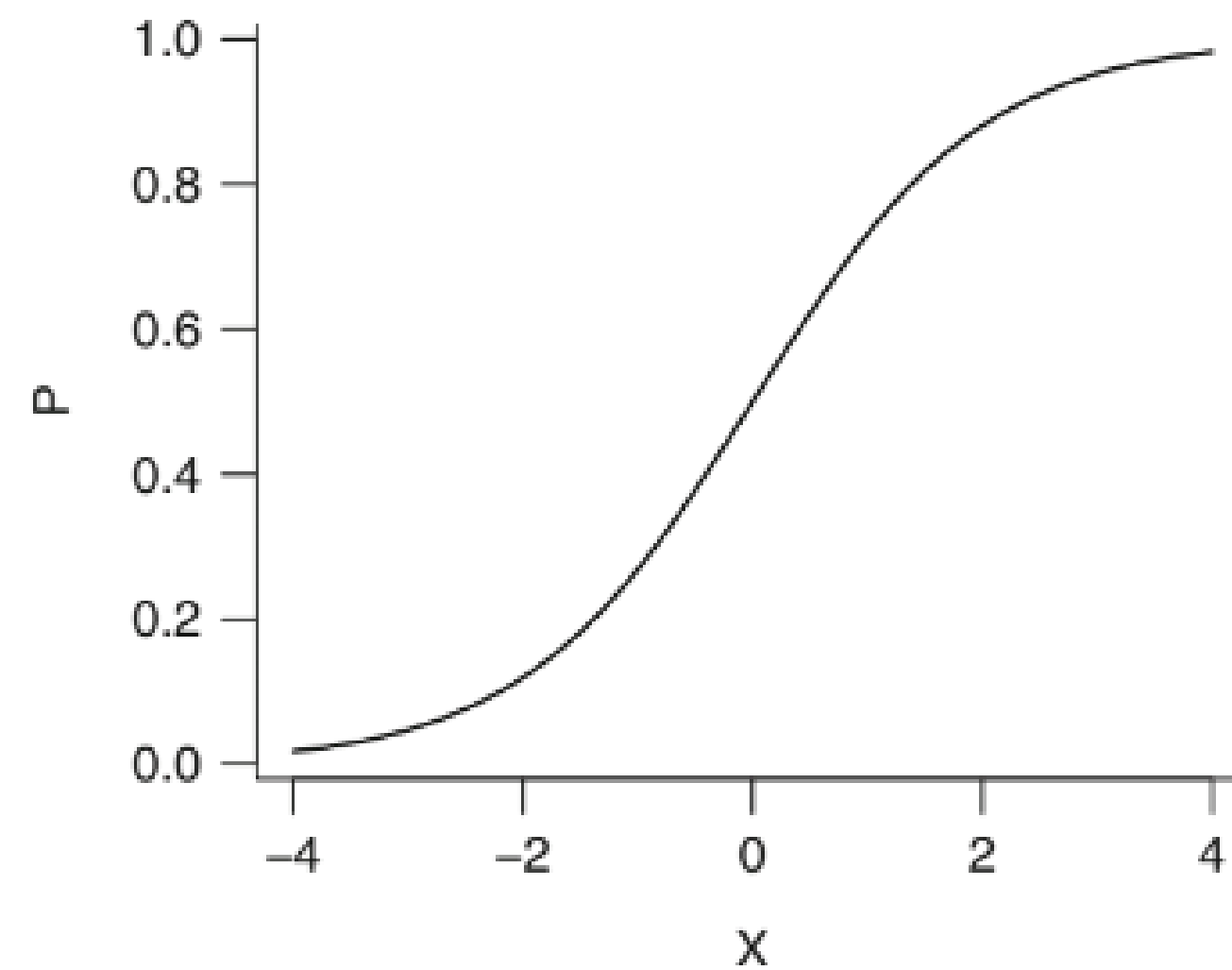
$$f(z) = \frac{1}{(1 + e^{-z})}$$

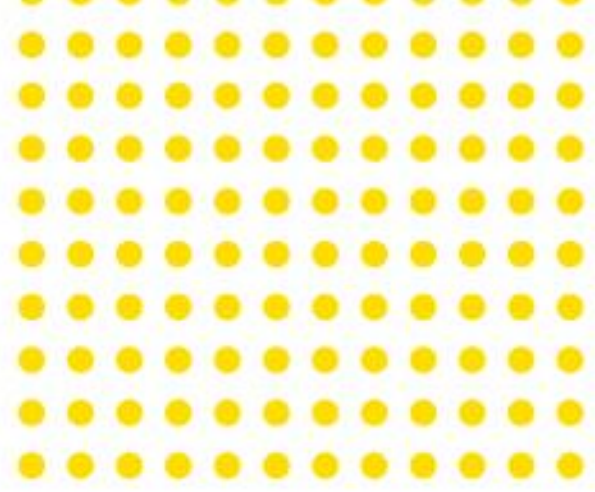
¡Justo lo que
necesitamos!

Una solución

Si le enchufamos el modelo lineal, ¡entonces tenemos predicciones acotadas de probabilidad!

$$P(Y = 1|X) = \frac{1}{(1 + e^{-(\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k)})}$$





Selección de modelos





Cuando escogemos un modelo

Tenemos que tomar decisiones:

- Identificamos **el problema**

¿supervisado? ¿no supervisado?...

- ¿**Qué modelo vamos a usar?**

¿Lineal? ¿Logit?...

- ¿**Qué características X** vamos a usar para predecir?

Características X

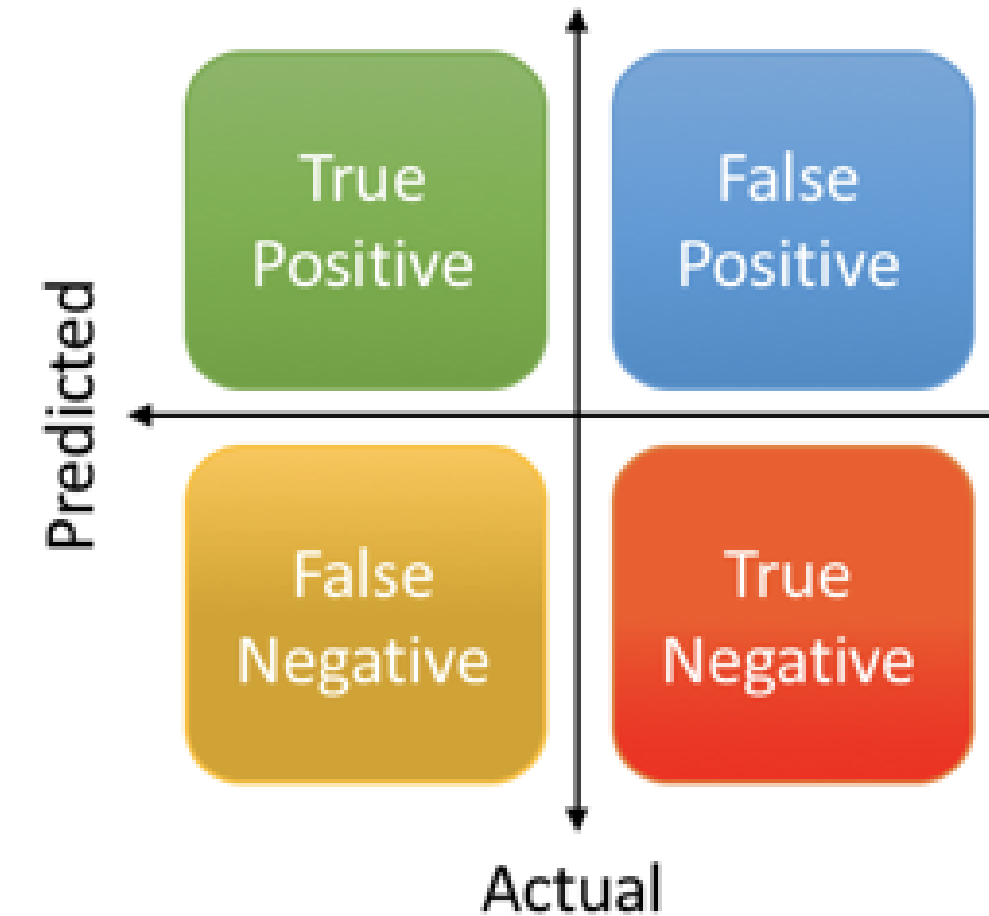
- ¿**Qué configuraciones** le vamos a cuadrar al modelo?

Usamos un umbral de 0.5?, 0.7?...



Qué tan bien le va al modelo ¿Qué tanto se equivoca?

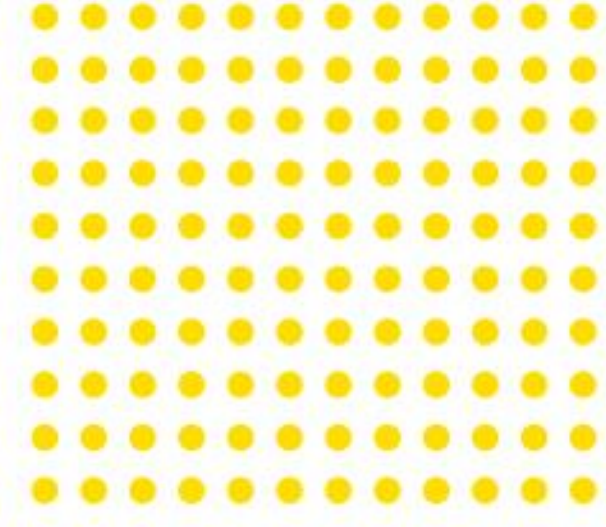
$$\begin{aligned}\text{Precision} &= \frac{\text{True Positive}}{\text{Actual Results}} \quad \text{or} \quad \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} \\ \text{Recall} &= \frac{\text{True Positive}}{\text{Predicted Results}} \quad \text{or} \quad \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}} \\ \text{Accuracy} &= \frac{\text{True Positive} + \text{True Negative}}{\text{Total}}\end{aligned}$$



Precisión: De la etiqueta que estoy prediciendo, que porcentaje de veces estuve en lo correcto.

Accuracy: De todas las predicciones que hice, en que proporción acerté.

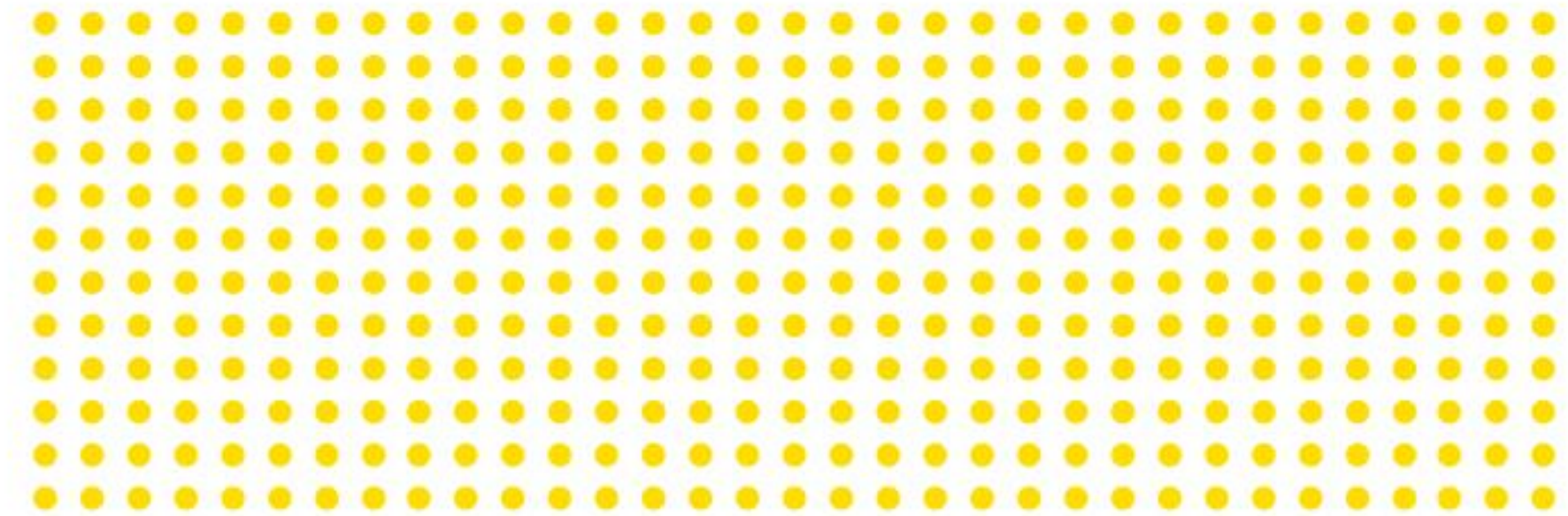
Recall: De la etiqueta que estoy prediciendo, que proporción de los registros encontré.



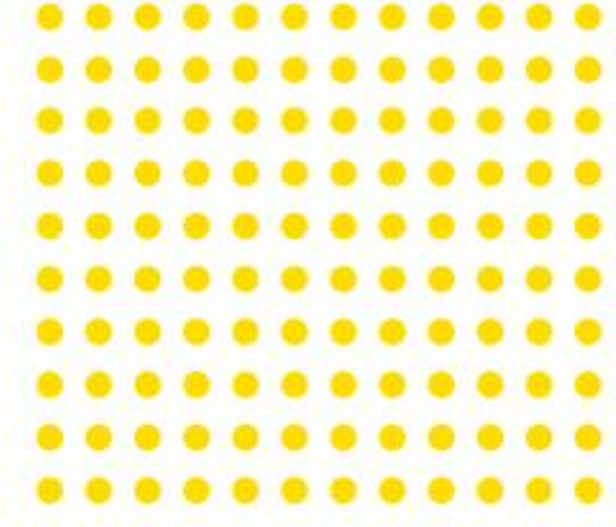
Pregunta

¿Qué es más grave?

- ¿Predecir algo como positivo, pero no lo era en realidad?
- ¿Predecir algo como negativo, pero en realidad era positivo?



2 minutos



Pregunta

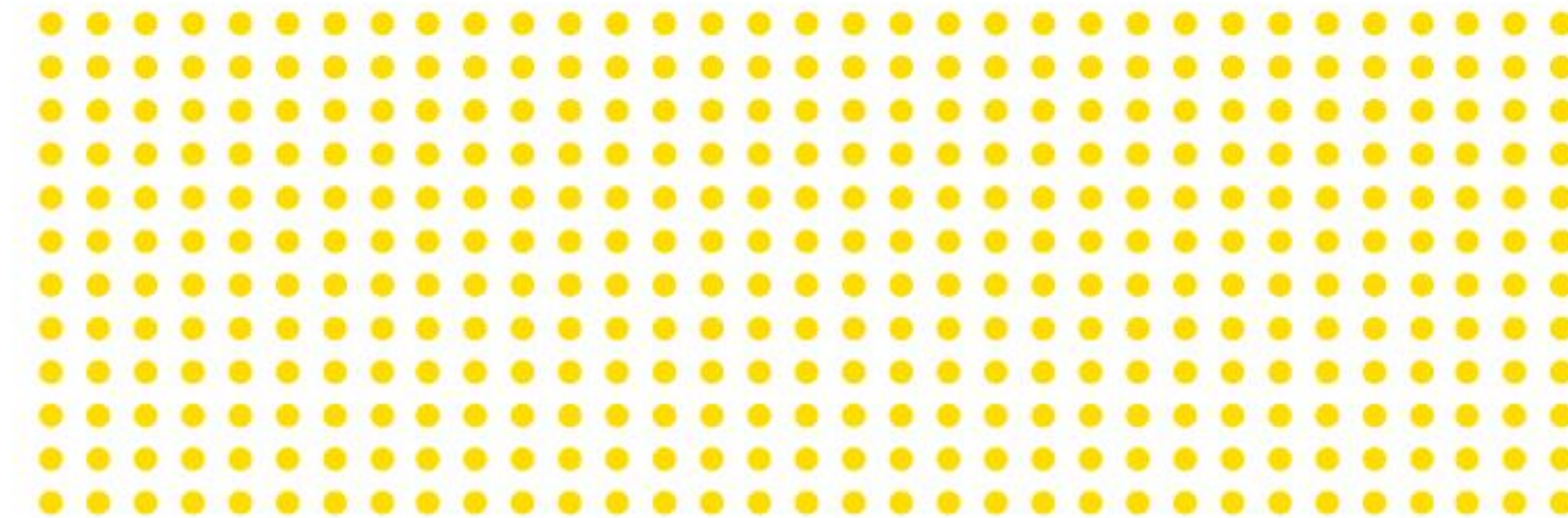
¿Qué es más grave?

- ¿Predecir algo como positivo, pero no lo era en realidad?

Identificó a un ladrón desde una foto.

- ¿Predecir algo como negativo, pero en realidad era positivo?

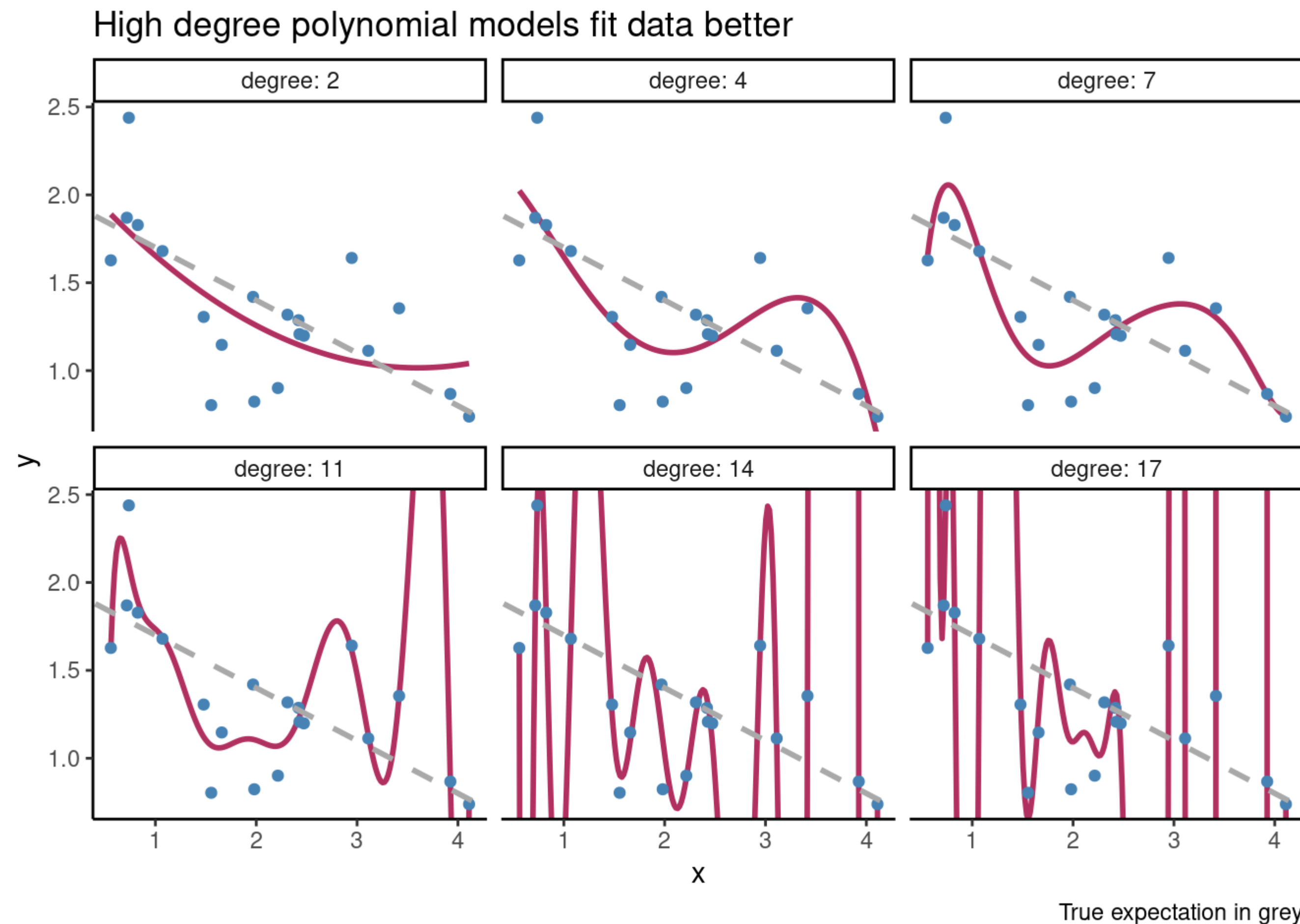
No admitido a la universidad.



Depende el contexto, a veces es menos grave cometer cierto tipo de error, o bien no es tan grave cometer otro.

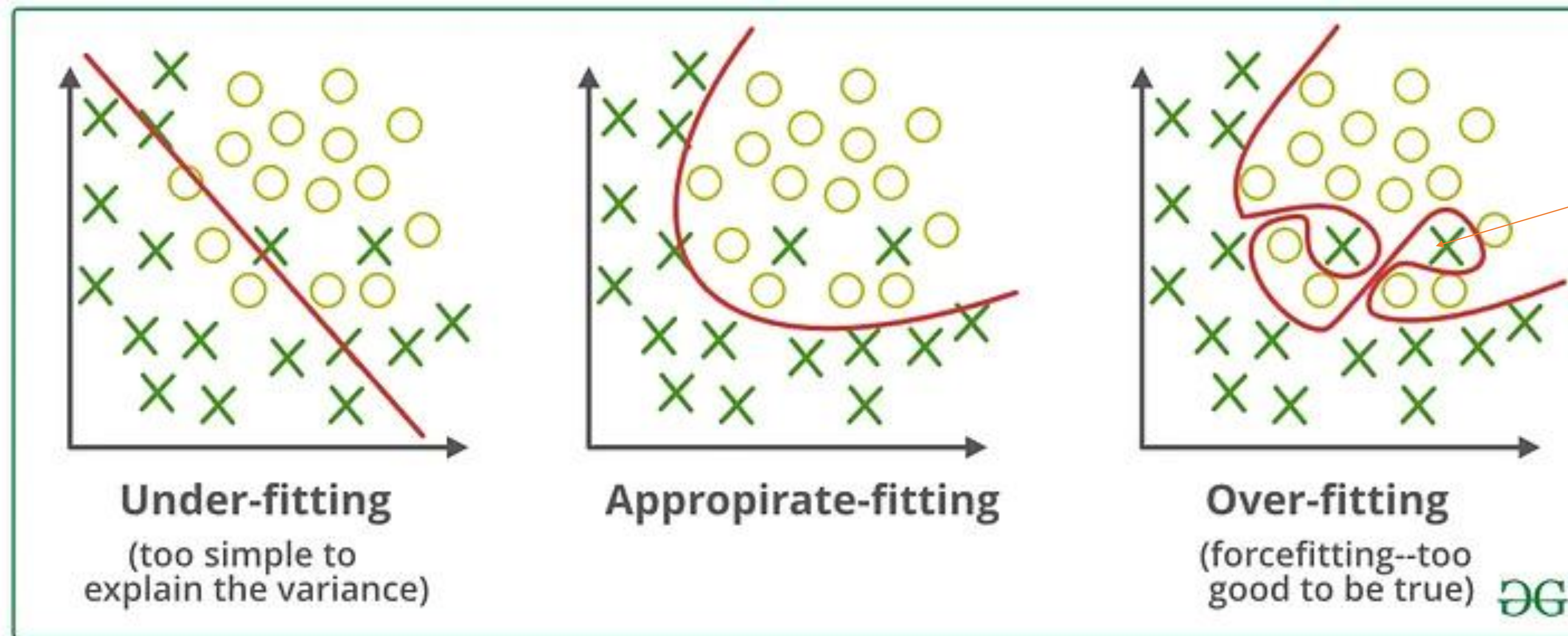
Sobreajuste

Con un modelo lo suficientemente complejo, podemos ajustar perfectamente la predicción a nuestros datos conocidos. ***¿Pero generalizando... Cómo le irá?***



Sobreajuste

Cuando armamos un clasificador, ensamblamos regiones

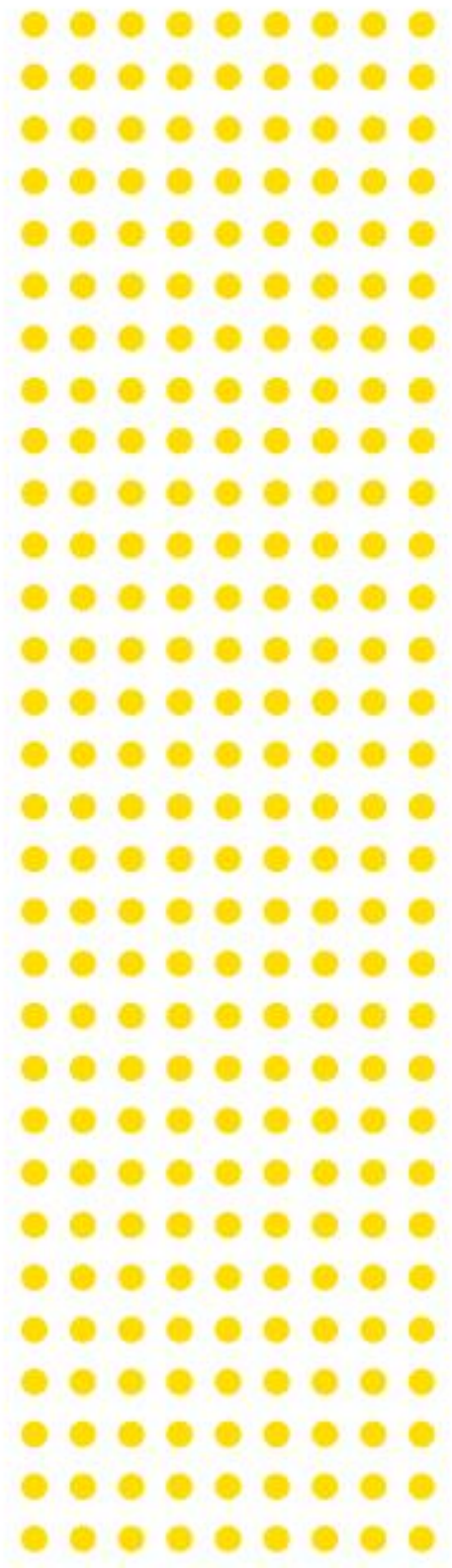


Todo dato que caiga aquí sería clasificado como verde.

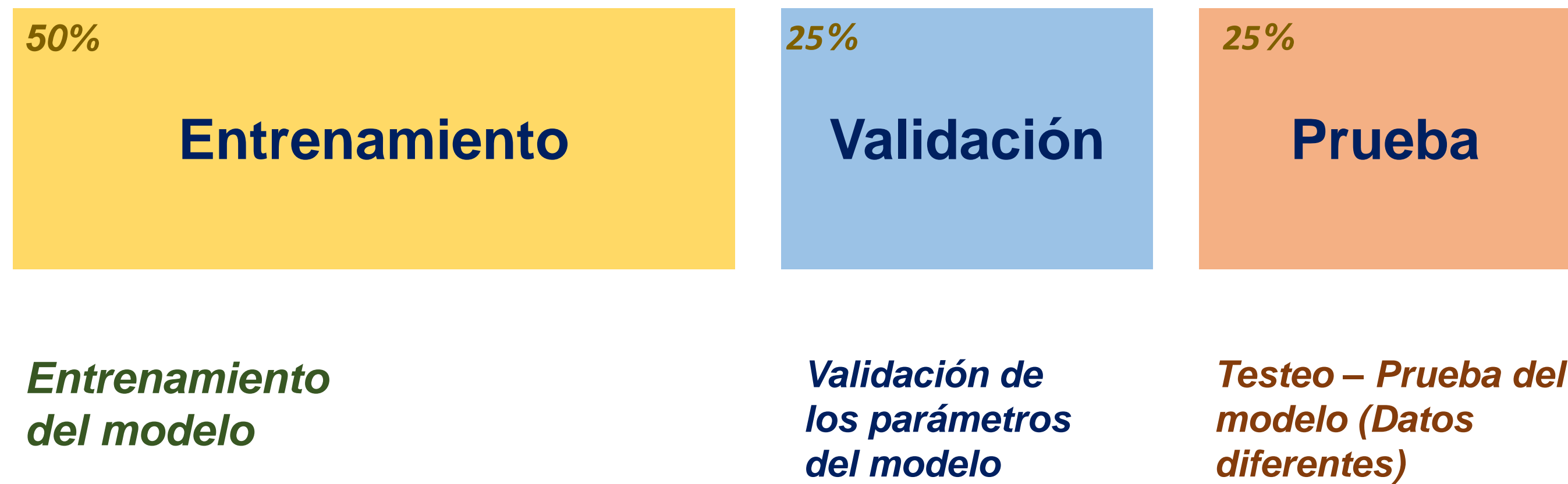
Para evitar esto separamos los datos

- Vamos a hablar más en detalle sobre esto más adelante.
- Práctica importante como mínimo: **separar nuestros datos en dos grupos: entrenamiento y prueba (70-30).**
- Estimamos el modelo con los **datos de entrenamiento**, y luego vemos qué tan bien predice **los de prueba**.
- Así nos cuidamos de que se acostumbre demasiado a los **datos que ya conoce**, y prediga bien **datos que nunca ha visto**.

Casado	Edad	Hijos
1	24	0
0	18	0
0	26	1
0	19	0
1	36	1
1	29	0
1	42	0
0	32	0
1	33	1



También podemos separar en tres grupos



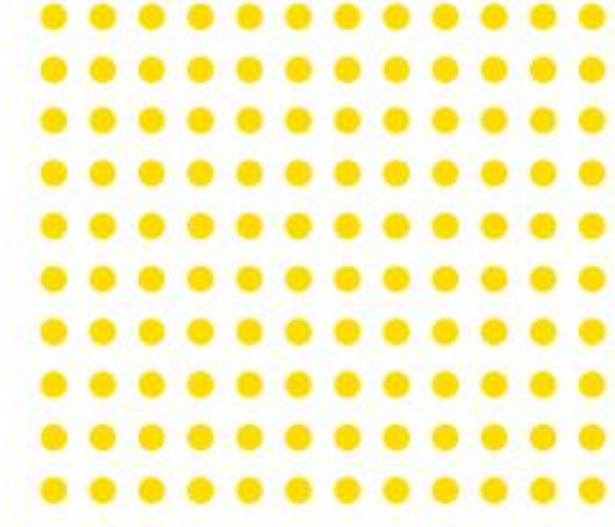
El problema es que dejamos de utilizar muchos datos para entrenar...

Validación cruzada

Es una técnica para evaluar modelos mediante el entrenamiento de varios modelos en subconjuntos de datos y probarlos con el subconjunto complementario de los datos.

Utilizamos la validación cruzada para detectar el sobreajuste, es decir, en aquellos casos en los que no se logre generalizar un patrón.

Iteraciones	Validación Cruzada (K = 5)				
1	Prueba	Entrenamiento	Entrenamiento	Entrenamiento	Entrenamiento
2	Entrenamiento	Prueba	Entrenamiento	Entrenamiento	Entrenamiento
3	Entrenamiento	Entrenamiento	Prueba	Entrenamiento	Entrenamiento
4	Entrenamiento	Entrenamiento	Entrenamiento	Prueba	Entrenamiento
5	Entrenamiento	Entrenamiento	Entrenamiento	Entrenamiento	Prueba



Configuraciones

Una vez escogimos modelo.

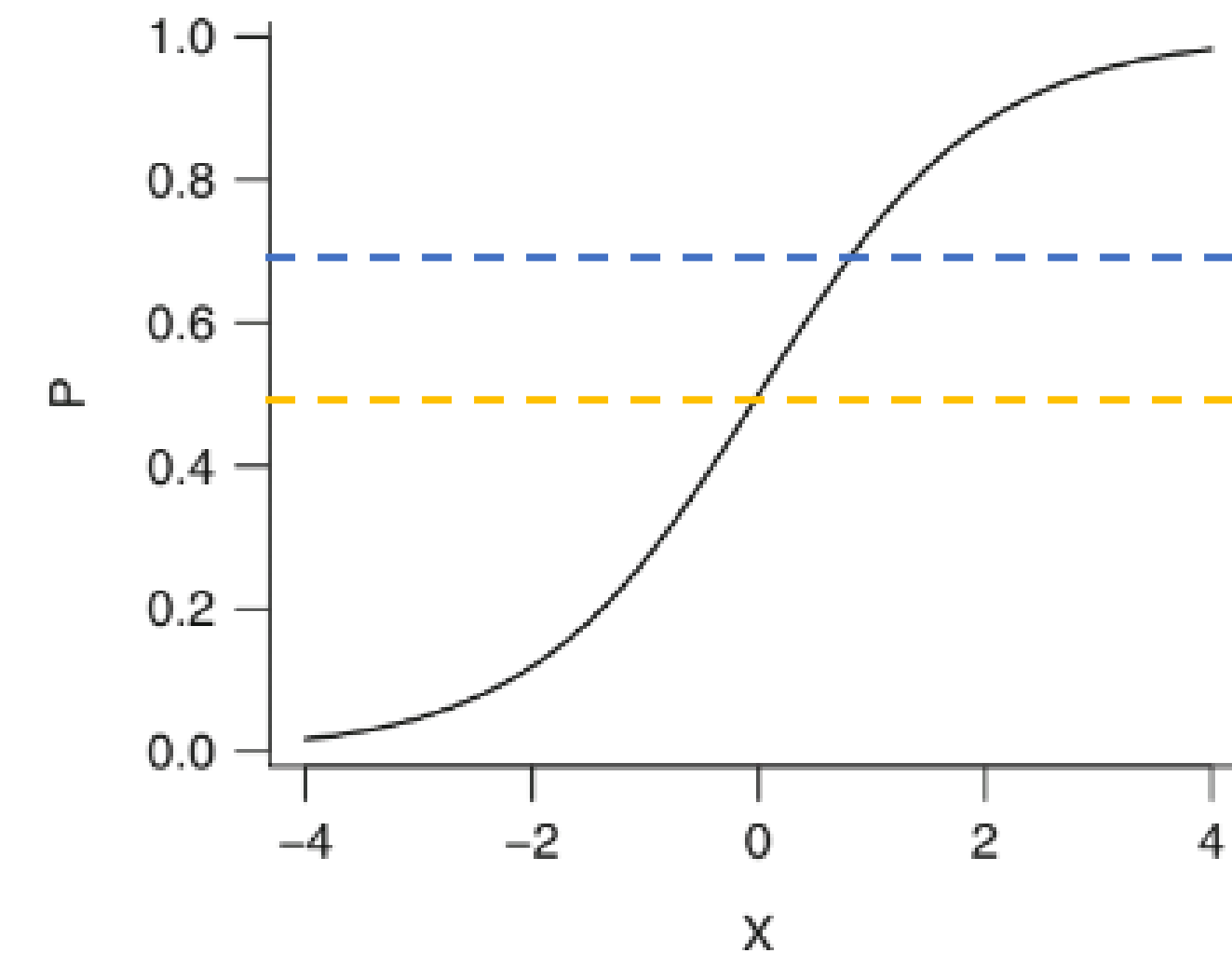
Y escogimos características X.

Tenemos que calibrar bien las configuraciones.

Cada modelo se configura diferente

Depende el contexto, a veces es menos grave cometer cierto tipo de error, o bien no es tan grave cometer otro.

Por ejemplo, si movemos el umbral, podemos mitigar cometer cierto tipo de error...



Hablaremos más de esto: hiperparámetros.



iAI
Colab!

¡Gracias!

Aprendiendo juntos a lo largo de la vida

educacioncontinua.uniandes.edu.co

Síguenos: **EdcoUniandes**     



**Educación
Continua**
Vicerrectoría Académica

Universidad de los Andes | Vigilada Mineducación. Reconocimiento como Universidad: Decreto 1297 del 30 de mayo de 1964. Reconocimiento personería jurídica: Resolución 28 del 23 de febrero de 1949 Minjusticia.

