Introducción a la Inteligencia Artificial

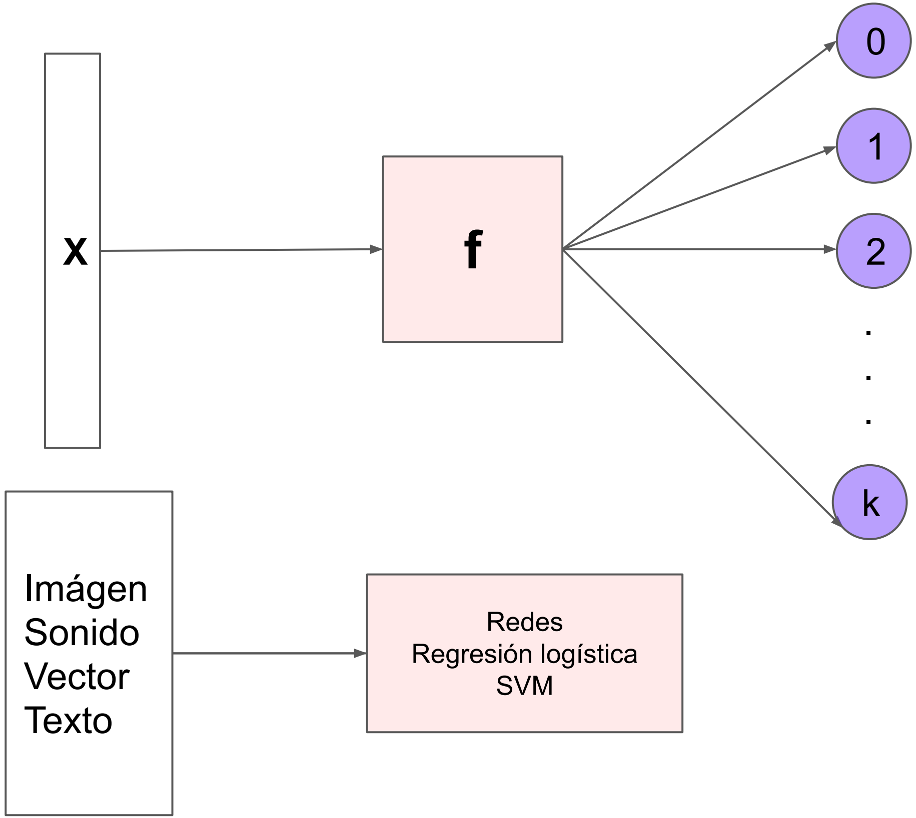
Facultad de Ingeniería
Universidad de Buenos Aires

Índice

Clase 6

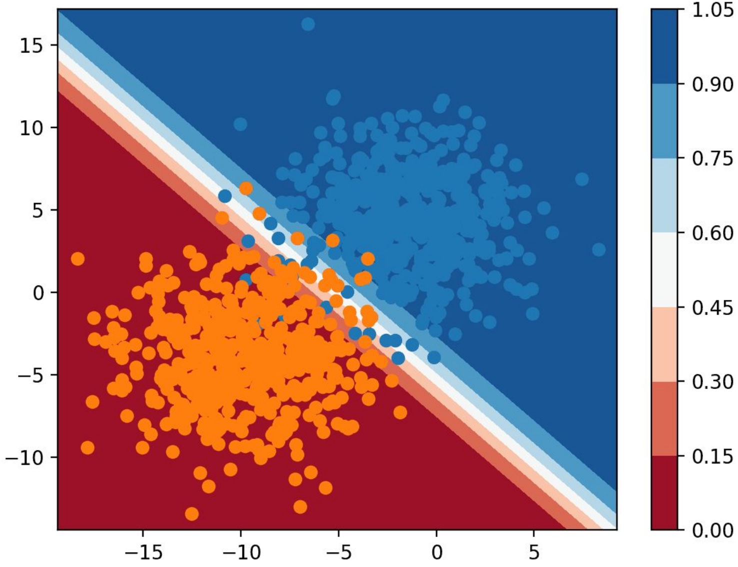
1. Clasificación Binaria
   a. Motivación
   b. Regresión Logística - Ejercicio de Aplicación
   c. Regresión Logística - Teoría
2. Clasificación Multiclase
   a. Motivación
   b. Softmax
   c. Ejercicio de Aplicación
3. Ejercicio integrador

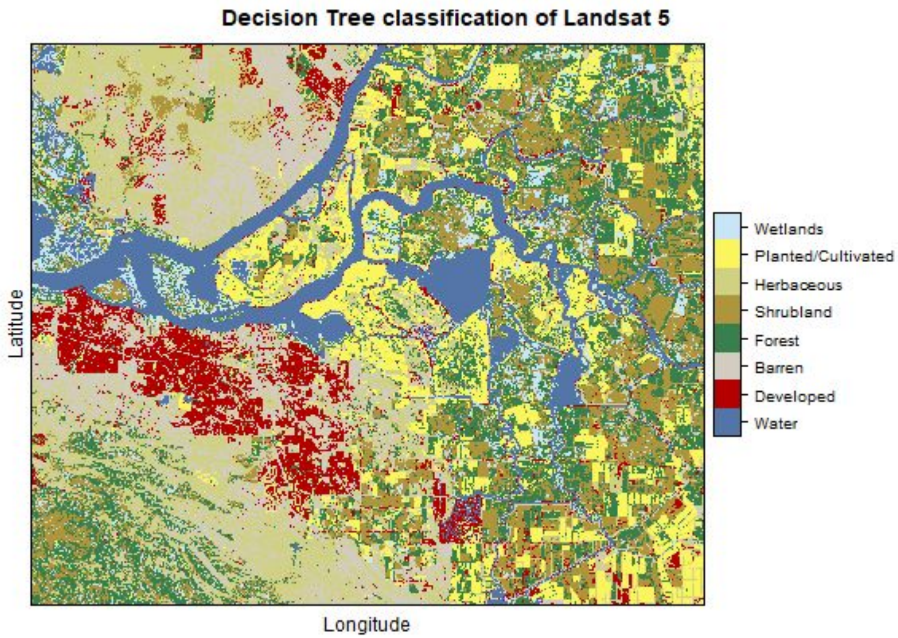
Clasificación



$$f : R^D \rightarrow \{0, 1, 2, \cdots, k\}$$

Clasificación

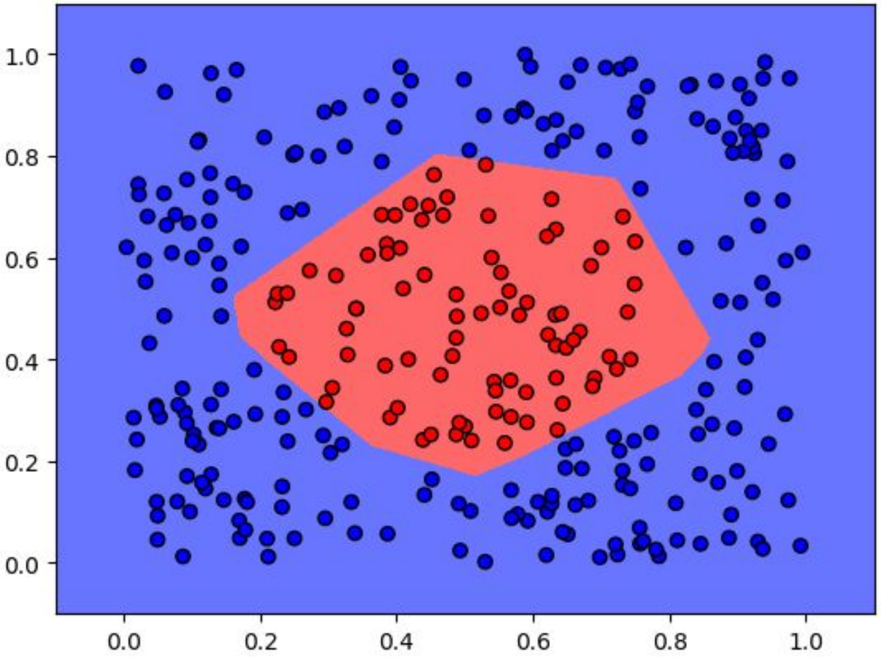




Clasificación binaria

Clasificación Binaria - Ejemplos

- Detección de fraudes
- Diagnóstico médico
- Detección de spam
- Sentiment Analysis
- Detección de objetos
- Outliers



Clasificación Binaria - Diagnóstico médico

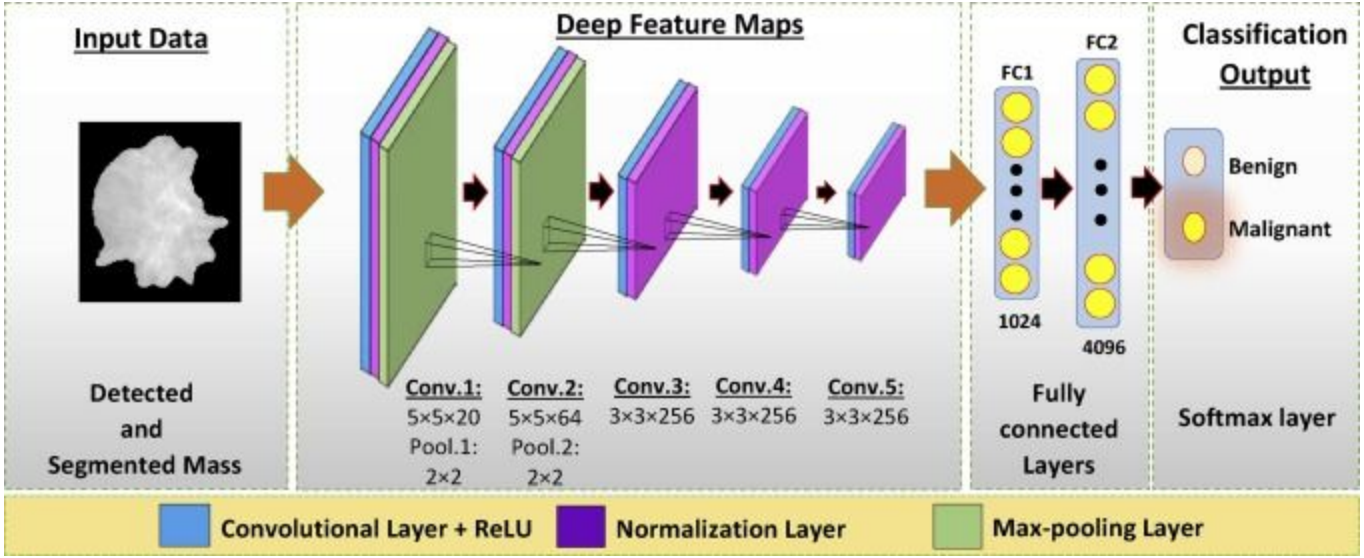


Imagen de: “A fully integrated computer-aided diagnosis system for digital X-ray mammograms via deep learning detection, segmentation, and classification“

Clasificación Binaria - Spam detection

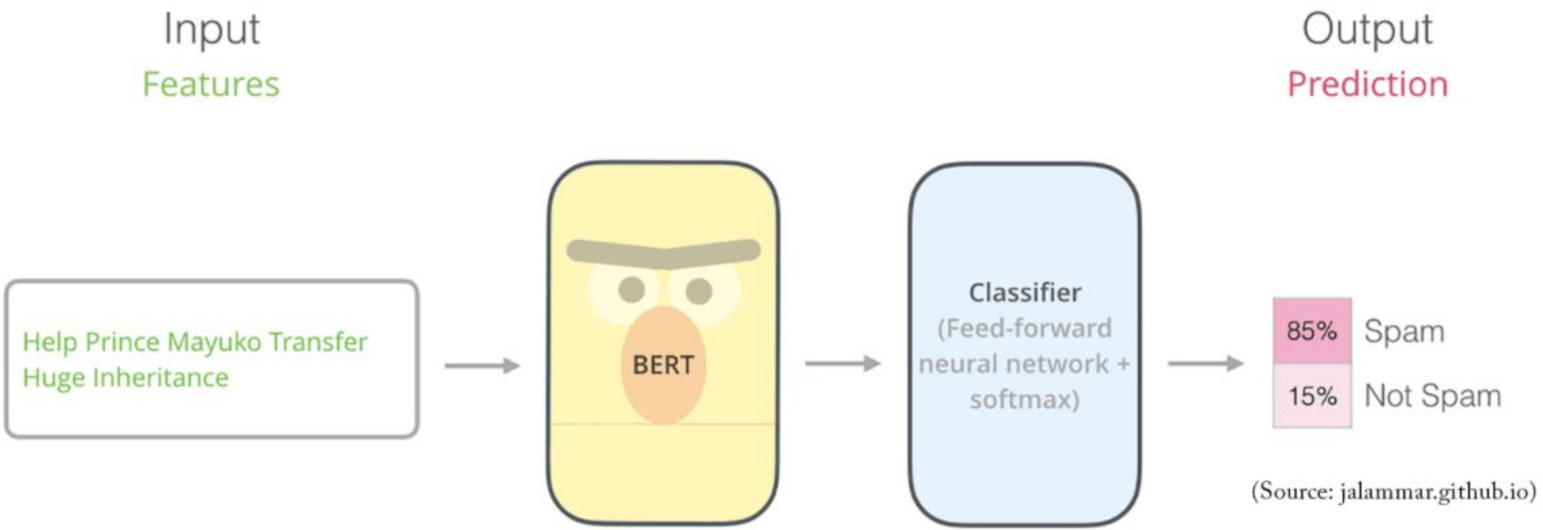
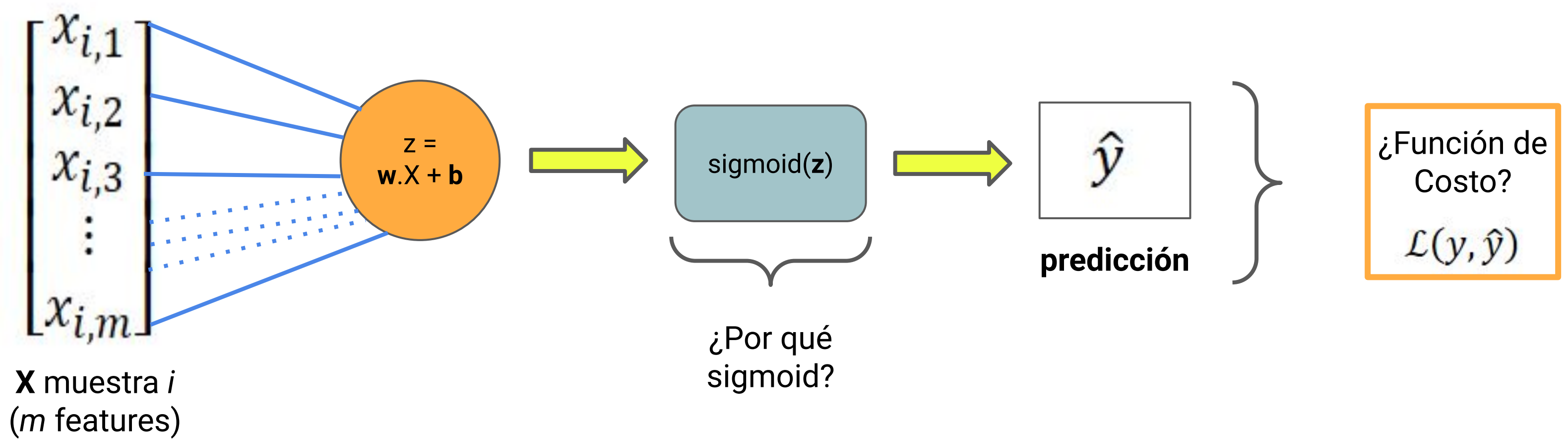


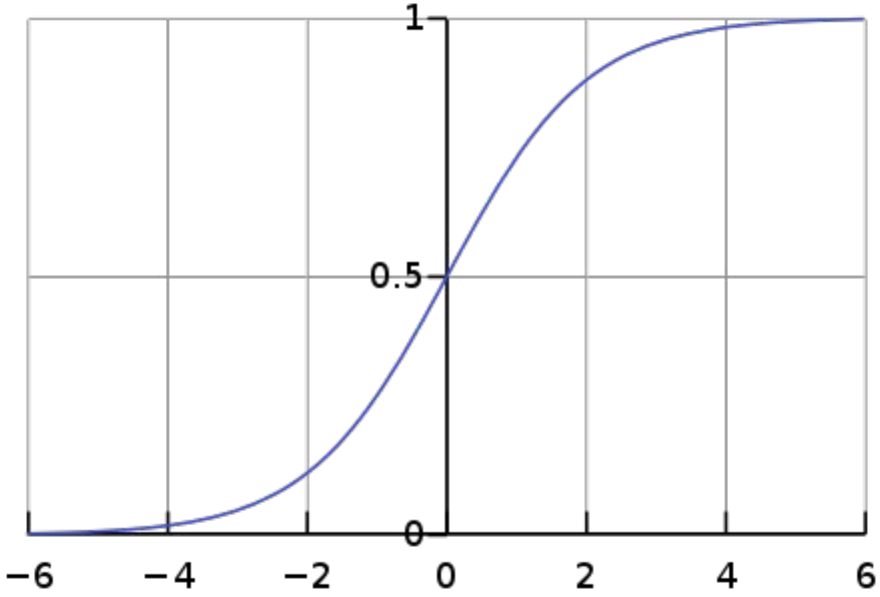
Imagen de: “Tutorial: Fine tuning BERT for Sentiment Analysis“

Regresión Logística



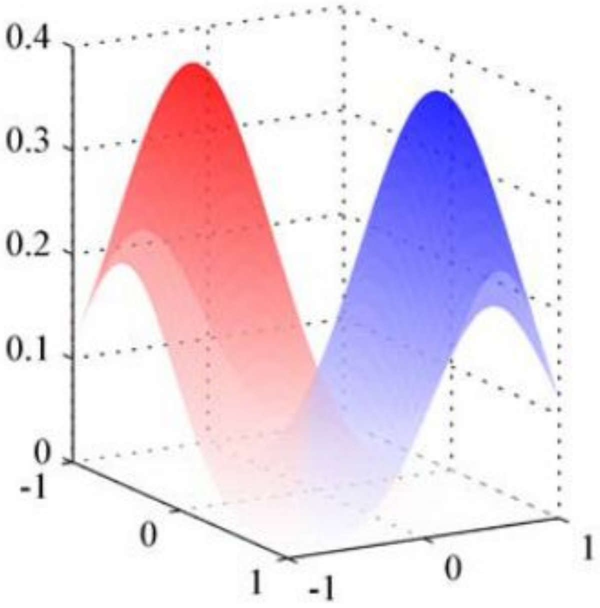
Logistic function

$$S(x) = \frac{1}{1+e^{-x}} = \frac{e^x}{e^x+1} = 1 - S(-x).$$

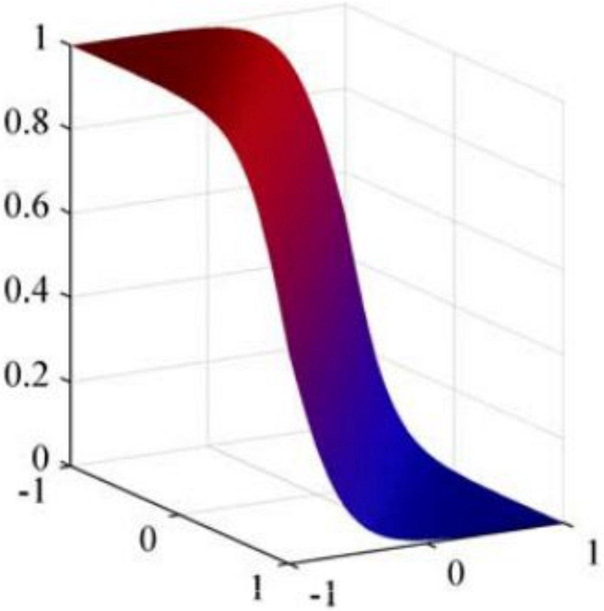


.UBAfiuba FACULTAD DE INGENIERÍA

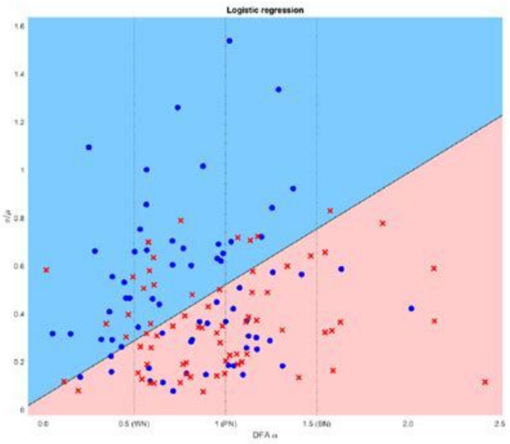
Regresión Logística



Class-conditional - P(x|Cn)

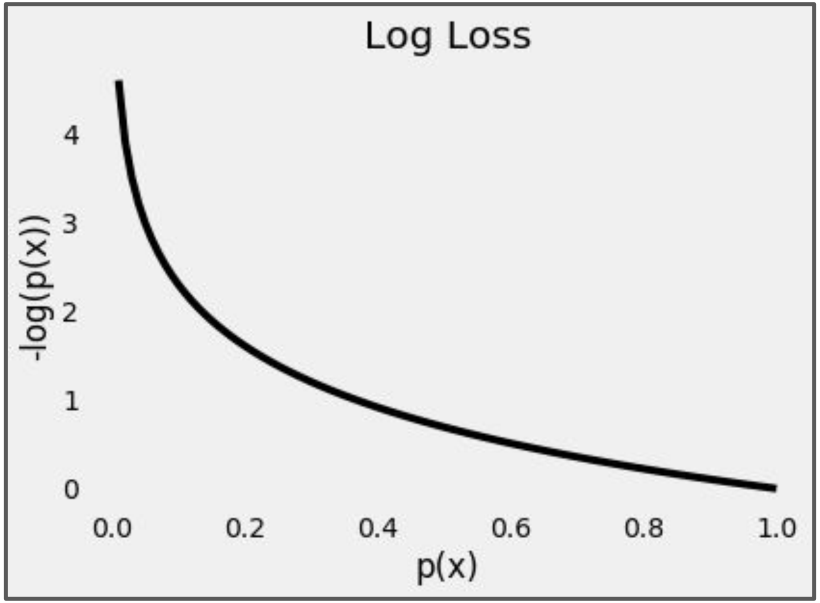


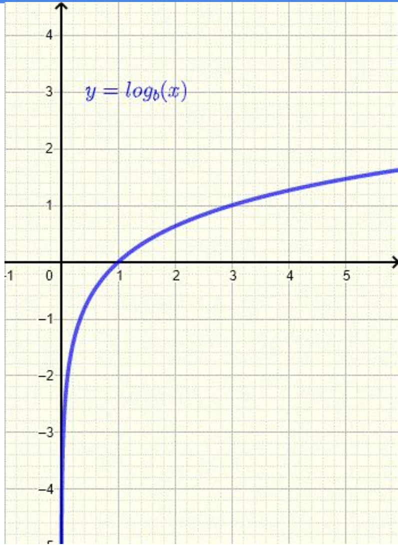
Posterior - P(Cn|x)



Función de costo - Binary cross entropy

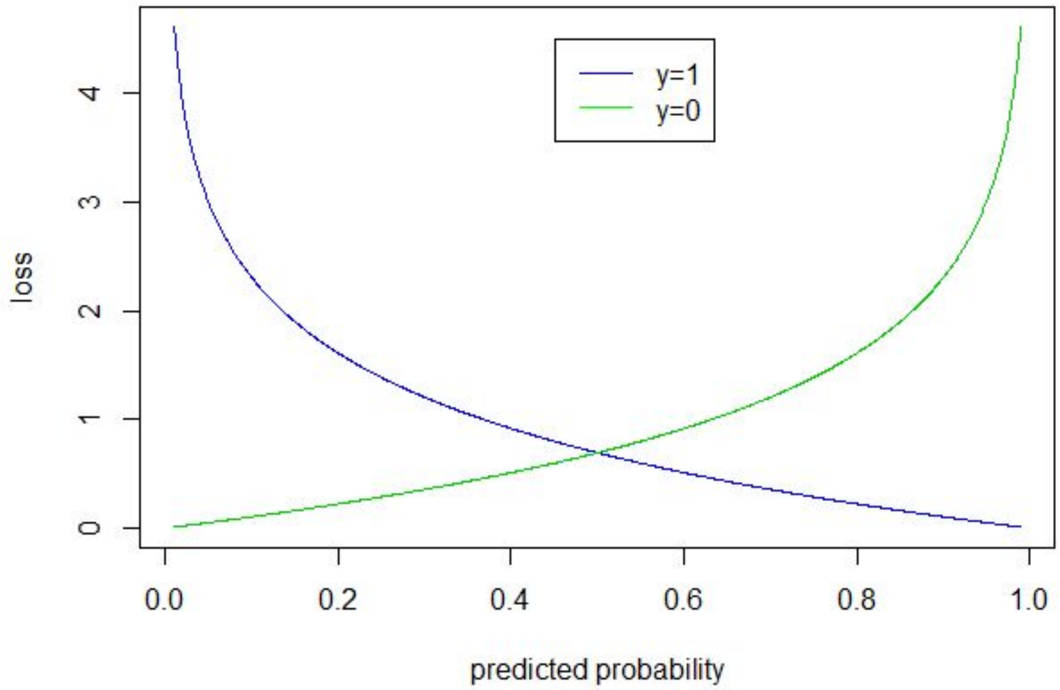
$$H_p(q) = -\frac{1}{N}\sum_{i=1}^{N} y_i \cdot log(p(y_i)) + (1 - y_i) \cdot log(1 - p(y_i))$$



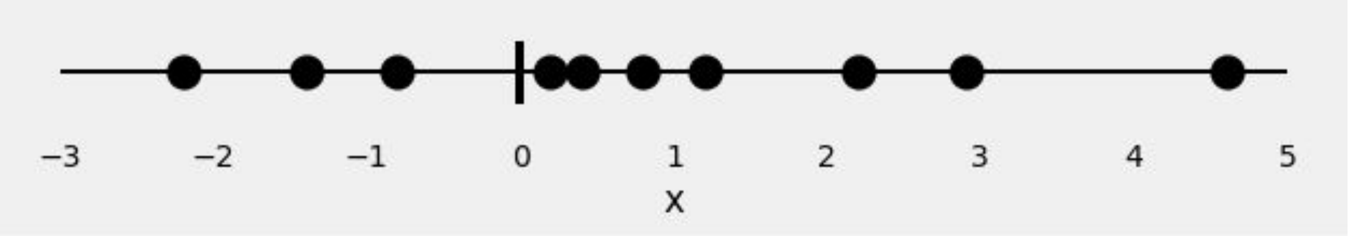


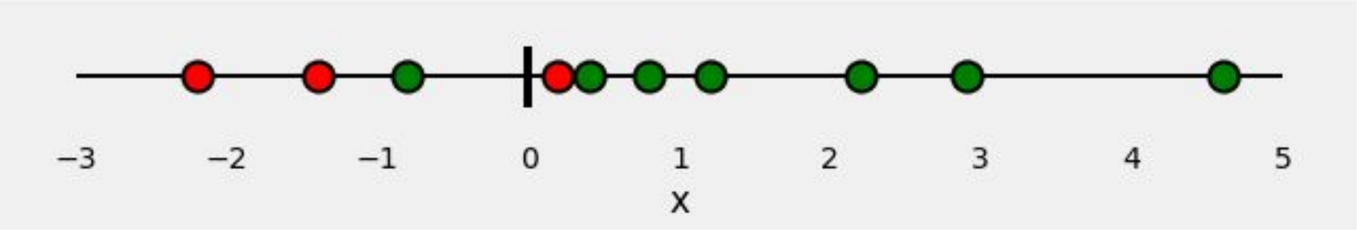
Función de costo - Binary cross entropy

$$H_p(q) = -\frac{1}{N}\sum_{i=1}^{N} \boxed{y_i \cdot log(p(y_i))} + \boxed{(1 - y_i) \cdot log(1 - p(y_i))}$$



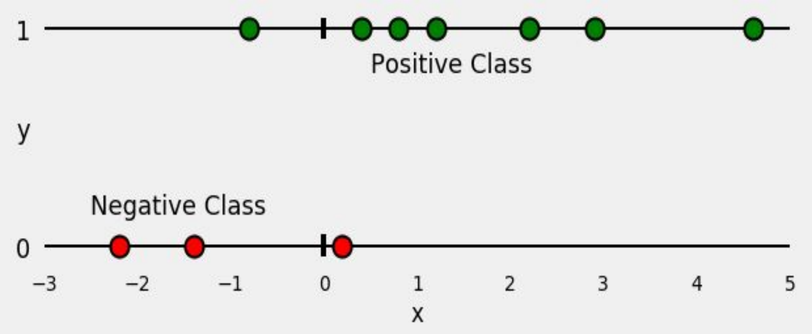
Regresión Logística

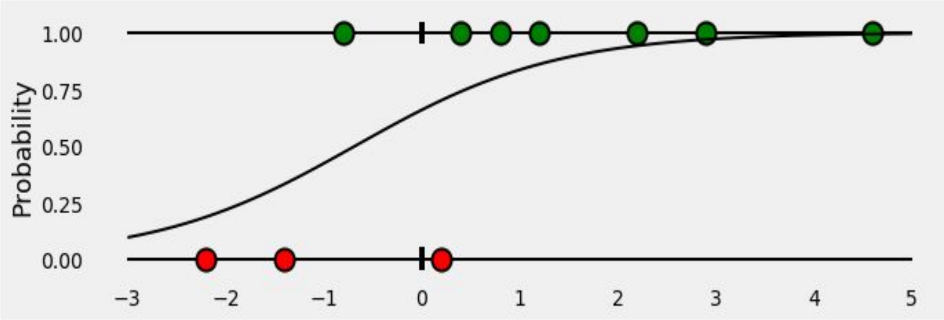




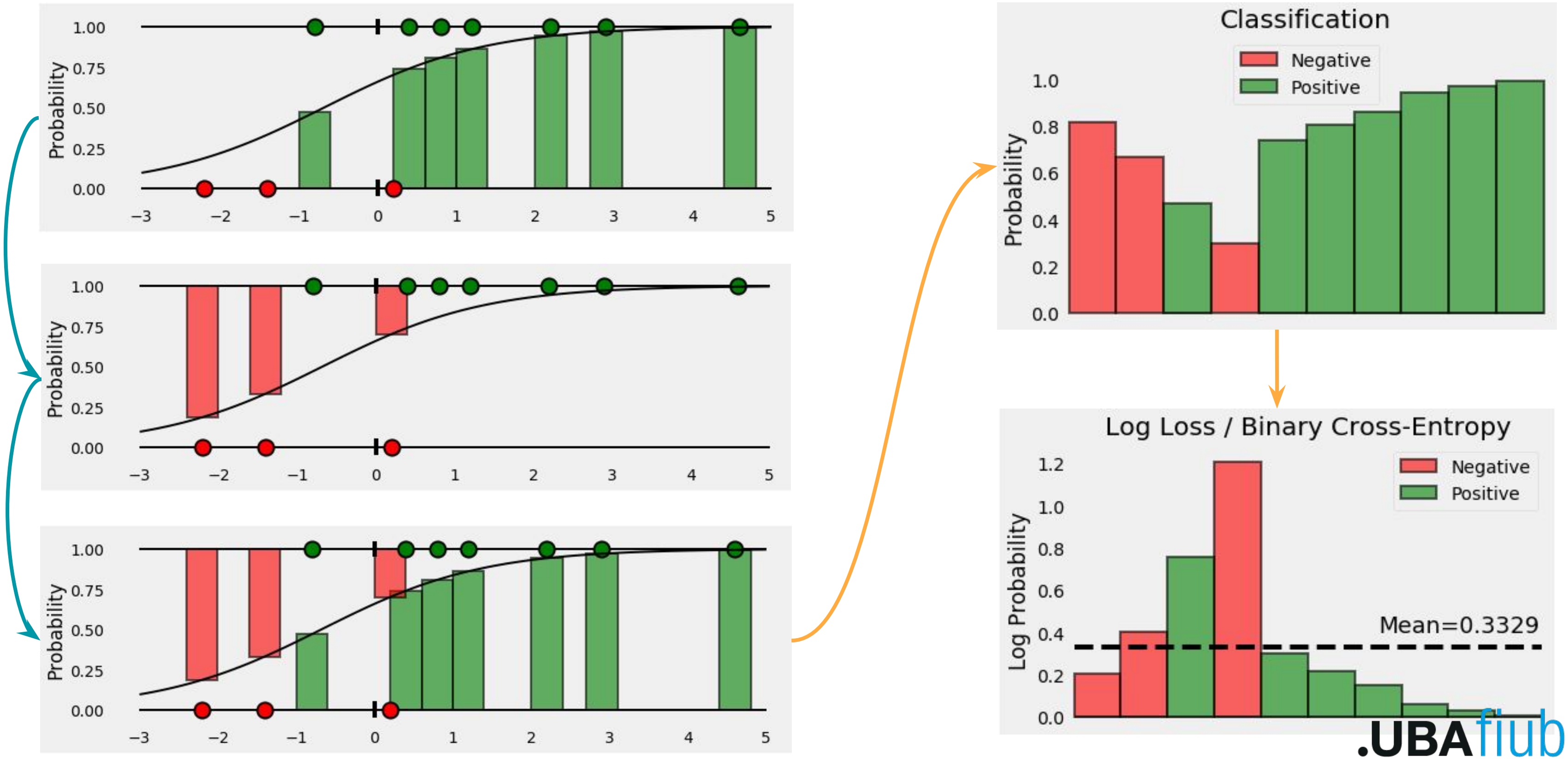
1: Verde, 0: Rojo

Regresión Logística





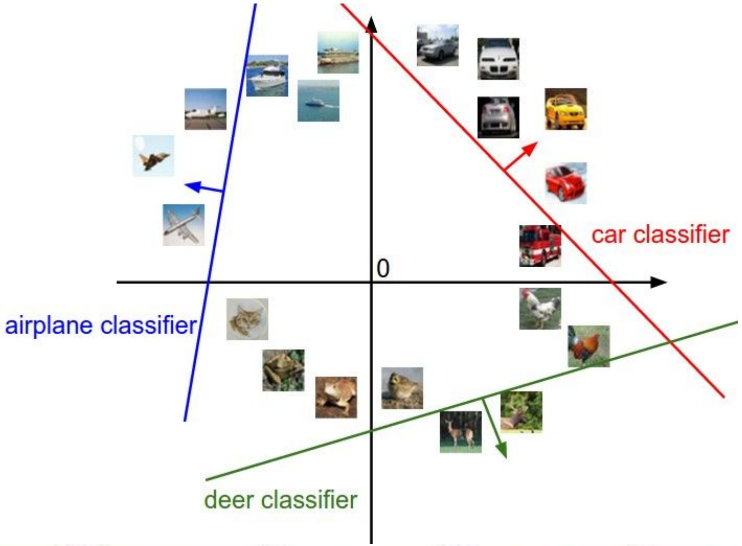
Regresión Logística

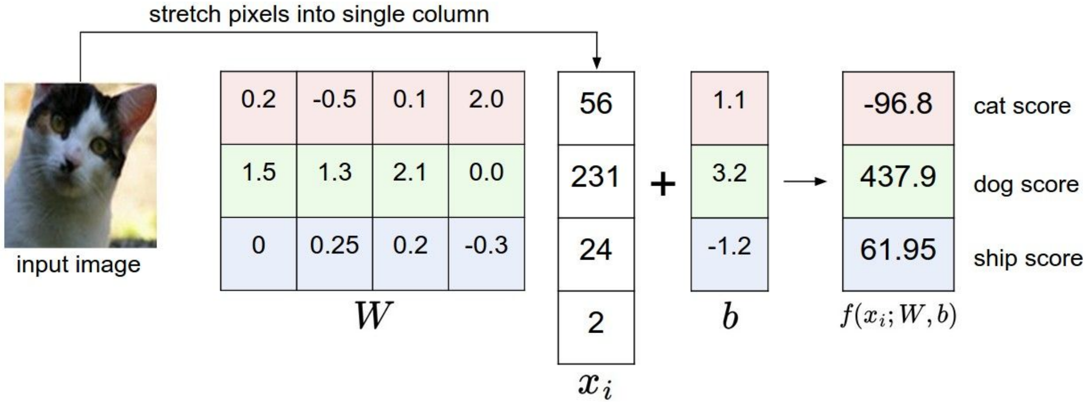


.UBAfiuba
FACULTAD DE INGENIERÍA

Clasificación multiclase

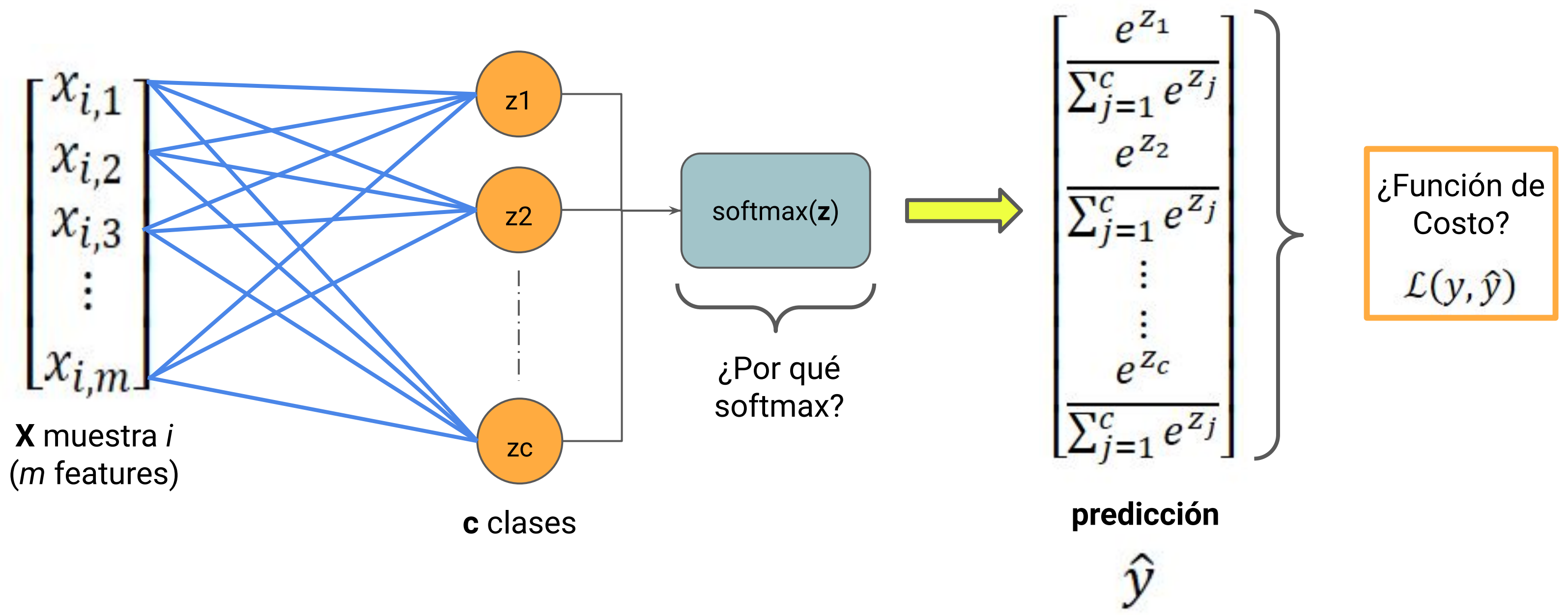
Clasificación Multiclase - Motivación





.UBA fiuba
FACULTAD DE INGENIERÍA

Softmax



$x_{i,1}$
$x_{i,2}$
$x_{i,3}$
$\vdots$
$x_{i,m}$
X muestra *i*
(*m* features)
z1
z2
zc
c clases
softmax(**z**)
¿Por qué softmax?
$\frac{e^{z_1}}{\sum_{j=1}^{c} e^{z_j}}$
$\frac{e^{z_2}}{\sum_{j=1}^{c} e^{z_j}}$
$\vdots$
$\frac{e^{z_c}}{\sum_{j=1}^{c} e^{z_j}}$
predicción
$\hat{y}$
¿Función de Costo?
$\mathcal{L}(y, \hat{y})$

Softmax

$$P(y_i \mid x_i; W) = \frac{e^{f_{y_i}}}{\sum_j e^{f_j}}$$

$$\underbrace{\frac{e^{f_{y_i}}}{\sum_j e^{f_j}} = \frac{Ce^{f_{y_i}}}{C\sum_j e^{f_j}} = \frac{e^{f_{y_i}+\log C}}{\sum_j e^{f_j+\log C}}}_{q(x)}$$

$$H(p,q) = -\sum_x p(x) \log q(x)$$

1 [Softmax Forma Gráfica]

2 [Softmax Visualización 3D]

Softmax

Derivación Softmax

$$p_i = \frac{e^{z_i}}{\sum_{j=1}^{C} e^{z_j}}$$

$$\frac{\partial p_i}{\partial z_k} = \frac{\partial \frac{e^{z_i}}{\sum_{j=1}^{C} e^{z_j}}}{\partial z_k}$$

$$\frac{\partial p_i}{\partial z_k} = p_i(\delta_{ik} - p_k) \qquad \delta ik = \begin{cases} 1, i = k \\ 0, i \neq k \end{cases}$$

Derivación Cross-Entropy

$$L = -\sum_i y_i log(p_i)$$

$$\frac{\partial L}{\partial z_i} = -\sum_j y_j \frac{\partial log(p_j)}{\partial z_i}$$

$$= -\sum_j y_j \frac{\partial log(p_j)}{\partial p_j} \times \frac{\partial p_j}{\partial z_i}$$

$$= -\sum_j y_j \frac{1}{p_j} \times \frac{\partial p_j}{\partial z_i}$$

$$\frac{\partial L}{\partial z_i} = -y_i(1-p_i) - \sum_{j \neq i} y_j \frac{1}{p_j}(-p_j.p_i)$$

$$= -y_i(1-p_i) + \sum_{j \neq i} y_j.p_i$$

$$= p_i\left(y_i + \sum_{j \neq i} y_j\right) - y_i$$

$$\frac{\partial L}{\partial z_i} = p_i - y_i$$

$$\frac{\partial z_i}{\partial W} = x_i$$

$$\frac{\partial L}{\partial W} = \sum_{i=1}^{N} (p_i - y_i)\, x_i$$

Usar gradiente descendente para actualizar W !!!

.UBA fiuba FACULTAD DE INGENIERÍA

Ejercicios

Ejercicio integrador

1. Implementar el algoritmo de regresión logística en NumPy.
2. Aplicar el modelo a un dataset de elección.
3. Comparar los resultados con Scikit-Learn .
4. Comparar los resultados agregando regularización.

Bibliografía

- The Elements of Statistical Learning | Trevor Hastie | Springer
- An Introduction to Statistical Learning | Gareth James | Springer
- Deep Learning | Ian Goodfellow | https://www.deeplearningbook.org/
- Mathematics for Machine Learning | Deisenroth, Faisal, Ong
- Artificial Intelligence, A Modern Approach | Stuart J. Russell, Peter Norvig
- Understanding binary cross-entropy: a visual explanation | Daniel Godoy
- Visual Information Theory | Link
- https://cs231n.github.io/
- Classification and Loss Evaluation-Softmax and Cross Entropy Loss | Paras Dahal