



13. Optimizacion del Entrenamiento Bootcamp **Visión Artificial para los ODS**

 M.Sc. Edwin Salcedo

 1 de Junio del 2022

 Zoom



UNIVERSIDAD
CATÓLICA
BOLIVIANA
LA PAZ





ML supervisado

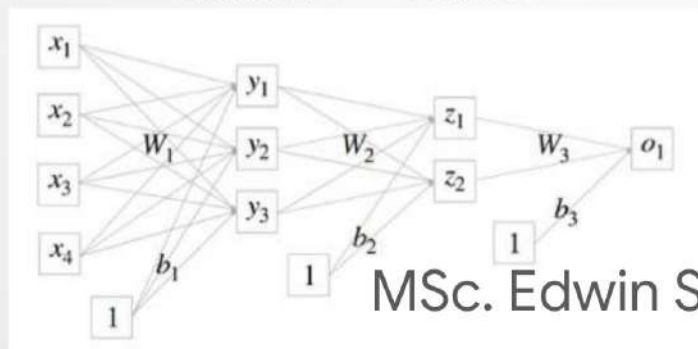
$$\hat{y}^{(i)} = f_{\theta}(x^{(i)})$$



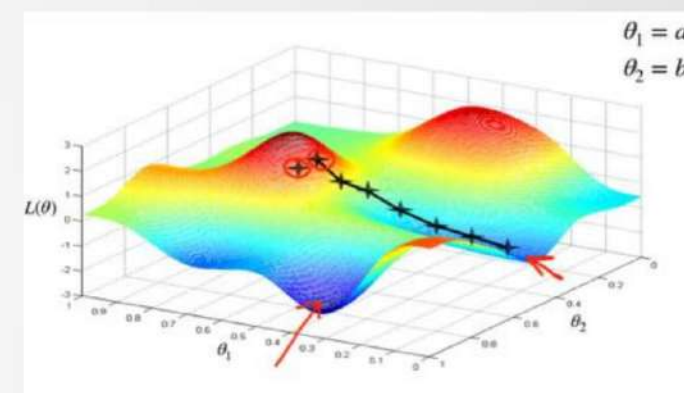
entrenamiento

pruebas

entrada Capa oculta 1 Capa oculta 2 salida



MSc. Edwin Salcedo



¡Gradient Descent al rescate!

$$\theta^{(i+1)} = \theta^{(i)} - \alpha \cdot \nabla L(\theta)$$

Funciones de Perdida

MSc. Edwin Salcedo

Funciones de costo



UNIVERSIDAD
CATÓLICA
BOLIVIANA
LA PAZ

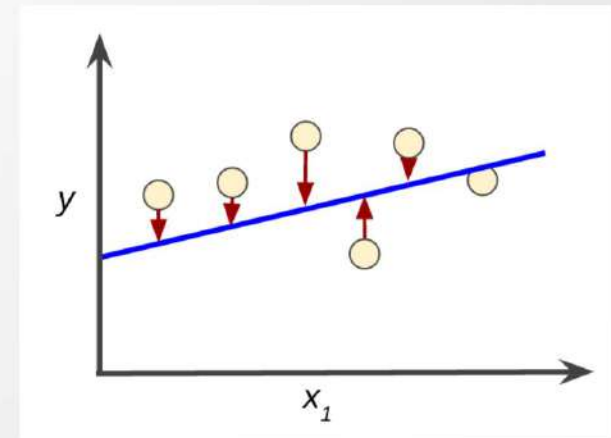
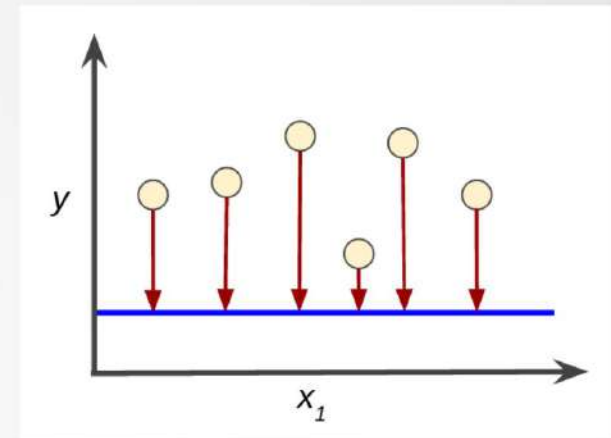
Perdida (Loss)

La pérdida (costo) es un número que indica qué tan mala fue la predicción del modelo en un solo ejemplo. Esta se puede encontrar con varias funciones, la más utilizada es la función Squared Loss, también conocida como L2 loss.

$$\text{Squared Loss} = (y_j - \hat{y}_j)^2$$

Donde:

- y_j es el valor real
- \hat{y}_j es el valor predicho



MSc. Edwin Salcedo

Funciones de costo para regresión



UNIVERSIDAD
CATÓLICA
BOLIVIANA
LA PAZ

Mean Absolute Error L_1 (MAE)

El error absoluto medio es una forma de encontrar el total de pérdidas y una buena opción si MSE no funciona para un problema específico. Ambos son buenos con tareas de regresión.

$$\text{Mean Absolute Error} = \frac{1}{N} \sum_{j=1}^N |y_j - \hat{y}_j|$$

Mean Squared Error L_2 (MSE)

El agregado de pérdidas generalmente se calcula mediante MSE porque facilita el cálculo del gradiente, mientras que el error absoluto medio requiere herramientas de programación lineal complicadas para calcular el gradiente.

$$\text{Mean Squared Error} = \frac{1}{N} \sum_{j=1}^N (y_j - \hat{y}_j)^2$$

MSc. Edwin Salcedo

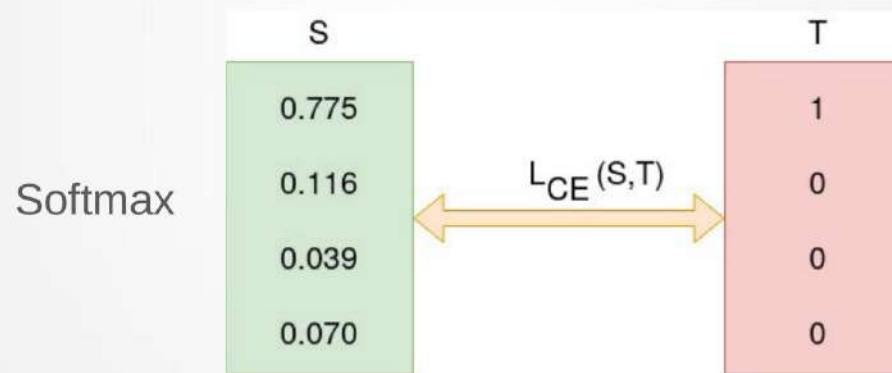
Funciones de pérdida para la clasificación



UNIVERSIDAD
CATÓLICA
BOLIVIANA
LA PAZ

Categorical Cross-Entropy

La entropía cruzada es la mejor función de pérdida para la clasificación multiclase y su propósito es tomar las probabilidades de salida (P) y medir la distancia de los valores de verdad.



Para el ejemplo anterior, el resultado deseado es [1,0,0,0] para el perro de la clase pero el modelo salidas [0,775, 0,116, 0,039, 0,070] .

MSc. Edwin Salcedo

Funciones de pérdida para la clasificación



UNIVERSIDAD
CATÓLICA
BOLIVIANA
LA PAZ

Categorical Cross-Entropy

$$L_{Cross-Entropy} = - \sum_{i=1}^n t_i * \log_2(p_i)$$

Donde:

- n es el numero de clases (etiquetas)
- t_i el la etiqueta real
- p_i es la probabilidad proveniente de Softmax para la i clase

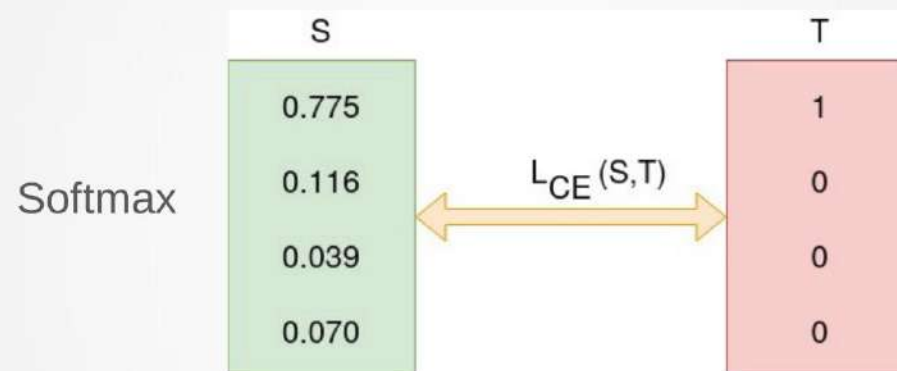
MSc. Edwin Salcedo

Funciones de pérdida para la clasificación



UNIVERSIDAD
CATÓLICA
BOLIVIANA
LA PAZ

Categorical Cross-Entropy



$$\begin{aligned} L_{CE} &= - \sum_{i=1} T_i \log(S_i) \\ &= - [1 \log_2(0.775) + 0 \log_2(0.126) + 0 \log_2(0.039) + 0 \log_2(0.070)] \\ &= - \log_2(0.775) \\ &= 0.3677 \end{aligned}$$

MSc. Edwin Salcedo

Funciones de pérdida para la clasificación



UNIVERSIDAD
CATÓLICA
BOLIVIANA
LA PAZ

Binary Cross-Entropy

$$L = -\frac{1}{N} \left[\sum_{j=1}^N [t_j * \log_2(p_j) + (1 - t_j) * \log(1 - p_j)] \right]$$

donde:

- N es el numero de muestras
- t_i es la etiqueta real, donde se toma en cuenta a los valores 0 o 1
- p_i es la salida de la función Softmax para la i clase

MSc. Edwin Salcedo

Funciones de Optimización

MSc. Edwin Salcedo



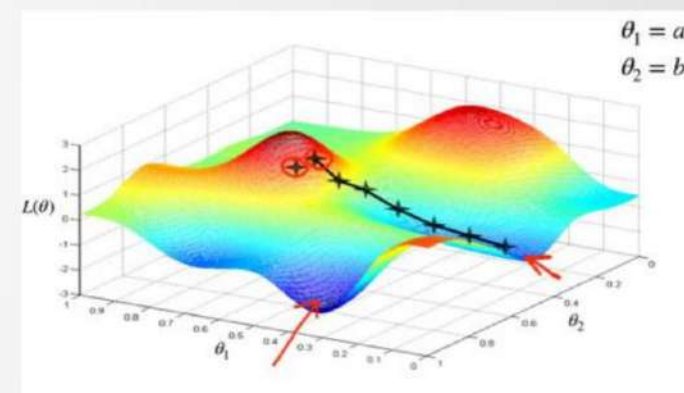
Descenso de la gradiente (Gradient Descent)

Los pasos del algoritmo son:

1. Definir los parámetros del modelo aleatoriamente.
2. **While** nos ubiquemos en una pendiente:
 - A. Encontrar el error (costo) de la función de pérdida.
 - B. Calcular el error a propagar para cada variable usando derivadas parciales del error encontrado con respecto a cada parámetro.
 - C. Calcular el nuevo valor para cada parámetro del modelo, usando la siguiente función, donde α será el grado de aprendizaje (su valor debería rondar entre 0.1 y 0.01).

$$\theta^{(i+1)} = \theta^{(i)} - \alpha \cdot \nabla L(\theta)$$

MSc. Edwin Salcedo





Descenso de la gradiente (Gradient Descent)

Cost function

$$J(\theta) = \frac{1}{2m} \sum_{i=1}^m \left(\underbrace{h_{\theta}(x^{(i)})}_{\text{Predicted value}}, \underbrace{y^{(i)}}_{\text{True value}} \right)^2$$

Square Error of data i

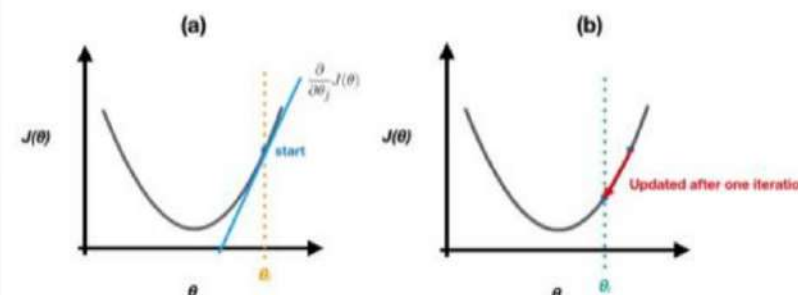
Mean Error Square

Problem Type	Output Type	Final Activation Function	Loss Function
Regression	Numerical value	Linear	Mean Squared Error (MSE)
Classification	Binary outcome	Sigmoid	Binary Cross Entropy
Classification	Single label, multiple classes	Softmax	Cross Entropy
Classification	Multiple labels, multiple classes	Sigmoid	Binary Cross Entropy

Repeat until converge {

$$\theta_j := \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta)$$

} where j represents the feature index number.



Back-propagation



UNIVERSIDAD
CATÓLICA
BOLIVIANA
LA PAZ

Error

$$E(W) = -\frac{1}{m} \sum_{i=1}^m y_i \ln(\hat{y}_i) + (1 - y_i) \ln(1 - \hat{y}_i)$$

Error a propagar

$$\nabla E = \begin{pmatrix} \frac{\partial E}{\partial W_{11}^{(1)}} & \frac{\partial E}{\partial W_{12}^{(1)}} & \frac{\partial E}{\partial W_{11}^{(2)}} \\ \frac{\partial E}{\partial W_{21}^{(1)}} & \frac{\partial E}{\partial W_{22}^{(1)}} & \frac{\partial E}{\partial W_{21}^{(2)}} \\ \frac{\partial E}{\partial W_{31}^{(1)}} & \frac{\partial E}{\partial W_{32}^{(1)}} & \frac{\partial E}{\partial W_{31}^{(2)}} \end{pmatrix}$$

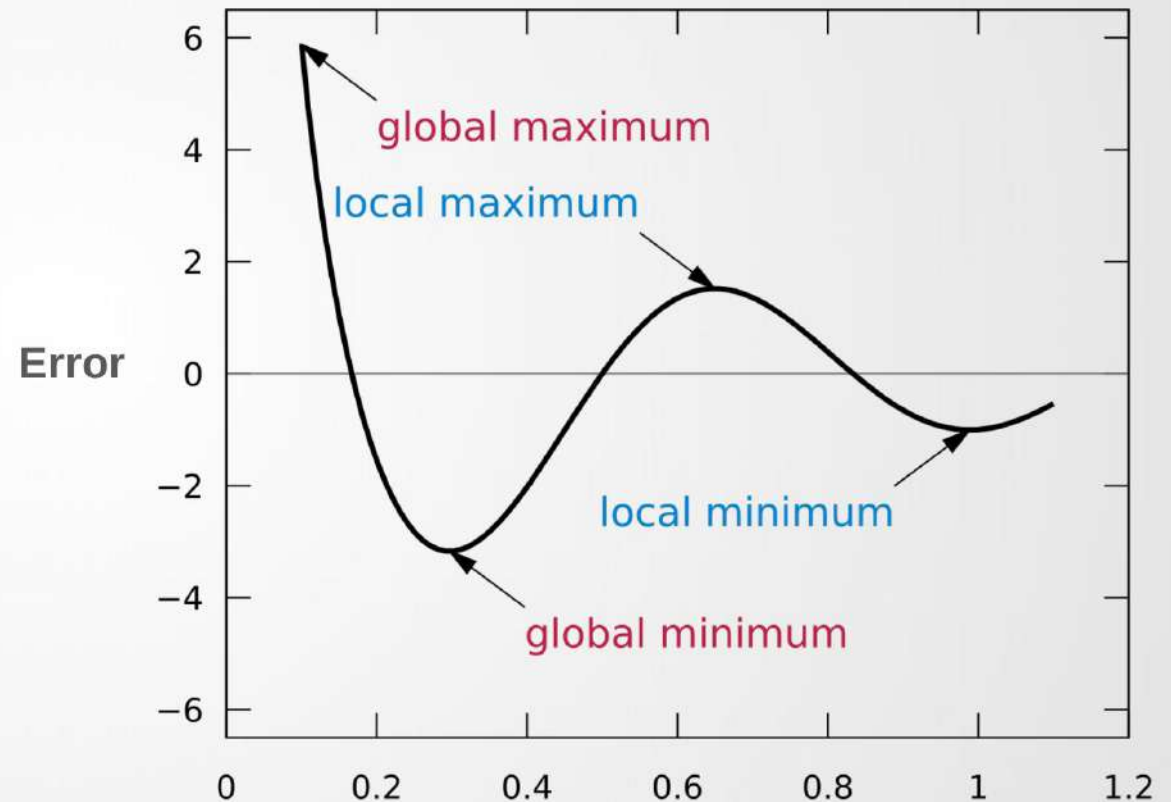
MSc. Edwin Salcedo

Local Minimum



UNIVERSIDAD
CATÓLICA
BOLIVIANA
LA PAZ

Nuestro Gradient Descent también puede tener problemas, por ejemplo, los mínimos locales son bastante comunes y pueden hacernos pensar que hemos llegado a un buen modelo.

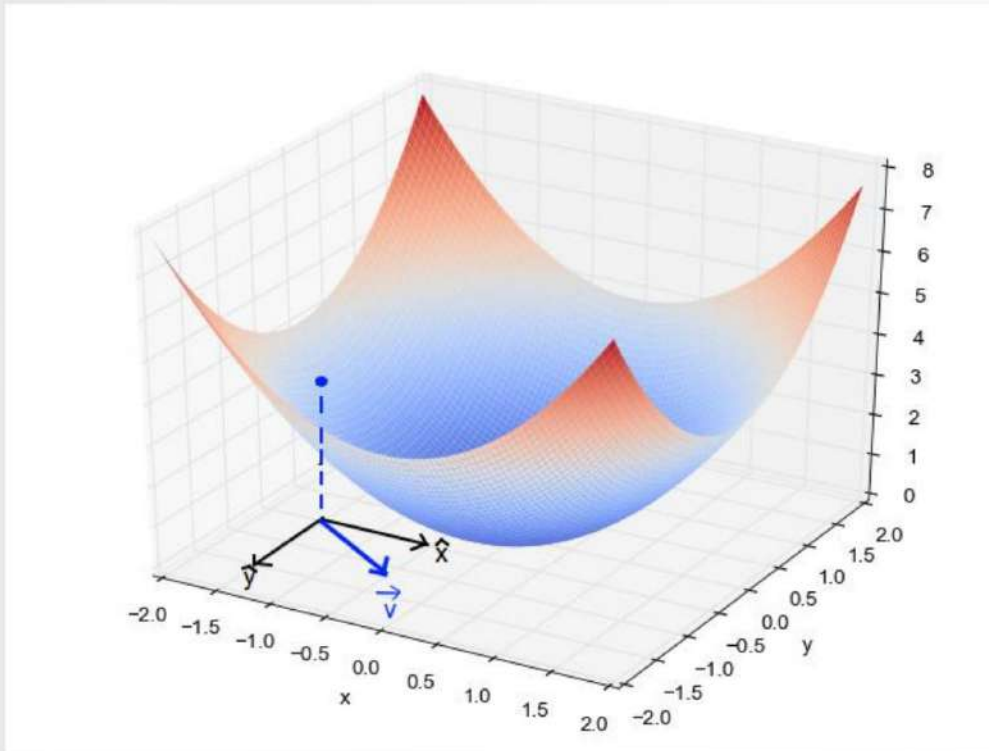


MSc. Edwin Salcedo

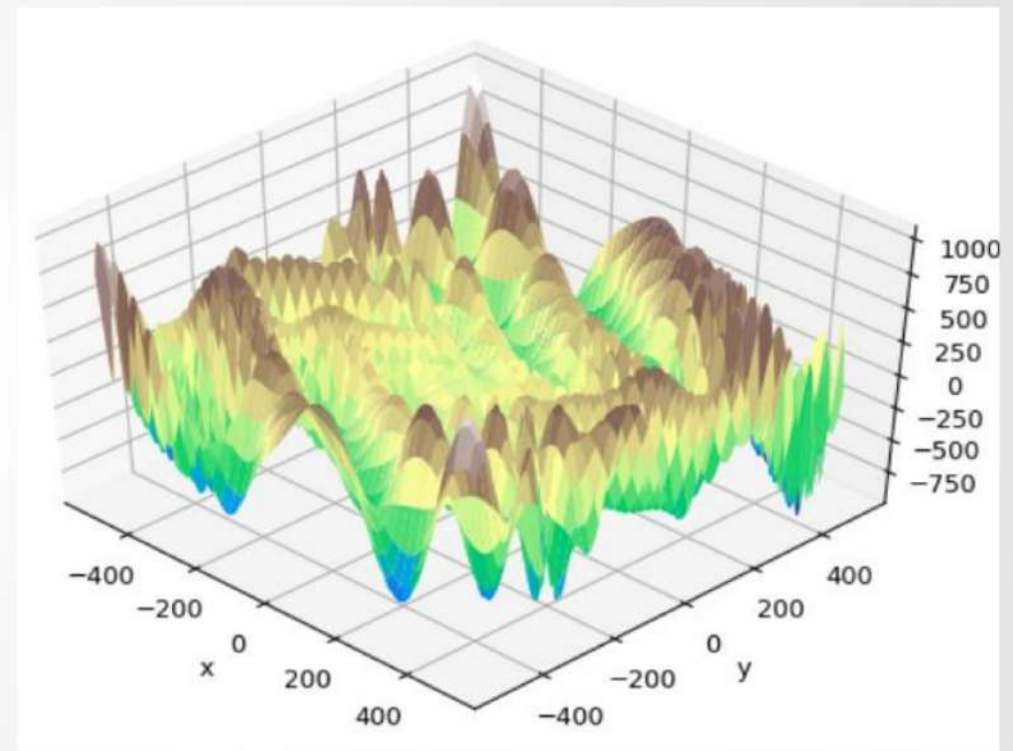
Gradient Descent



UNIVERSIDAD
CATÓLICA
BOLIVIANA
LA PAZ



Ejemplo perfecto



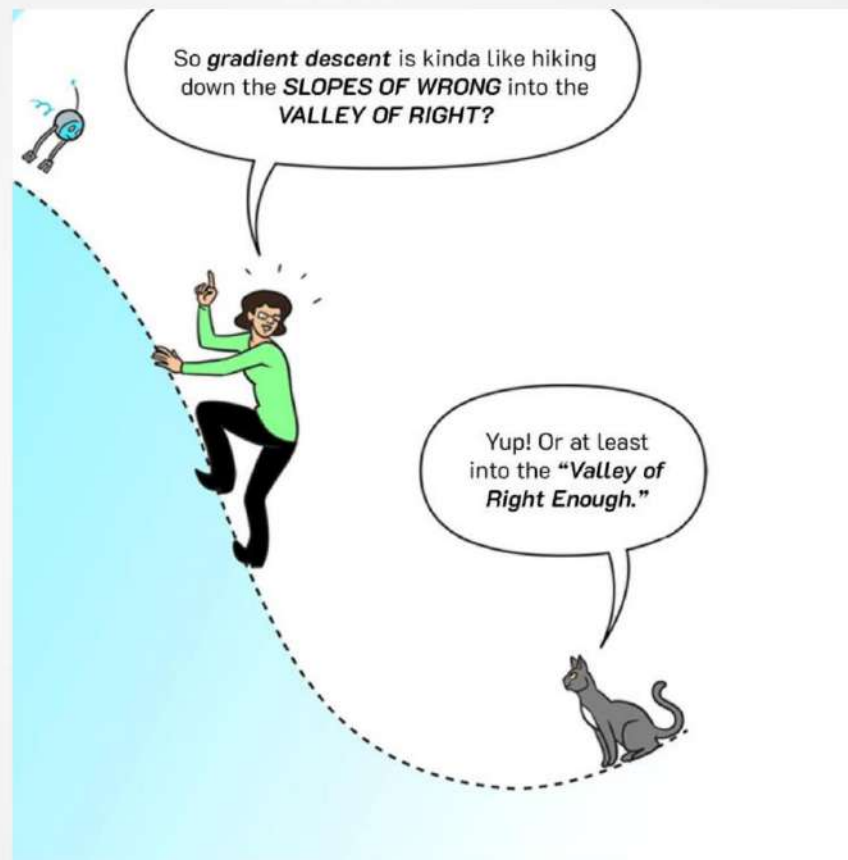
Vida real

MSc. Edwin Salcedo

Gradient Descent



UNIVERSIDAD
CATÓLICA
BOLIVIANA
LA PAZ



<https://cloud.google.com/solutions/ai/ml-comic-1>

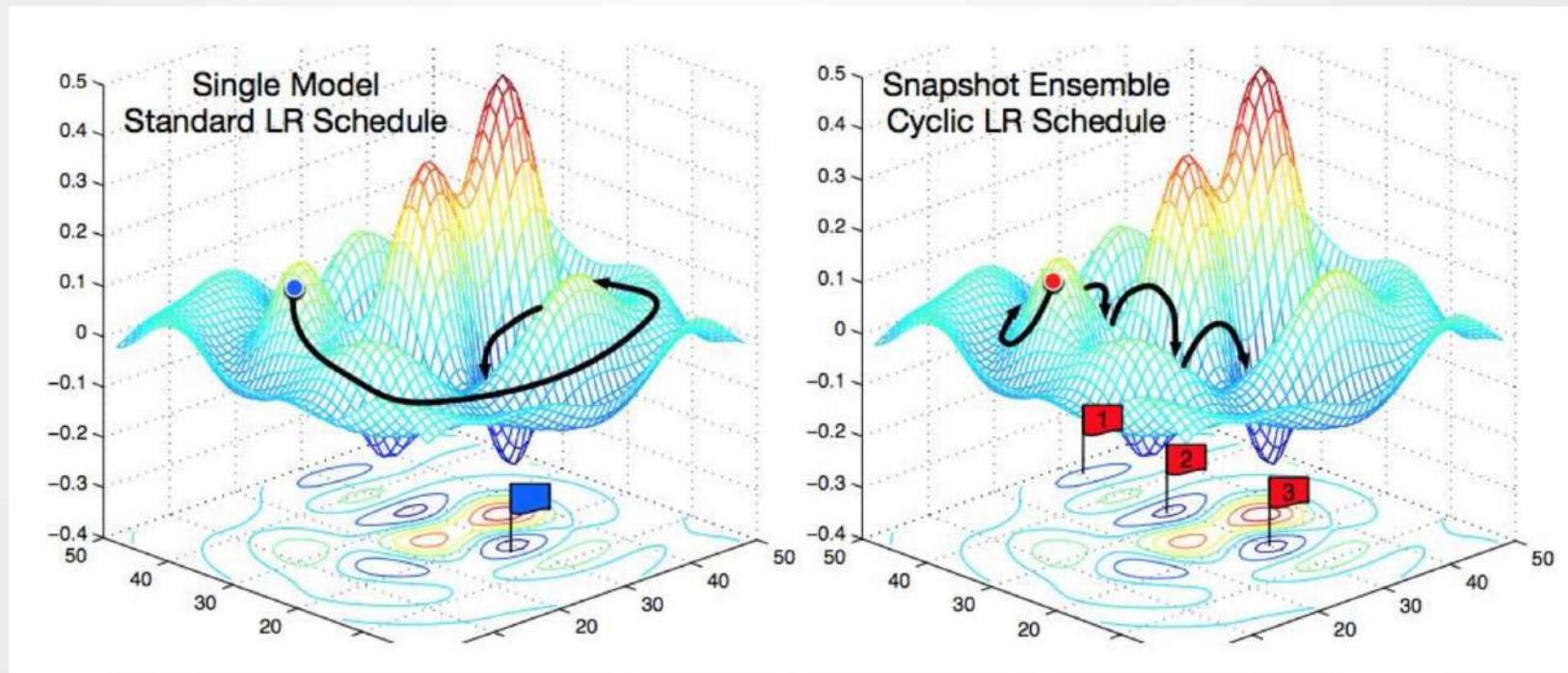
MSc. Edwin Salcedo

Variantes del algoritmo Gradient Descent



UNIVERSIDAD
CATÓLICA
BOLIVIANA
LA PAZ

Diferentes estrategias para el descenso de gradiente pueden aliviar el enorme procesamiento informático necesario para encontrar una solución óptima.



MSc. Edwin Salcedo

A Newbie's Guide to Stochastic Gradient Descent With Restarts by Chi-Peng Wang on Medium

Variantes del algoritmo Gradient Descent



UNIVERSIDAD
CATÓLICA
BOLIVIANA
LA PAZ

Batch Gradient Descent

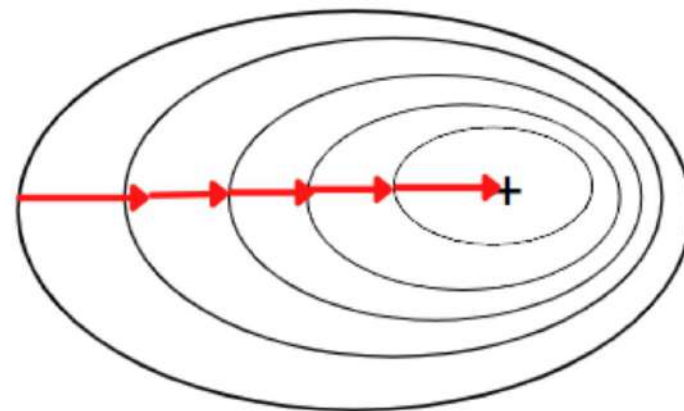
Ventajas

- Menos actualizaciones del modelo significan que esta variante es mas eficiente desde el puntos de vista computacional.
- La frecuencia de actualización reducida da como resultado un gradiente de error más estable, lo que conlleva una convergencia más estable en algunos problemas.

Desventajas

- Comúnmente, requiere todo el conjunto de datos de entrenamiento en la memoria y disponible para el algoritmo.
- Es propenso a encontrar mínimos locales.

Batch Gradient Descent



Calcula el error para cada ejemplo en el conjunto de datos de entrenamiento, pero solo actualiza el modelo después de que se hayan evaluado todos los ejemplos de entrenamiento.

MSc. Edwin Salcedo

Variantes del algoritmo Gradient Descent



UNIVERSIDAD
CATÓLICA
BOLIVIANA
LA PAZ

Mini-Batch Gradient Descent

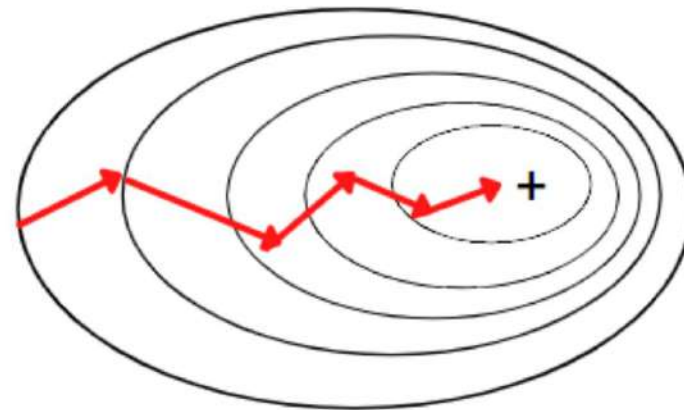
Ventajas

- La frecuencia de actualización del modelo es mayor que la anterior estrategia, lo que permite una convergencia más robusta, evitando mínimos locales.
- El procesamiento por lotes permite no tener todos los datos de entrenamiento en la memoria como la anterior estrategia.

Desventajas

- El mini-batch requiere la configuración de un hiperparámetro adicional de "tamaño de minibatch", el cual también requiere tuning.

Mini-Batch Gradient Descent



Calcula el error y actualiza el modelo por cada batch procesado.

MSc. Edwin Salcedo

Variantes del algoritmo Gradient Descent



UNIVERSIDAD
CATÓLICA
BOLIVIANA
LA PAZ

Stochastic Gradient Descent

Ventajas

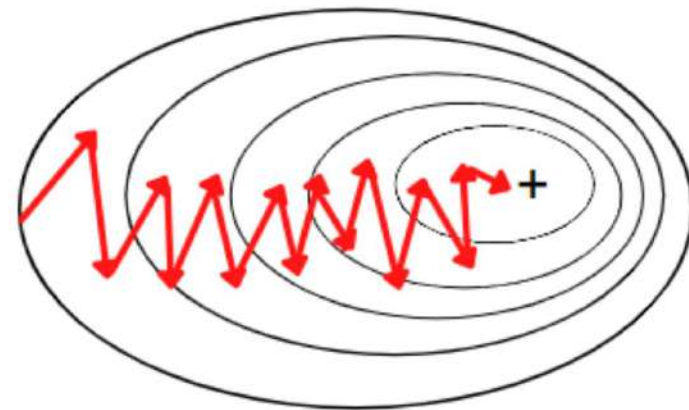
- Las actualizaciones frecuentes permiten encontrar una solución óptima en menor tiempo.
- La aleatoriedad del modelo permite evitar mínimos locales.

Desventajas

- Actualizar el modelo tan frecuentemente requiere de mas recursos computacionales.
- La aleatoriedad también podría causar que el modelo tenga sobresaltos de error.
- La aleatoriedad puede causar que el error y la gradiente no pueda establecerse en un mínimo global.

MSc. Edwin Salcedo

Stochastic Gradient Descent



Calcula el error y actualiza el modelo para cada ejemplo en el conjunto de datos de entrenamiento. La selección es aleatoria.

Momentum



UNIVERSIDAD
CATÓLICA
BOLIVIANA
LA PAZ

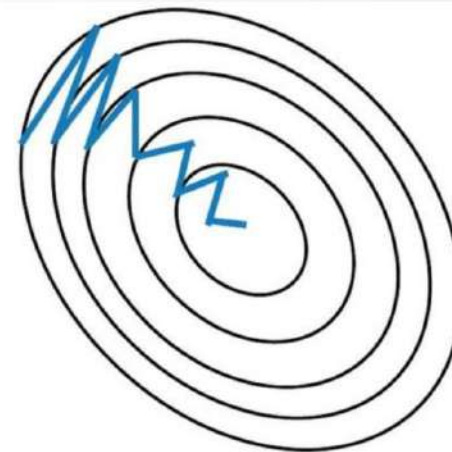
Paso: Promedio de los pasos anteriores

β : momentum

$$\text{Paso}(n) \rightarrow \text{Paso}(n) + \beta * \text{Paso}(n - 1) + \beta^2 * \text{Paso}(n - 2) + \dots$$



Stochastic Gradient
Descent **without**
Momentum



Stochastic Gradient
Descent **with**
Momentum

MSc. Edwin Salcedo

Adaptive Moment Estimation (ADAM)



UNIVERSIDAD
CATÓLICA
BOLIVIANA
LA PAZ

For each Parameter w^j

(j subscript dropped for clarity)

$$\nu_t = \beta_1 * \nu_{t-1} + (1 - \beta_1) * g_t$$

$$s_t = \beta_2 * s_{t-1} + (1 - \beta_2) * g_t^2$$

$$\Delta \omega_t = -\eta \frac{\nu_t}{\sqrt{s_t + \epsilon}} * g_t$$

$$\omega_{t+1} = \omega_t + \Delta \omega_t$$

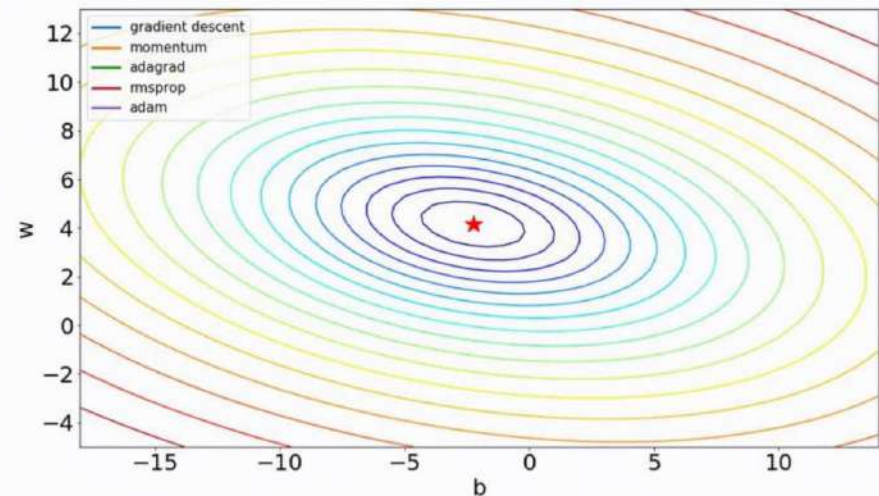
η : Initial Learning rate

g_t : Gradient at time t along ω^j

ν_t : Exponential Average of gradients along ω_j

s_t : Exponential Average of squares of gradients along ω_j

β_1, β_2 : Hyperparameters

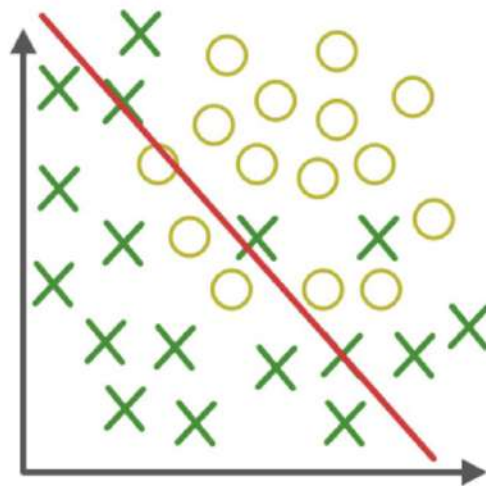


MSc. Edwin Salcedo

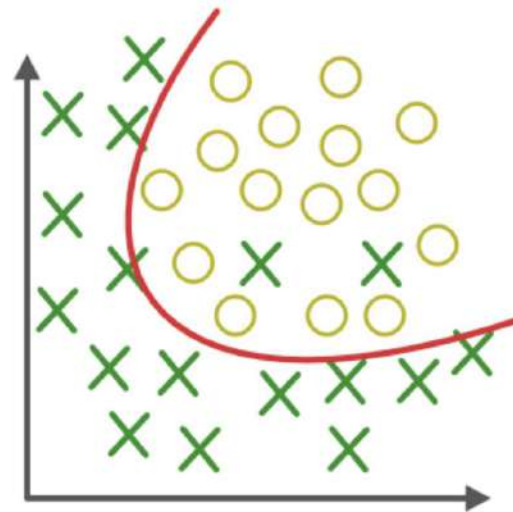
Ajuste apropiado



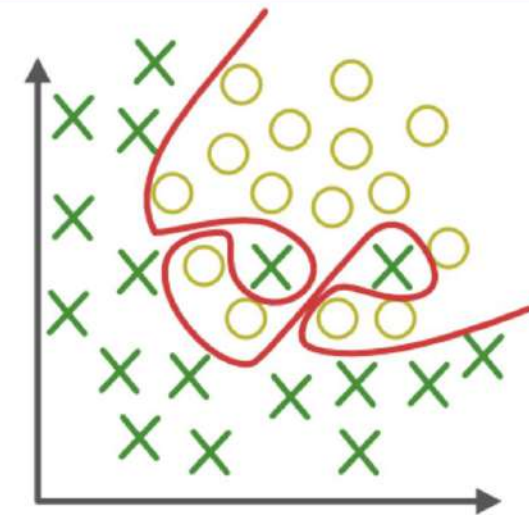
UNIVERSIDAD
CATÓLICA
BOLIVIANA
LA PAZ



Under-fitting
(too simple to
explain the variance)



Appropriate-fitting

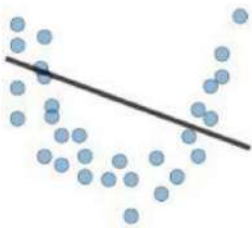


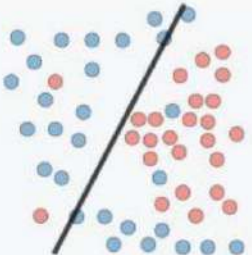
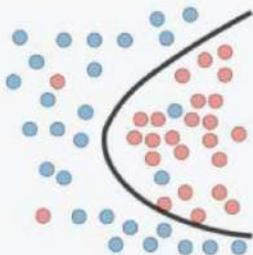
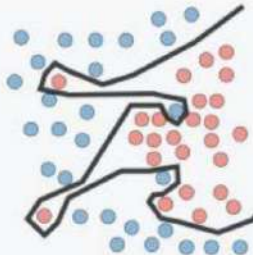

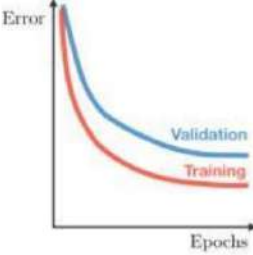
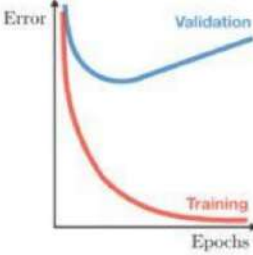


Over-fitting
(forcefitting--too
good to be true)



MSc. Edwin Salcedo

¿Como resolver problemas de entrenamiento?

	Underfitting	Just right	Overfitting
Symptoms	<ul style="list-style-type: none"> • High training error • Training error close to test error • High bias 	<ul style="list-style-type: none"> • Training error slightly lower than test error 	<ul style="list-style-type: none"> • Very low training error • Training error much lower than test error • High variance
Regression illustration			
Classification illustration			
Deep learning illustration			
Possible remedies	<ul style="list-style-type: none"> • Complexify model • Add more features • Train longer 		<ul style="list-style-type: none"> • Perform regularization • Get more data

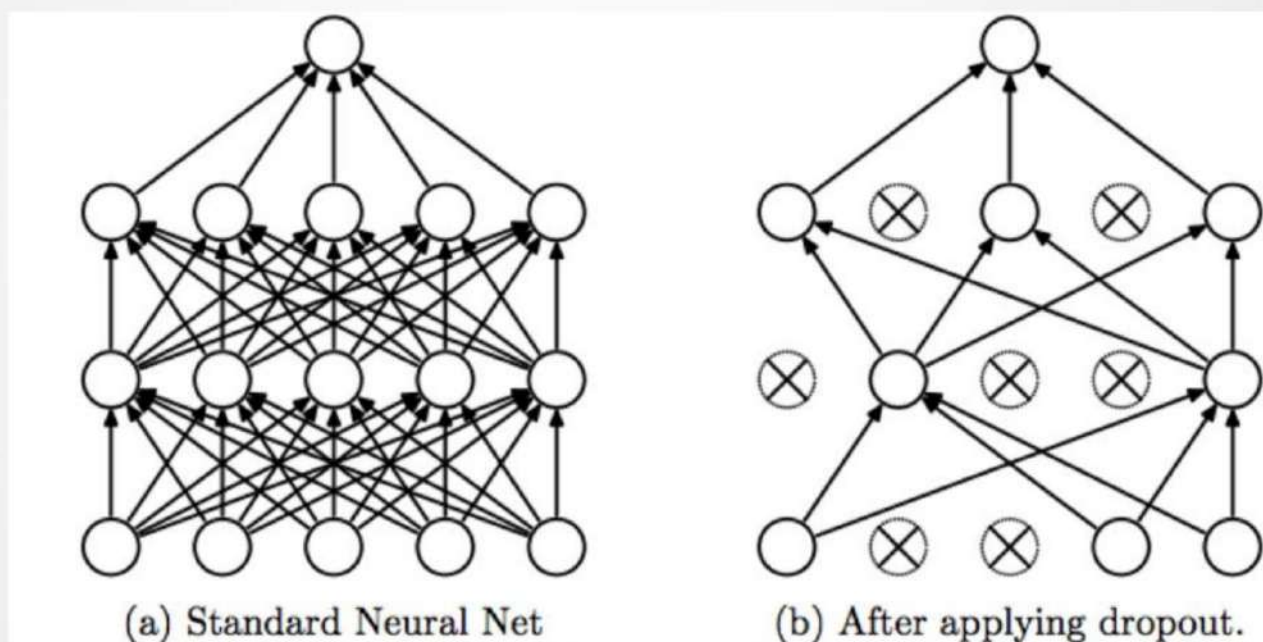
MSc. Edwin Salcedo

Dropout



UNIVERSIDAD
CATÓLICA
BOLIVIANA
LA PAZ

Dropout ayuda a apagar algunos nodos al azar en una capa para que podamos evitar el sobreajuste.

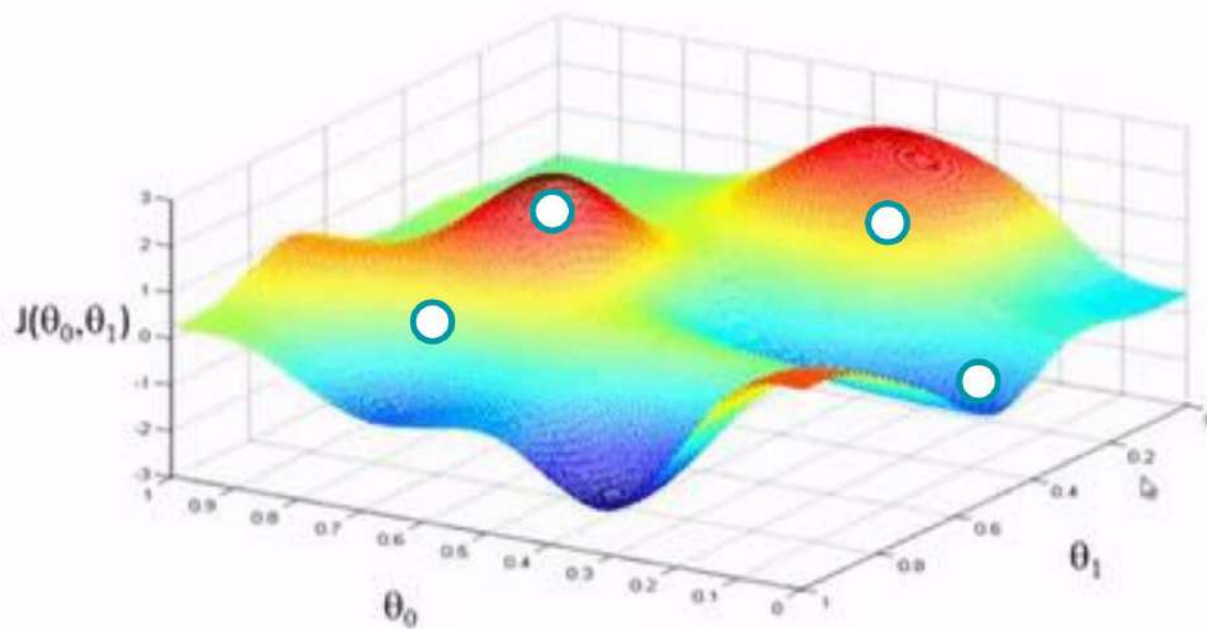


MSc. Edwin Salcedo

Random restart



UNIVERSIDAD
CATÓLICA
BOLIVIANA
LA PAZ



Andrew Ng

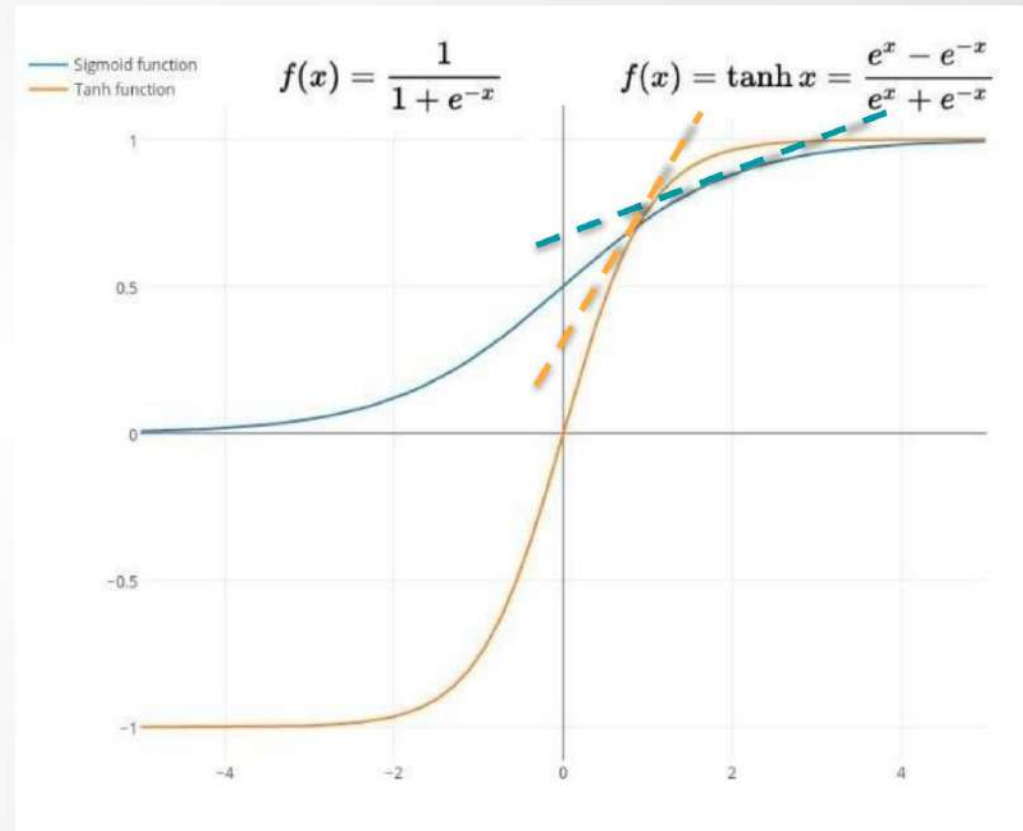
MSc. Edwin Salcedo

Vanishing gradient



UNIVERSIDAD
CATÓLICA
BOLIVIANA
LA PAZ

El fuga de gradiente (Vanishing gradient) ocurre cuando nuestra función de activación no es la más adecuada para una determinada capa o tarea.

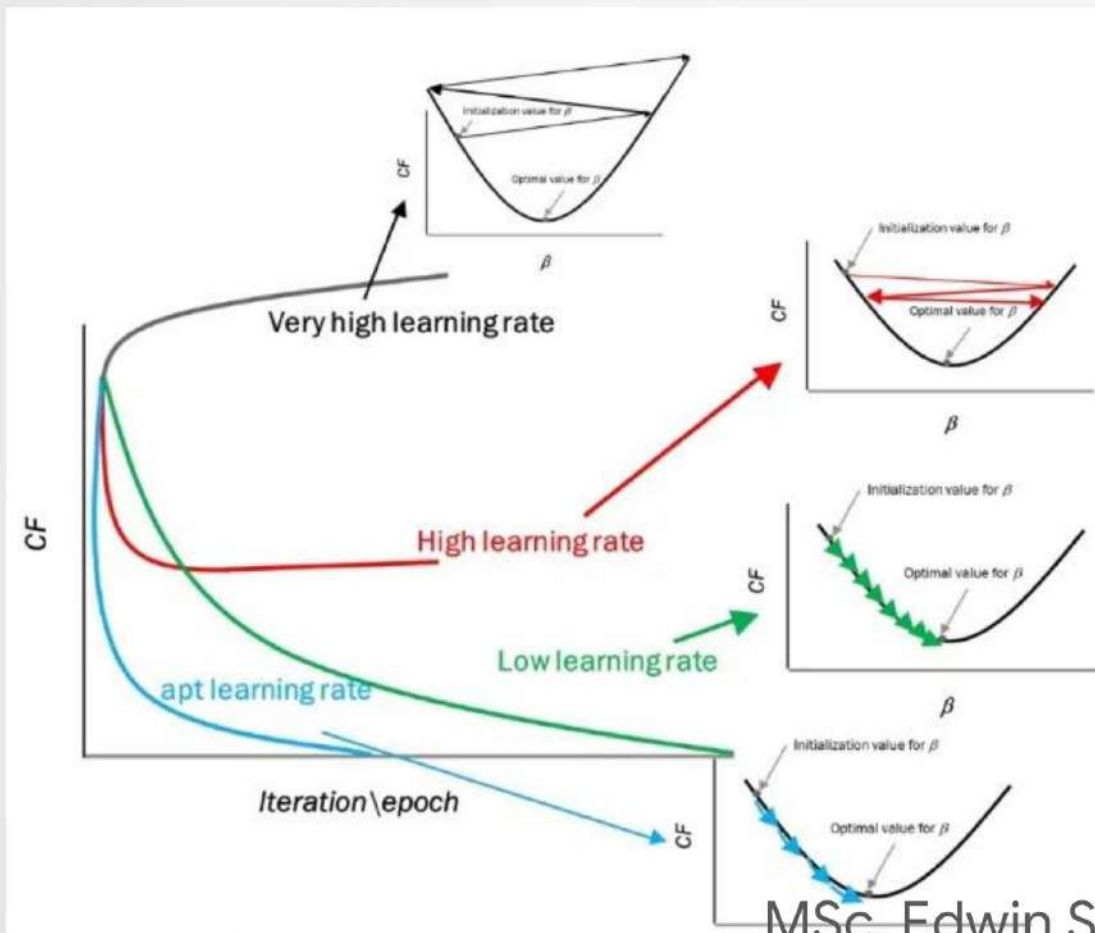


MSc. Edwin Salcedo

Efectos del Learning Rate



UNIVERSIDAD
CATÓLICA
BOLIVIANA
LA PAZ



MSc. Edwin Salcedo

Understanding learning rate in machine learning, Medium

Learning Rate Schedules: Step Decay

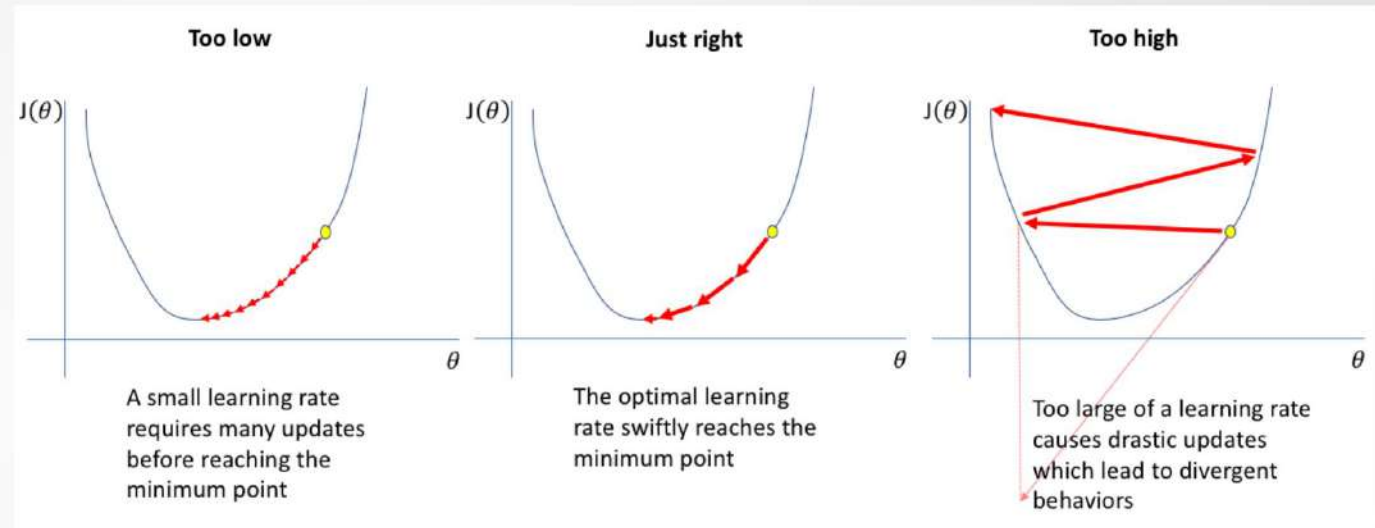


UNIVERSIDAD
CATÓLICA
BOLIVIANA
LA PAZ

Actualización de pesos:

$$w_x = w_x - \alpha * \left(\frac{\partial \text{LossFunction}}{\partial w_x} \right)$$

Donde α es el grado de aprendizaje.



Regla:

- Si estamos en una pendiente: realizar pasos largos
- Si estamos en un lugar plano: realizar pasos cortos

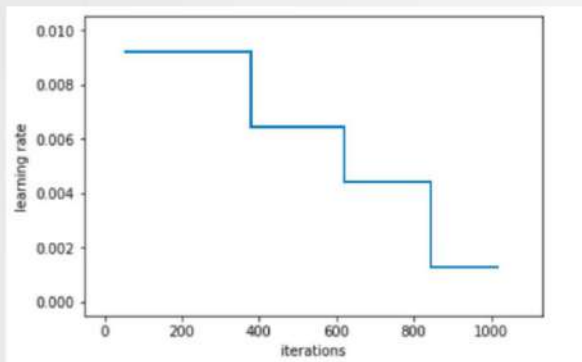
Reading: Learning Rate Schedule in Practice: an example with Keras and TensorFlow 2.0 on Medium

MSc. Edwin Salcedo

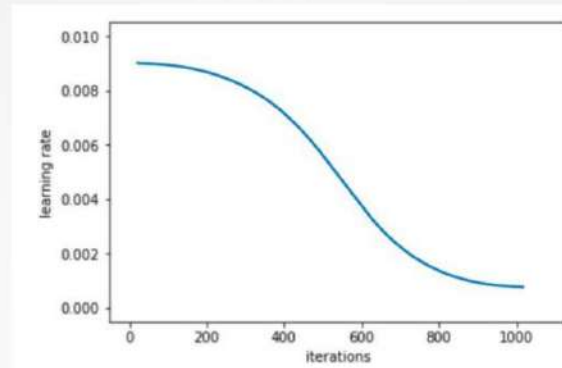
Learning Rate Schedules: Time-based



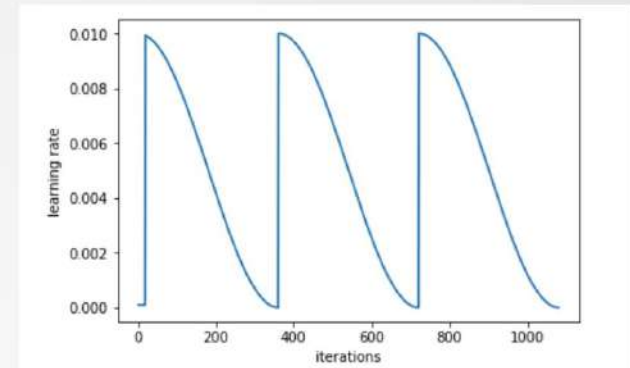
UNIVERSIDAD
CATÓLICA
BOLIVIANA
LA PAZ



Step Decay

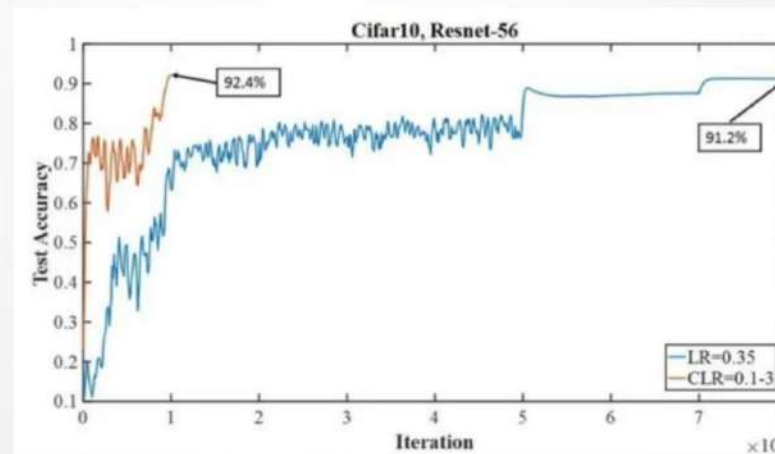


Cosine Annealing



Cosine Annealing with Restart

Reading: Learning Rate
Schedule in Practice: an
example with Keras and
TensorFlow 2.0 on
Medium



Example of superconvergence

Funciones de Optimización

MSc. Edwin Salcedo

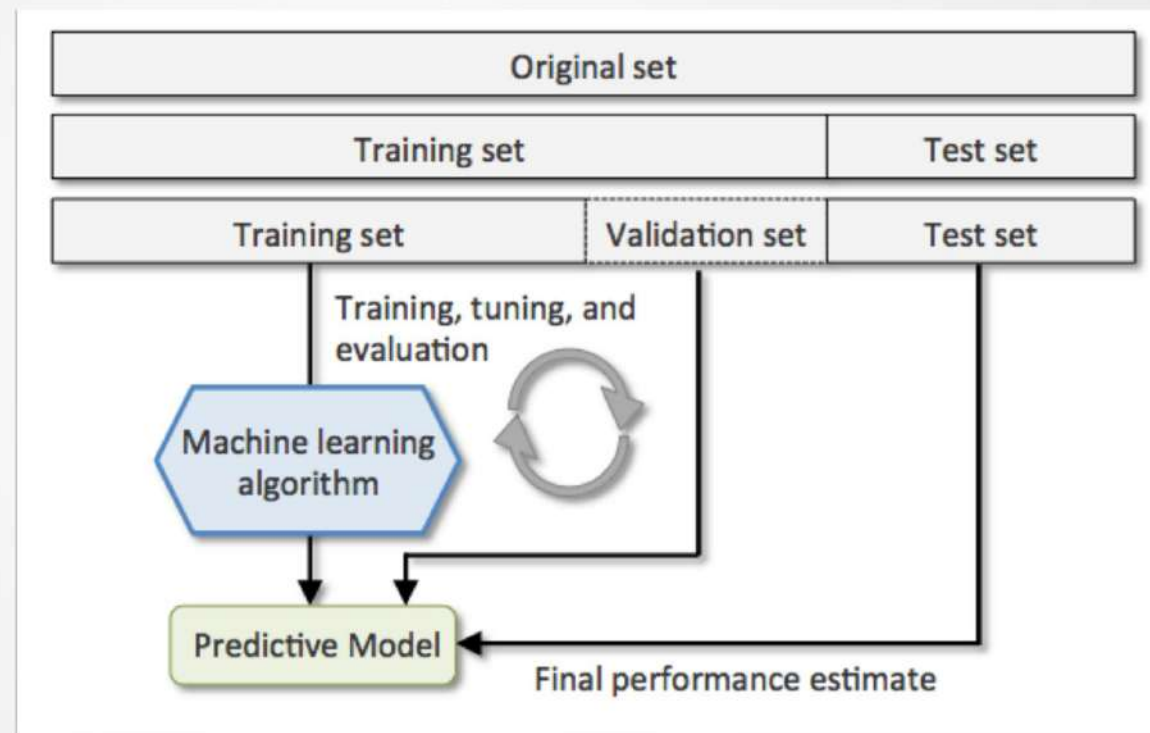
Conjunto de entrenamiento, prueba y validación



UNIVERSIDAD
CATÓLICA
BOLIVIANA
LA PAZ



Dataset



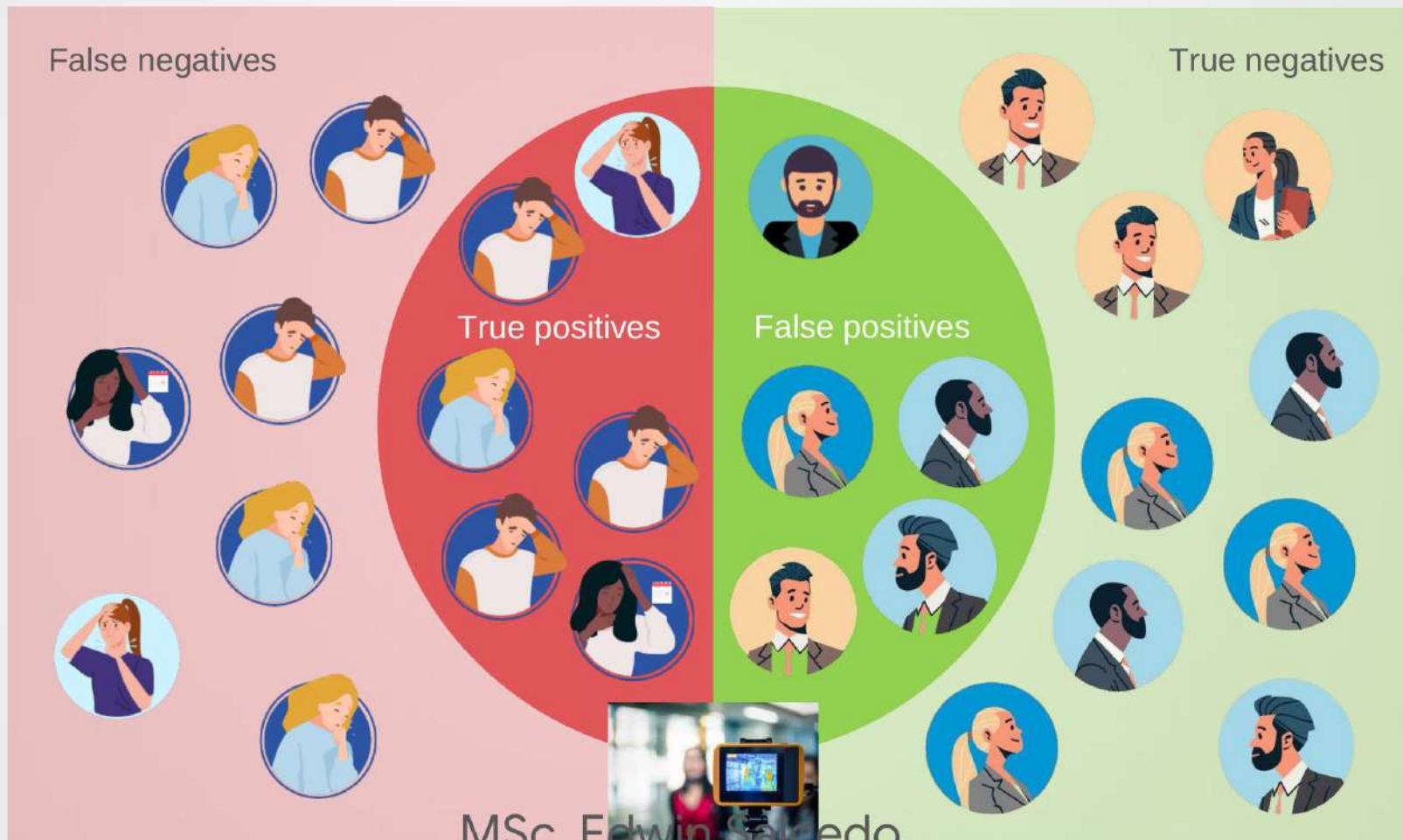
MSc. Edwin Salcedo

Métricas de evaluación



UNIVERSIDAD
CATÓLICA
BOLIVIANA
LA PAZ

Elementos
Relevantes



Métricas de evaluación



UNIVERSIDAD
CATÓLICA
BOLIVIANA
LA PAZ

Classification Accuracy

La precisión de la clasificación es lo que generalmente queremos decir cuando usamos el término precisión. Es la relación entre el número de predicciones correctas y el número total de muestras de entrada.

$$\text{Accuracy} = \frac{\text{Numero de predicciones bien hechas}}{\text{Numero total de predicciones hechas}}$$

Funciona bien solo si hay igual número de muestras pertenecientes a cada clase.

MSc. Edwin Salcedo

Métricas de evaluación



UNIVERSIDAD
CATÓLICA
BOLIVIANA
LA PAZ

Confusion Matrix

Confusion Matrix nos da una matriz como salida y describe el rendimiento completo del modelo.

		Real label	
		Positive	Negative
Predicted label	Positive	True Positive	False positive
	Negative	False negative	True Negative

MSc. Edwin Salcedo

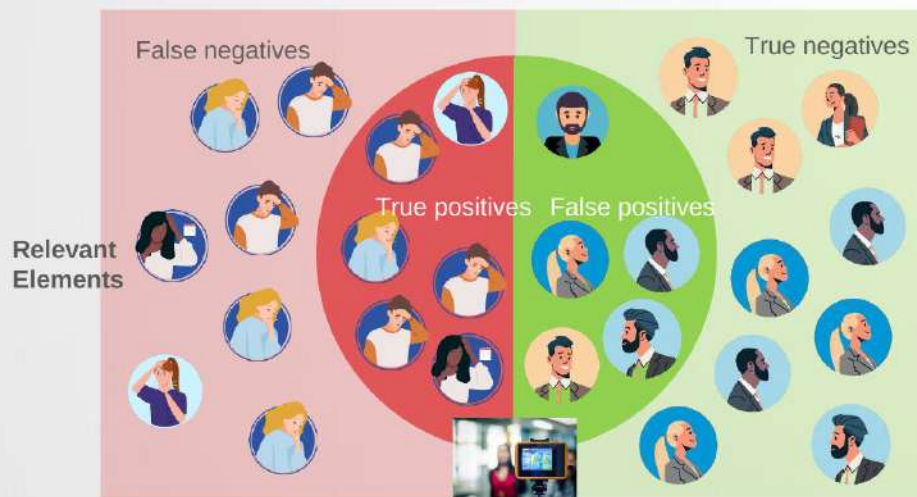
Métricas de evaluación: Confusion Matrix



UNIVERSIDAD
CATÓLICA
BOLIVIANA
LA PAZ

PASO 1: Identificar components

Confusion Matrix nos da una matriz como salida y describe el rendimiento completo del modelo.



Predicted
label

Predicted row	Real label		
	Positive	Negative	
Positive	6	5	= 11
Negative	7	9	= 16
		= 13	
		= 14	
			Actual row
All adds up = 27			

MSc. Edwin Salcedo

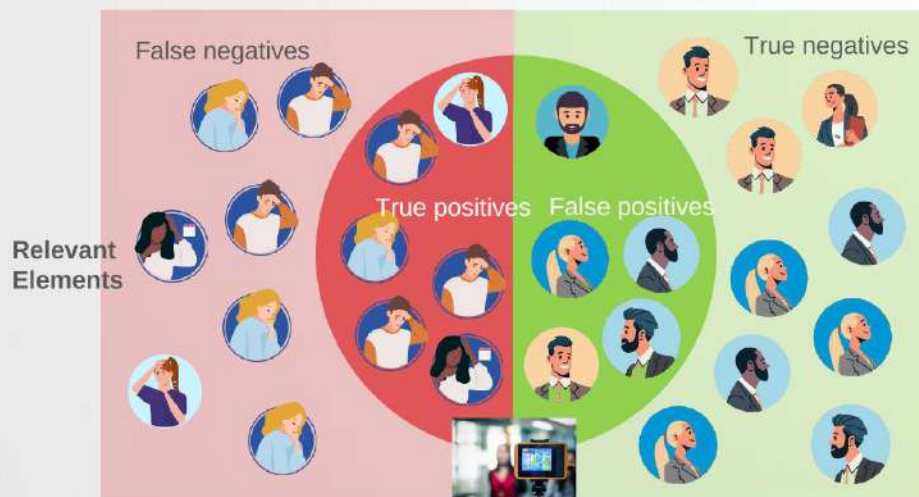
Métricas de evaluación: Confusion Matrix



UNIVERSIDAD
CATÓLICA
BOLIVIANA
LA PAZ

PASO 2: Normalizar valores

Confusion Matrix nos da una matriz como salida y describe el rendimiento completo del modelo.



		Real label		
		Positive	Negative	
Predicted label	Positive	0,55	0,45	= 1
	Negative	0,44	0,56	= 1

MSc. Edwin Salcedo

Métricas de evaluación: Confusion Matrix



UNIVERSIDAD
CATÓLICA
BOLIVIANA
LA PAZ

PASO 3: Calcule la precisión promediando la diagonal principal

Confusion Matrix nos da una matriz como salida y describe el rendimiento completo del modelo.

		Real label		
		Positive	Negative	
Predicted label	Positive	0,55	0,45	= 1
	Negative	0,44	0,56	= 1

$$(0,55 + 0,56)/2 = 0,55 = \text{Accuracy}$$

MSc. Edwin Salcedo

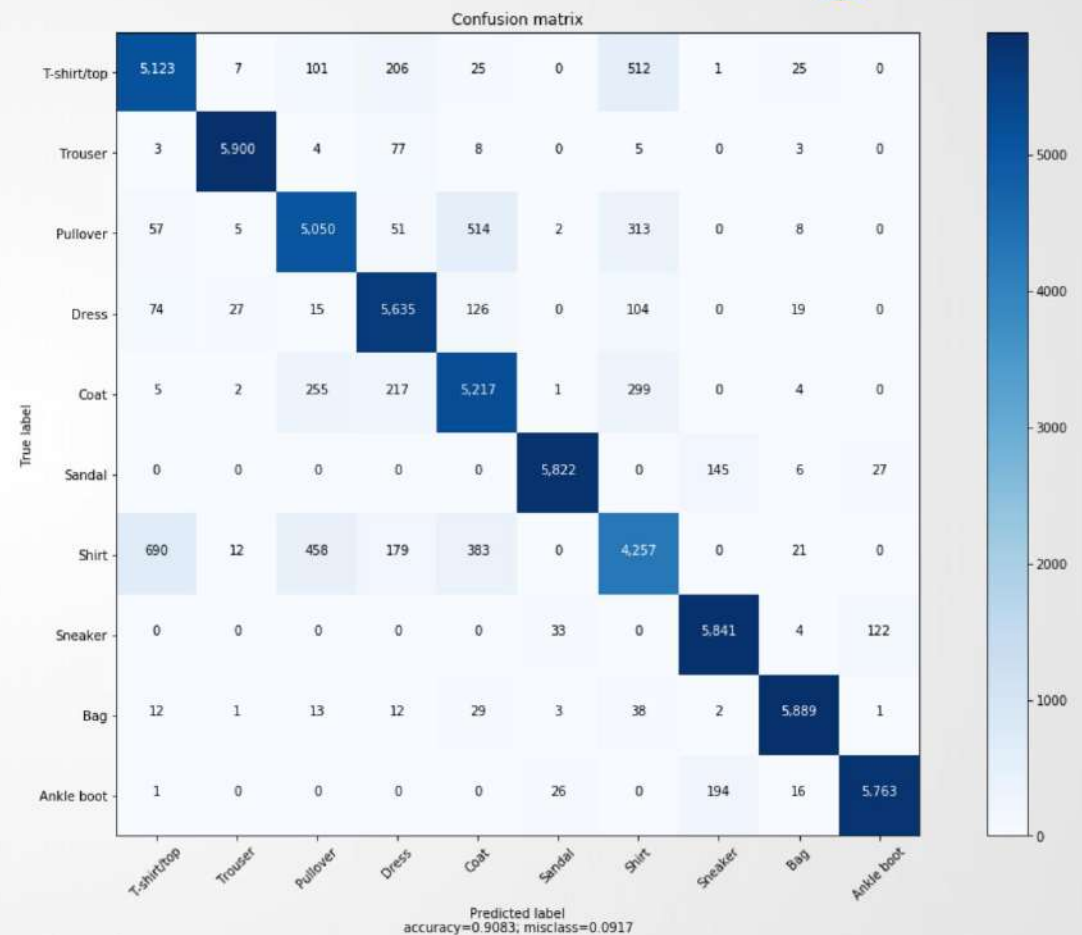
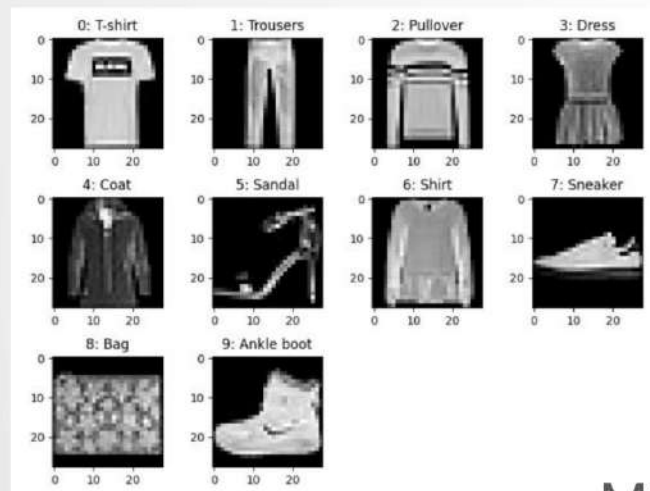
Métricas de evaluación: Confusion Matrix



UNIVERSIDAD
CATÓLICA
BOLIVIANA
LA PAZ

Fashion-MNIST Dataset

Dataset de 70000 imágenes de ropa para clasificarlas según 10 categorías: camiseta, pantalón, jersey, vestido, abrigo, sandalia, camisa, zapatilla, bolso, botín.



MSc. Edwin Salcedo

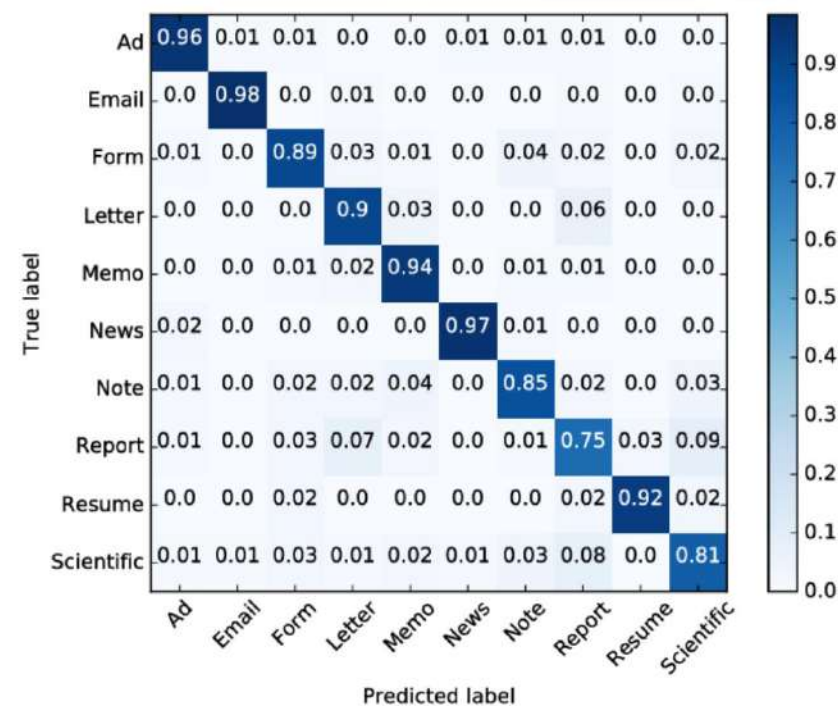
Métricas de evaluación: Confusion Matrix



UNIVERSIDAD
CATÓLICA
BOLIVIANA
LA PAZ

Tobacco-3482 Dataset

Conjunto de datos para clasificar texto en función de su contenido visual.



MSc. Edwin Salcedo

Jayant Kumar, Peng Ye and David Doermann. "Learning Document Structure for Retrieval and Classification." (ICPR 2012), 2012.

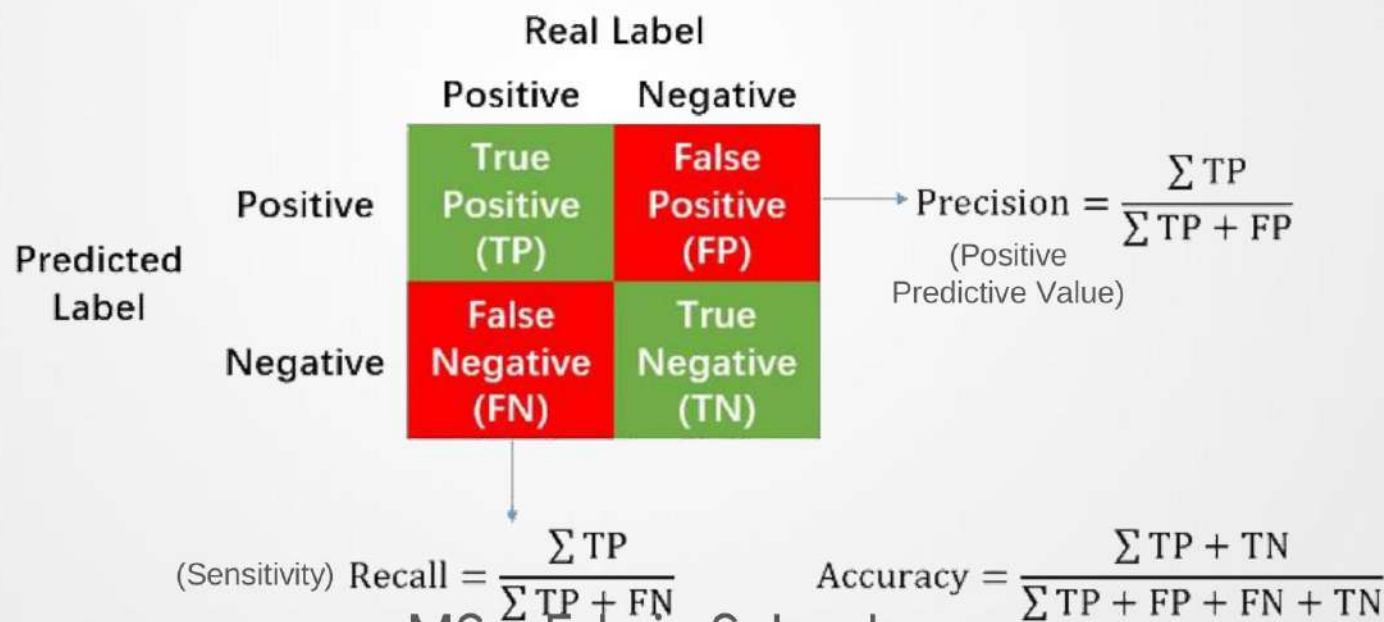
Métricas de evaluación



UNIVERSIDAD
CATÓLICA
BOLIVIANA
LA PAZ

Precision and recall

La precisión es la fracción de instancias relevantes entre las instancias recuperadas, mientras que la recuperación es la fracción de instancias relevantes que se recuperaron.



MSc. Edwin Salcedo

Métricas de evaluación



UNIVERSIDAD
CATÓLICA
BOLIVIANA
LA PAZ

Logarithmic Loss

Logarithmic Loss o Log Loss, funciona penalizando las clasificaciones falsas. Funciona bien para la clasificación multiclase.

$$\text{LogarithmicLoss} = \frac{-1}{N} \sum_{i=1}^N \sum_{j=1}^M y_{ij} * \log(p_{ij})$$

Where:

y_{ij} indicates whether sample i belongs to class j or not

p_{ij} indicates the probability of sample i belonging to class j $[0,1]$

LogarithmicLoss has no upper bound $[0, \infty)$

MSc. Edwin Salcedo

Métricas de evaluación

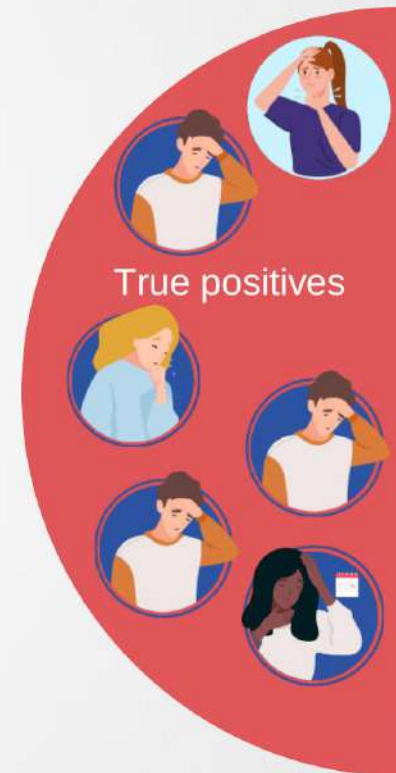


UNIVERSIDAD
CATÓLICA
BOLIVIANA
LA PAZ

AUC

El área bajo la curva (AUC) es una de las métricas más utilizadas para la evaluación. Se utiliza para problemas de clasificación binaria.

$$TPR \text{ (True Positive Rate / Recall / Sensitivity)} = \frac{TP}{TP + FN}$$



MSc. Edwin Salcedo

Métricas de evaluación



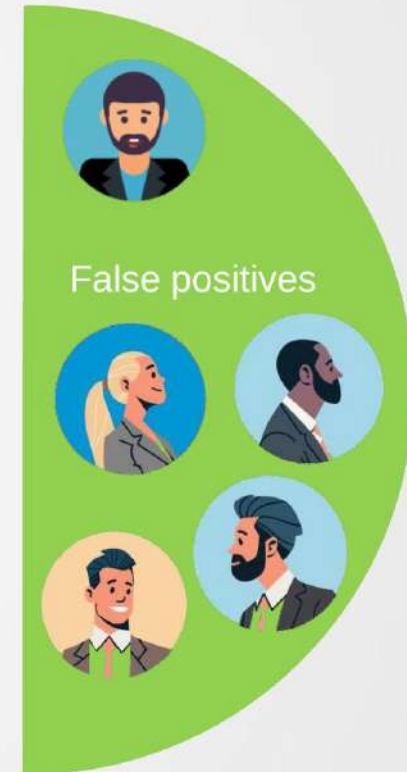
UNIVERSIDAD
CATÓLICA
BOLIVIANA
LA PAZ

AUC

El área bajo la curva (AUC) es una de las métricas más utilizadas para la evaluación. Se utiliza para problemas de clasificación binaria.

$$\textit{Specificity} = \frac{TN}{TN + FP}$$

$$\textit{FPR (False Positive Rate)} = 1 - \textit{Specificity}$$



MSc. Edwin Salcedo

Métricas de evaluación

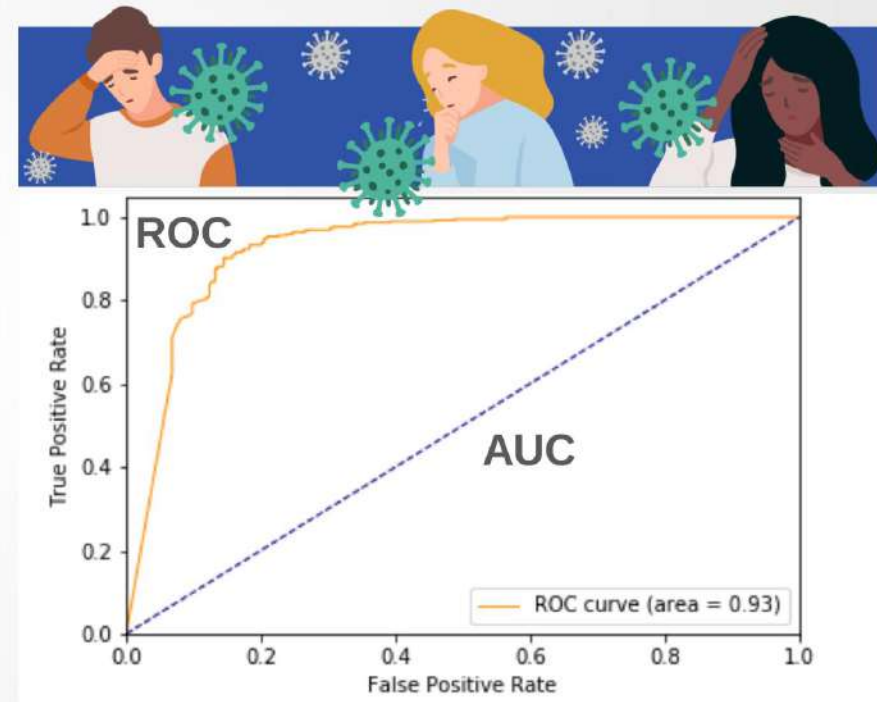


UNIVERSIDAD
CATÓLICA
BOLIVIANA
LA PAZ

Area Under Curve (AUC)

Samples	%	Class
Kiril	0.21	Healthy
Hans	0.67	Infected
Carl	0.56	Healthy
Jane	0.31	Healthy
Julie	0.70	Infected
Carol	0.84	Healthy

TPR



FPR

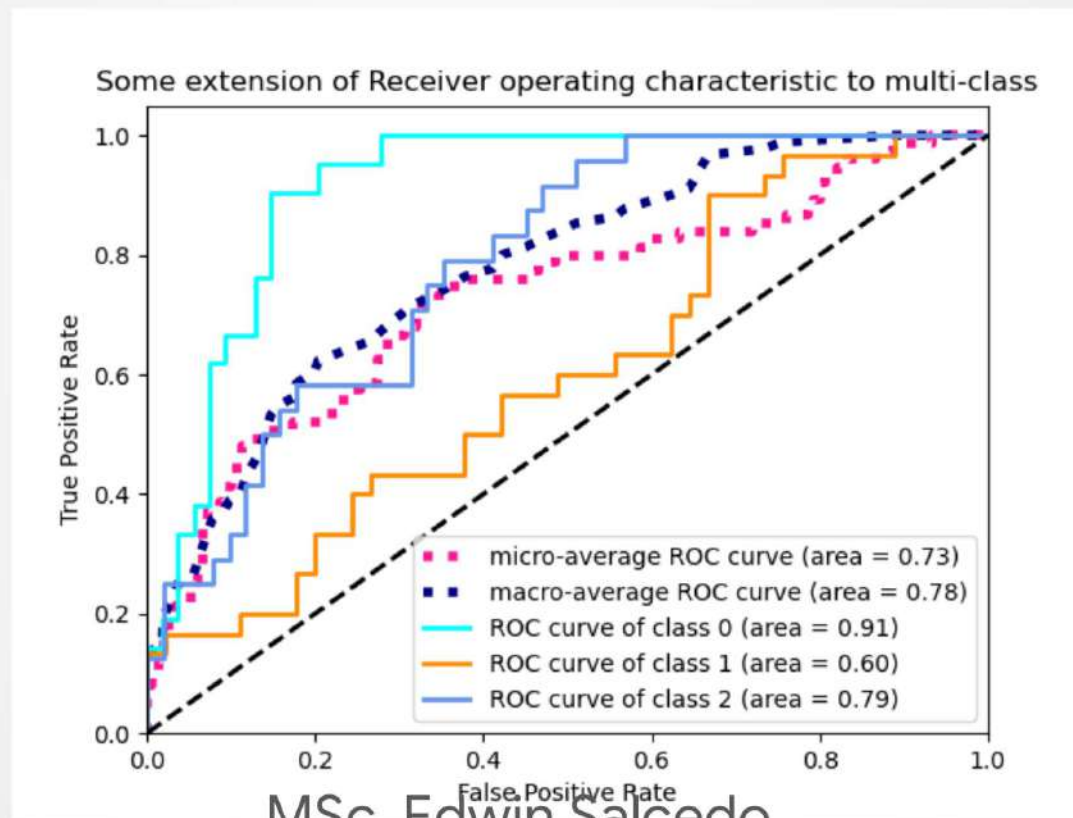
MSc. Edwin Salcedo

Métricas de evaluación



UNIVERSIDAD
CATÓLICA
BOLIVIANA
LA PAZ

Area Under Curve (AUC) for Multi-Category Classification



Métricas de evaluación



UNIVERSIDAD
CATÓLICA
BOLIVIANA
LA PAZ

F1 Score

F1 es una función de Precisión y Recuperación. El valor más alto posible de una puntuación F es 1,0, lo que indica precisión y recuperación perfectas.

$$F1\ Score = 2 * \frac{1}{\frac{1}{precision} + \frac{1}{recall}}$$

MSc. Edwin Salcedo

Logs de entrenamiento



UNIVERSIDAD
CATÓLICA
BOLIVIANA
LA PAZ

```
48000/48000 [=====] - 55s 1ms/step - loss: 0.0375 - acc: 0.8041 - val_loss: 0.0241 - val_acc: 0.8078
Epoch 2/50
48000/48000 [=====] - 49s 1ms/step - loss: 0.0219 - acc: 0.8085 - val_loss: 0.0207 - val_acc: 0.8088
Epoch 3/50
48000/48000 [=====] - 59s 1ms/step - loss: 0.0196 - acc: 0.8096 - val_loss: 0.0190 - val_acc: 0.8113
Epoch 4/50
48000/48000 [=====] - 65s 1ms/step - loss: 0.0185 - acc: 0.8101 - val_loss: 0.0181 - val_acc: 0.8106
Epoch 5/50
48000/48000 [=====] - 71s 1ms/step - loss: 0.0177 - acc: 0.8105 - val_loss: 0.0174 - val_acc: 0.8113
Epoch 6/50
48000/48000 [=====] - 73s 2ms/step - loss: 0.0171 - acc: 0.8108 - val_loss: 0.0170 - val_acc: 0.8122
Epoch 7/50
48000/48000 [=====] - 73s 2ms/step - loss: 0.0167 - acc: 0.8110 - val_loss: 0.0166 - val_acc: 0.8120
Epoch 8/50
48000/48000 [=====] - 74s 2ms/step - loss: 0.0163 - acc: 0.8111 - val_loss: 0.0162 - val_acc: 0.8121
Epoch 9/50
48000/48000 [=====] - 80s 2ms/step - loss: 0.0161 - acc: 0.8113 - val_loss: 0.0160 - val_acc: 0.8119
Epoch 10/50
48000/48000 [=====] - 76s 2ms/step - loss: 0.0158 - acc: 0.8114 - val_loss: 0.0158 - val_acc: 0.8128
Epoch 11/50
48000/48000 [=====] - 80s 2ms/step - loss: 0.0156 - acc: 0.8115 - val_loss: 0.0156 - val_acc: 0.8129
Epoch 12/50
48000/48000 [=====] - 79s 2ms/step - loss: 0.0153 - acc: 0.8116 - val_loss: 0.0153 - val_acc: 0.8127
Epoch 13/50
48000/48000 [=====] - 86s 2ms/step - loss: 0.0151 - acc: 0.8117 - val_loss: 0.0152 - val_acc: 0.8120
Epoch 14/50
48000/48000 [=====] - 89s 2ms/step - loss: 0.0150 - acc: 0.8118 - val_loss: 0.0151 - val_acc: 0.8120
Epoch 15/50
48000/48000 [=====] - 86s 2ms/step - loss: 0.0148 - acc: 0.8118 - val_loss: 0.0148 - val_acc: 0.8130
Epoch 16/50
48000/48000 [=====] - 94s 2ms/step - loss: 0.0147 - acc: 0.8119 - val_loss: 0.0147 - val_acc: 0.8130
Epoch 17/50
48000/48000 [=====] - 98s 2ms/step - loss: 0.0145 - acc: 0.8120 - val_loss: 0.0145 - val_acc: 0.8129
```

MSc. Edwin Salcedo
Keras

Herramientas de implementación



UNIVERSIDAD
CATÓLICA
BOLIVIANA
LA PAZ



Programación
Pura



Composición de capas y
modelos desechables.



MSc. Edwin Salcedo



Arquitectura
tuneable

