Aprendizaje Automático Profundo (Deep Learning)

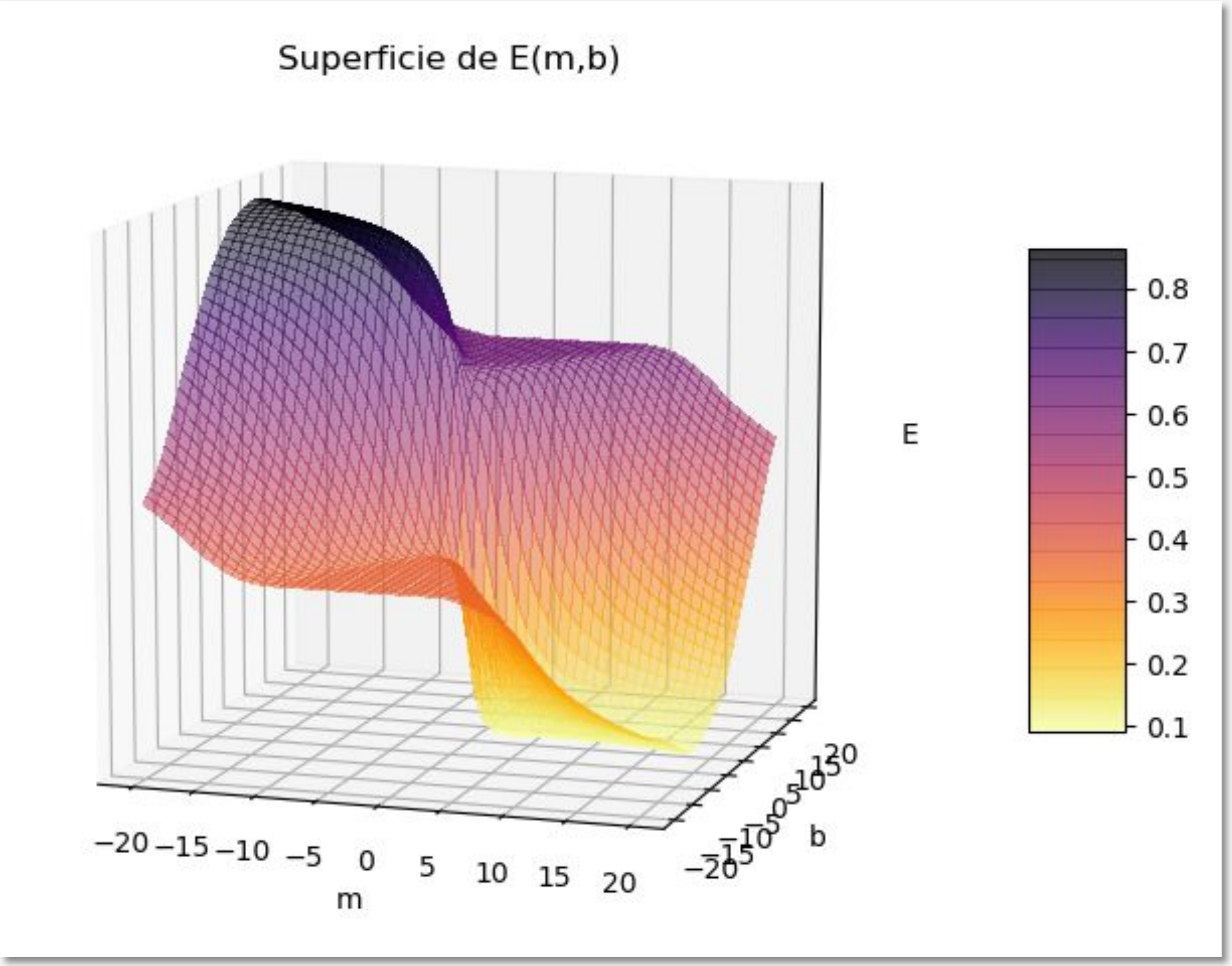
Dr. Facundo Quiroga - Dr. Franco Ronchetti

Regresión Logística Función de Error Entropía Cruzada

Error cuadrático medio

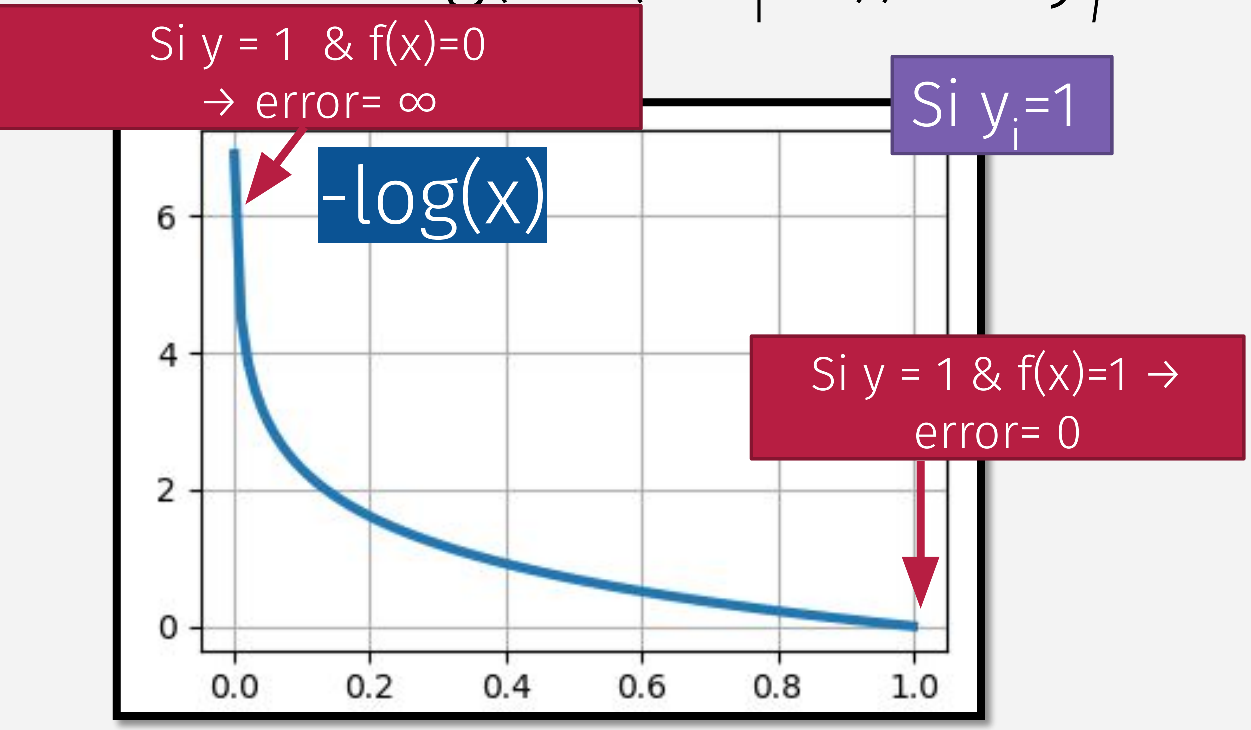
- $E = 1/n \sum_i^n Ei$
- $Ei = (\sigma(mx_i+b)-y_i)^2$
- Funciona, pero
  - E es ECM de σ(mx+b)
  - E no es convexa
- Podemos definir otro E que sea convexa
  - **Entropía Cruzada**
    - Función de error
    - Convexa para $\sigma(mx_i+b)$

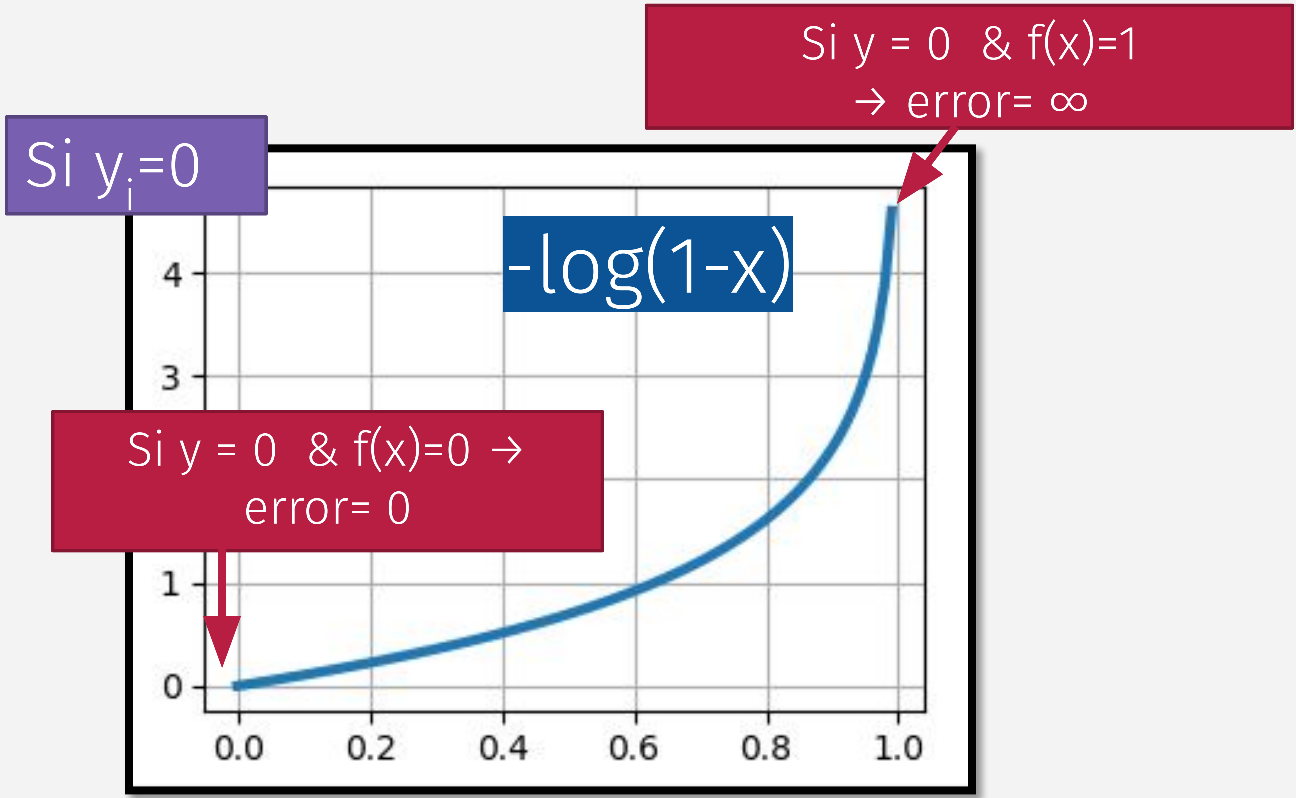
ECM de σ(mx+b)



Entropía Cruzada

- $E = 1/n \sum_i^n Ei$
- $Ei =$
  - $-\log(\sigma(mx_i+b))$ *si* $y_i=1$
  - $-\log(1-\sigma(mx_i+b))$ *si* $y_i=0$
- $Ei =$
  - $-\log(f(x_i))$ *si* $y_i=1$
  - $-\log(1-f(x_i))$ *si* $y_i=0$





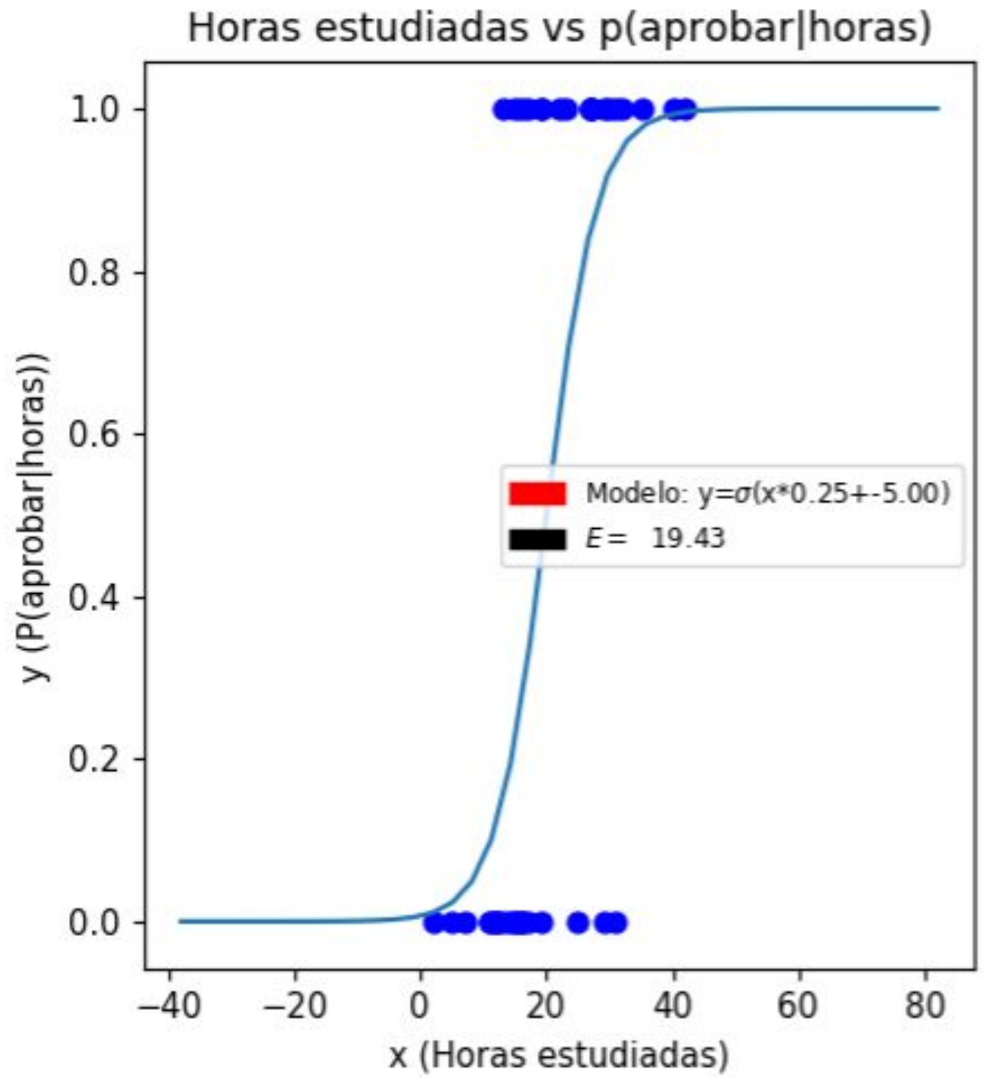
$f(x) = \sigma(mx_i+b)$ es un valor entre 0 y 1 → entrada a $\log(x)$

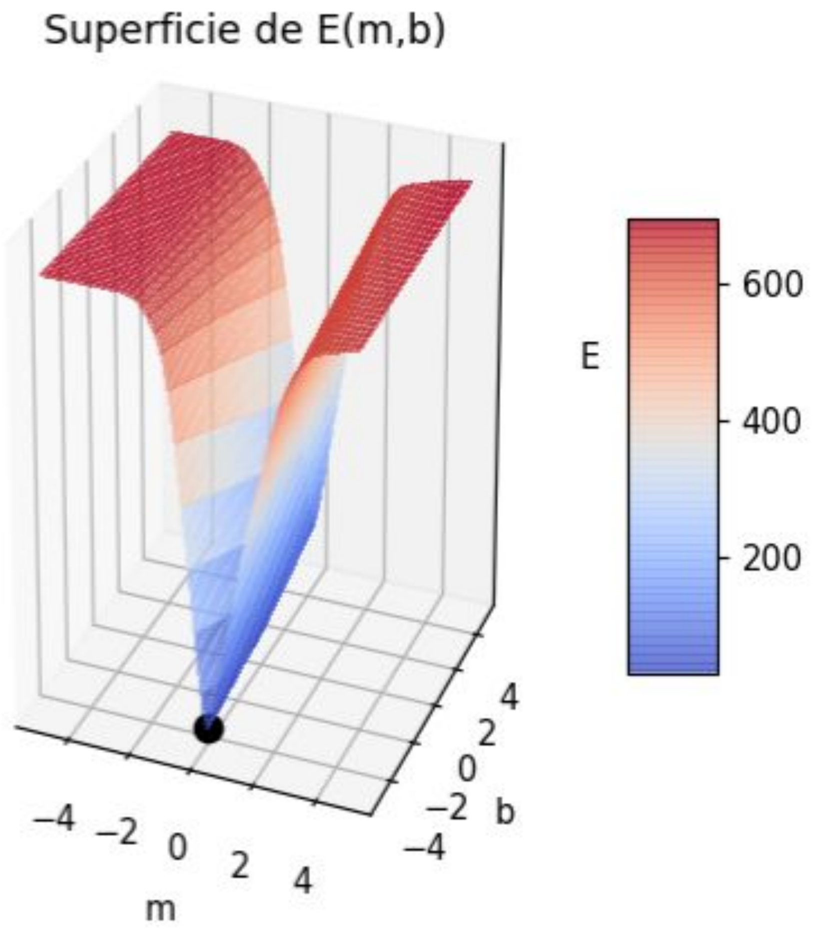
Entropía Cruzada - Fórmula vectorial

- Ei =
  - $-\log(f(x_i))$ *si* $y_i=1$
  - $-\log(1-f(x_i))$ *si* $y_i=0$
- Podemos escribir Ei como
  - Ei = $y_i\,[-\log(f(x_i))]$ + $(1-y_i)\,[-\log(1-f(x_i))]$
    - Ejemplo
      - $y_i=1$
      - $y_i\,[-\log(f(x_i))] + (1-y_i)\,[-\log(1-f(x_i))]$
        $= 1\,[-\log(f(x_i))] + 0\,[-\log(1-f(x_i))]$
        $= [-\log(f(x_i))]$

- y_i actua como función indicadora
  - Si $y_i=1$, contribuye $[-\log(f(x_i))]$
  - Si $y_i=0$, contribuye 0
- Idem $1-y_i$ al revés

Entropía Cruzada





Interpretación de entropía cruzada

- Distancia entre distribuciones
  - f(x) = 0.3
    - P(x es de clase = 1) = 0.3
    - P(x es de clase = 0) = 1 - P(x es de clase = 1) = 1 - 0.3 = 0.7
      - P(x es de clase = 1) y P(x es de clase = 0) forman una **distribución de probabilidad**
- ECM
  - Mide distancia entre **puntos**
  - Distancia **euclídea al cuadrado**
- Entropía Cruzada
  - Mide distancia entre **distribuciones de probabilidad**
  - Distancia **Kullback-Leib**ler (o Divergencia)

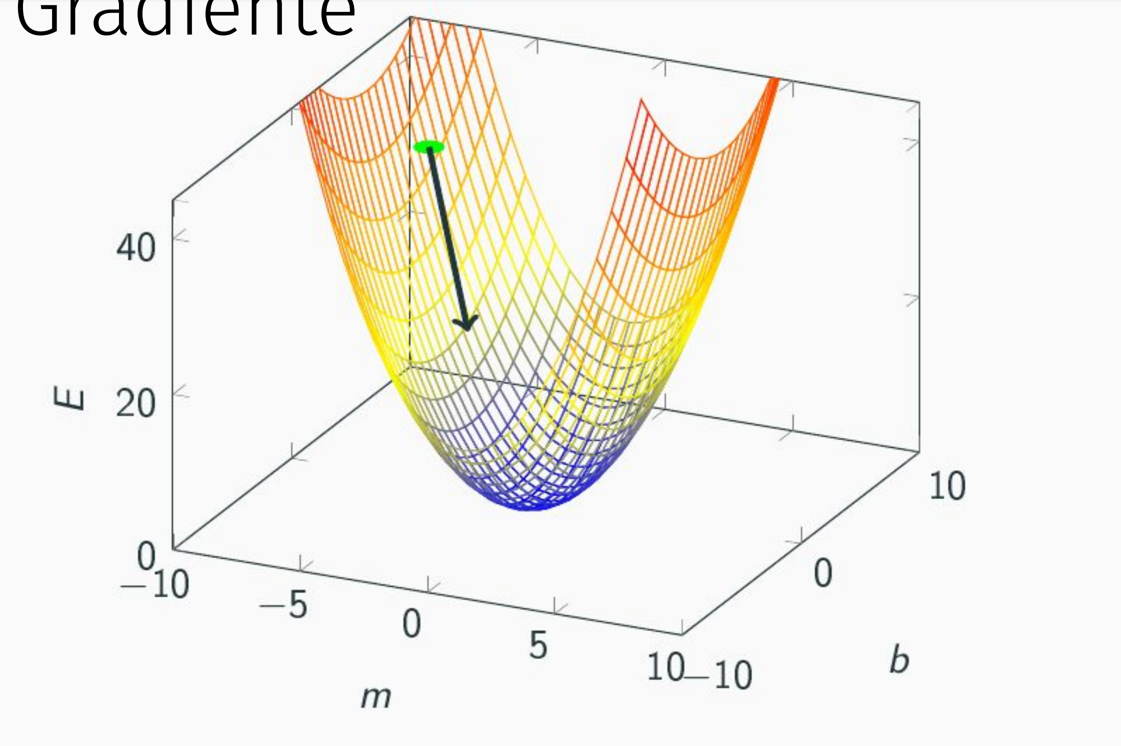
Distancia entre distribuciones

- Sean p y q distribuciones sobre n valores
  - p_i y q_i = probabilidad de tomar el valor i
  - KullbackLeibler(p,q) = $\Sigma_i^n p_i \log(q_i/p_i)$
- KullbackLeibler(B,C)
  = KullbackLeibler((0,1), (1,0))
  = $B_1\log(C_1/B_1) + B_2\log(C_2/B_2)$
  = $0^*\infty + 1^*\ \infty\ = \infty$
  - Peor caso
- KullbackLeiber(B,A)
  = KullbackLeiber((1,0),(0.3,0.7))
  = 1 log(0.3/1) + 0 log(0.7/0) = log(0.3)
- KullbackLeiber(A,B)
  = KullbackLeiber((0.3,0.7), (1,0))
  = 0.3 log(1/0.3) + 0.7 log(0/0.7) = ...

Distribución	1	2
A	0.3	0.7
B	0.0	1.0
C	1.0	0.0

Derivadas de Entropía Cruzada

- $\delta E(m,b)/\delta b = 1/n \sum_i^n y_i - f(x_i)$
- $\delta E(m,b)/\delta m = 1/n \sum_i^n (y_i - f(x_i))\, x_i$
  - Son iguales que las de Regresión Lineal!
  - [**ojo**: $f(x)=\sigma(mx+b)$]
- Mismas ecuaciones de Descenso de Gradiente
  - $b = b - \alpha\, 1/n \sum_i^n (y_i - f(x_i))$
  - $m = m - \alpha\, 1/n \sum_i^n (y_i - f(x_i))\, x_i$



Accuracy: *Métrica* de clasificación

- Valores de entropía cruzada: 0 a +∞
  - Difíciles de interpretar
- Accuracy = % de ejemplos que clasificó correctamente

Aprobado	Predicción	Acierto
0	0	1
1	1	1
1	0	0
0	1	0
0	0	1
0	0	1
1	0	0
0	1	0

- **Acierto**
  - 1 si Aprobado = Predicción
- **Accuracy**
  - Promedio de columna **Acierto**
- Ejemplo
  - N = 8
  - aciertos=4
  - Accuracy = aciertos/N
    - = 4/8 = 0.5
  - Accuracy(%) = 50%

Resumen

- **Error Cuadrático Medio**
  - No es convexo
    - Respecto de σ(mx+b)
- **Entropía Cruzada**
  - Es convexa
  - Derivadas simples
    - E iguales a Regresión Lineal + ECM
  - Difícil de interpretar
    - Accuracy: *métrica* de clasificación
- 2 o m variables de entrada?
  - No lo vimos
  - Generalización similar a la de Regresión Lineal
  - Mismos resultados

Regresión Logística - Notebooks

- Abrir el archivo **Regresion Logistica 2D - Modelo.ipynb**
  - Modificar los parámetros m y b , observar cómo cambia la función de predicción y el error.
- Abrir el archivo **Regresion Logistica 2D - Modelo.ipynb**
  - Idem caso anterior
- Abrir el archivo **Regresion Logística - Aprendizaje.ipynb**
  - Experimentar con los valores iniciales de m , b, así como los hiperparámetros α y el número de iteraciones
- Abrir el archivo **Regresion Logística 2D -Aprendizaje.ipynb**
  - Idem caso anterior