

# Machine Learning

El problema del aprendizaje

Error y ruido

Angel Vázquez-Patiño  
angel.vazquezp@ucuenca.edu.ec

Departamento de Ciencias de la Computación  
Universidad de Cuenca

16 de octubre de 2017

# Objetivos

1. Entender qué significa aproximación al comparar una hipótesis con la función objetivo  $f$
2. Entender que hay ruido en la salida de  $f$  que no está determinada únicamente por la entrada

# Contenido

Medidas del error

Objetivo  $f$  con ruido

# Medidas del error

# Medidas del error

- Nunca se replica  $f$
- Se debe definir una medida del error que cuantifique qué tan bien  $g$  aproxima  $f$
- La elección define la salida del proceso de aprendizaje
- Diferentes hipótesis finales
- ¿Entonces qué criterios usar para escogerla?

# Medidas del error

## Medida del error

- Una medida del error o función del error cuantifica qué tan bien cada hipótesis  $h \in \mathcal{H}$  en el modelo aproxima la función objetivo  $f$

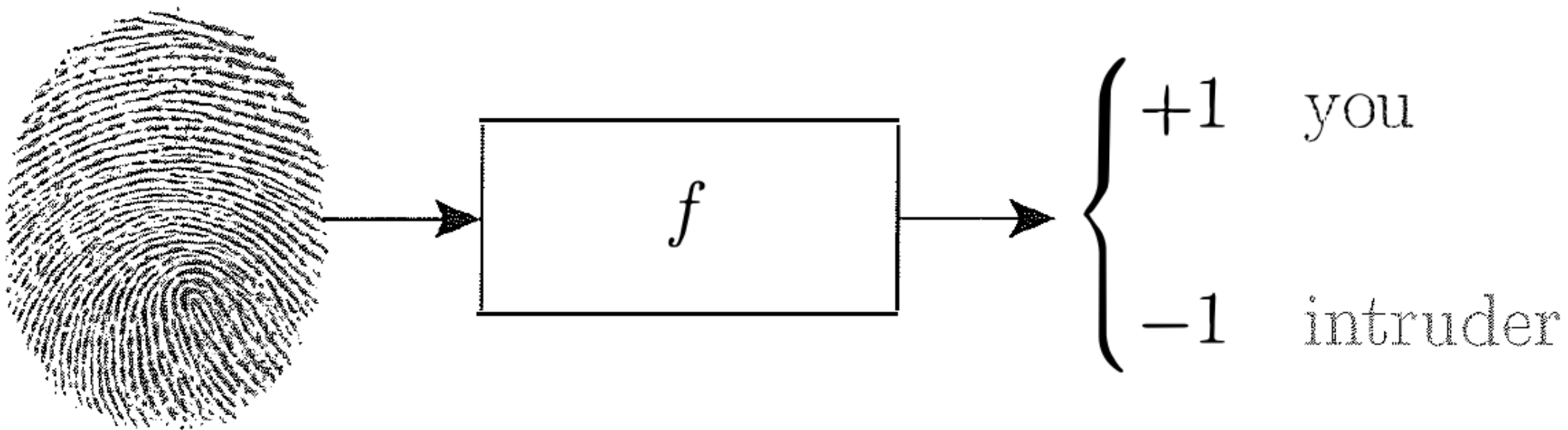
$$\text{Error} = E(h, f)$$

- Aunque definida en el  $h$  y  $f$  completo, es casi universalmente definido basado en los errores de los data points  $\mathbf{x}$  individuales

$$e(h(\mathbf{x}), f(\mathbf{x})) = \llbracket h(\mathbf{x}) \neq f(\mathbf{x}) \rrbracket$$

# Medidas del error

- La misma tarea de aprendizaje en diferentes contextos pueden garantizar el uso de diferentes medidas del error
- Ejemplo



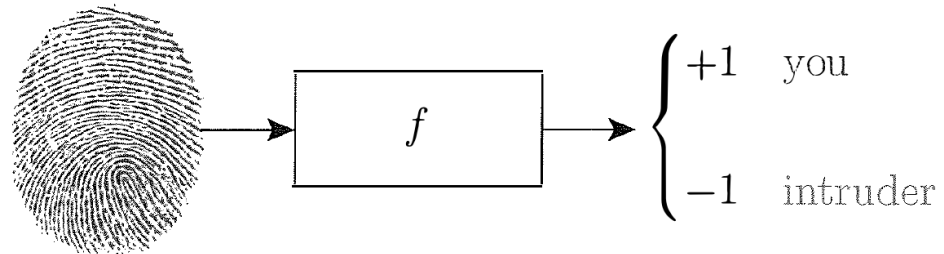
# Medidas del error

		$f$	
		+1	-1
$h$	+1	no error	false accept
	-1	false reject	no error

¿Cómo definir la medida del error?



# Medidas del error



## Casos de aplicación

- Descuento en el bus
  - Falso rechazo
  - Falsa aceptación
- Sala de servidores de la universidad
  - Falso rechazo
  - Falsa aceptación

		$f$	
		$+1$	$-1$
$h$	$+1$	no error	false accept
	$-1$	false reject	no error

# Medidas del error

- Los costos de cometer los diferentes errores se pueden tabular

		$f$	
		+1	-1
$h$	+1	0	1
	-1	10	0

		$f$	
		+1	-1
$h$	+1	0	1000
	-1	1	0

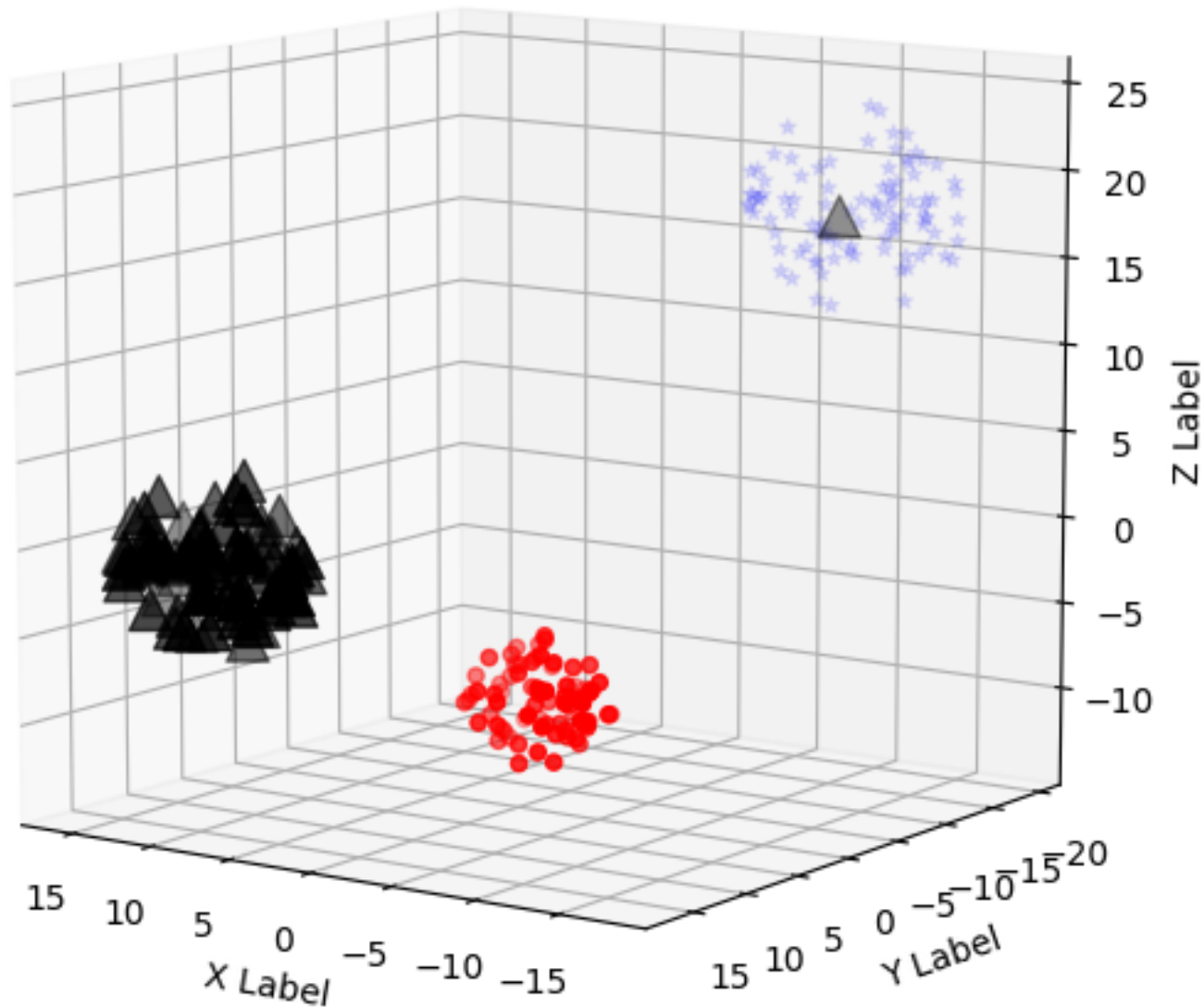
- Esto daría el peso de los diferentes tipos de errores
- El algoritmo de aprendizaje **minimiza** la función

# Objetivo $f$ con ruido

# Objetivo con ruido

- En la práctica  $\mathcal{D}$  que usamos para aprender no es generada por una función objetivo  $f$  determinística
- Para dar un crédito, dos personas con exactamente la misma información ( $x_d$ ) pueden haber tenido una respuesta diferente (e.g. decisión de un humano)

# Objetivo con ruido



# Objetivo con ruido

## Modelamiento

- Y sería una variable afectada por  $\mathbf{x}$  y no determinada por  $\mathbf{x}$

*distribution*  $P(y \mid \mathbf{x})$  instead of a target function  $y = f(\mathbf{x})$

- Un data point es generado por la distribución conjunta

- $$P(\mathbf{x}, y) = P(\mathbf{x})P(y \mid \mathbf{x})$$

# Conceptos y términos importantes

# Conceptos y términos importantes

- Lo que se entiende al decir que los datos tienen ruido
- Outlier (valor atípico)
- Lo que significa aproximación cuando se dice que la hipótesis aproxima bien la función objetivo  $f$
- Qué otras cosas determinan que haya ruido en  $f$



# Referencias

- Abu-Mostafa, Y.S., Magdon-Ismail, M., Lin, H.-T., 2012. Learning from data: a short course. AMLbook.com, USA.
- Raschka, S., 2016. Python machine learning, Community experience distilled. Packt Publishing, Birmingham Mumbai.

# Preguntas

