



UNIVERSIDAD  
**NACIONAL**  
DE COLOMBIA

# Clustering – Aprendizaje No Supervisado

**3007855 - Inteligencia Artificial**

**3010476 - Introducción a la Inteligencia Artificial**

**Semestre: 02/2021**

**Prof. Demetrio Arturo Ovalle Carranza**  
**Departamento de Ciencias de la Computación**  
**y de la Decisión**  
**Facultad de Minas**

**Noviembre 23 de 2021**

**LMS:** <https://minaslap.net/user/index.php?id=560>

**Link Clases:** [meet.google.com/quy-okvi-ugq](https://meet.google.com/quy-okvi-ugq)

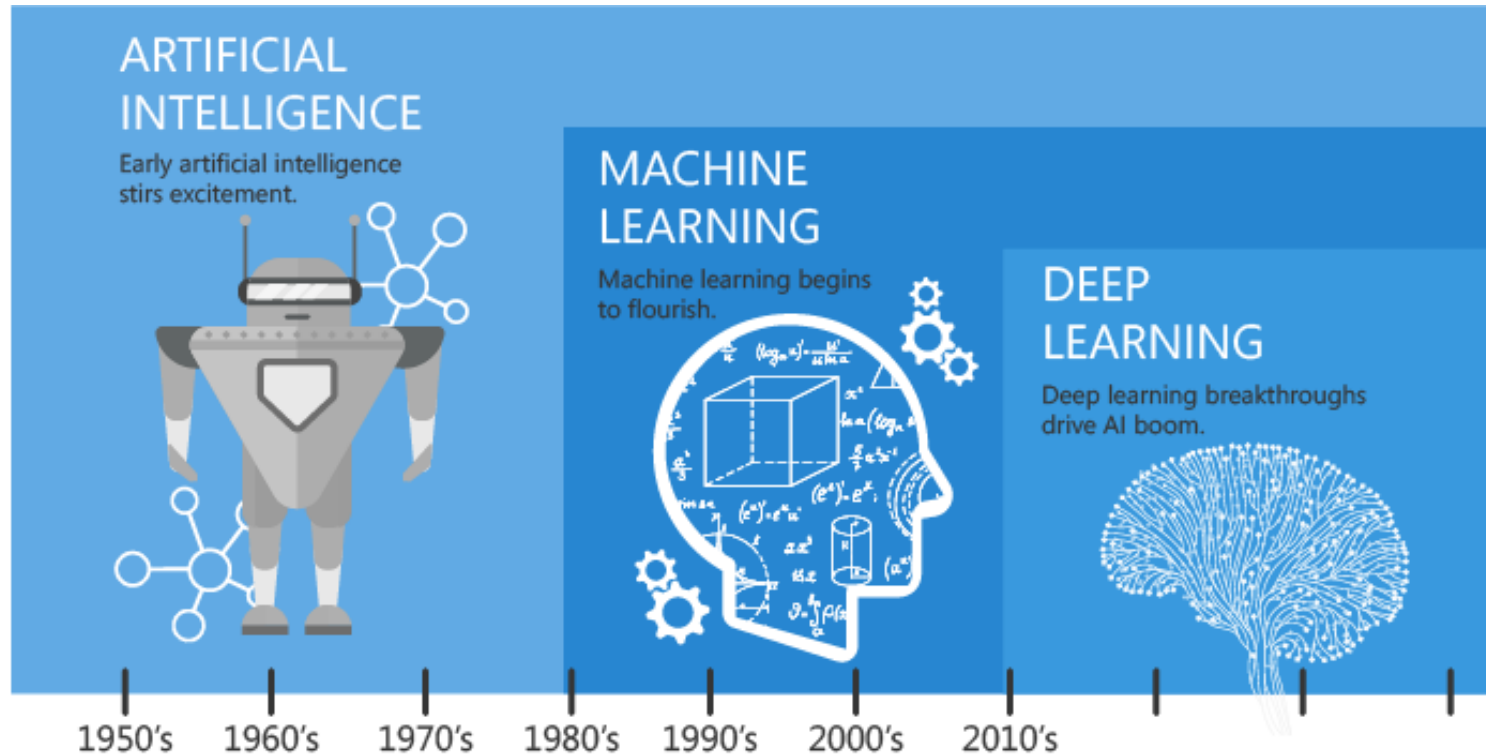


# Contenido

Introducción al Machine Learning (ML):

- Aprendizaje No Supervisado
- Aprendizaje por Refuerzo
- Aprendizaje Supervisado

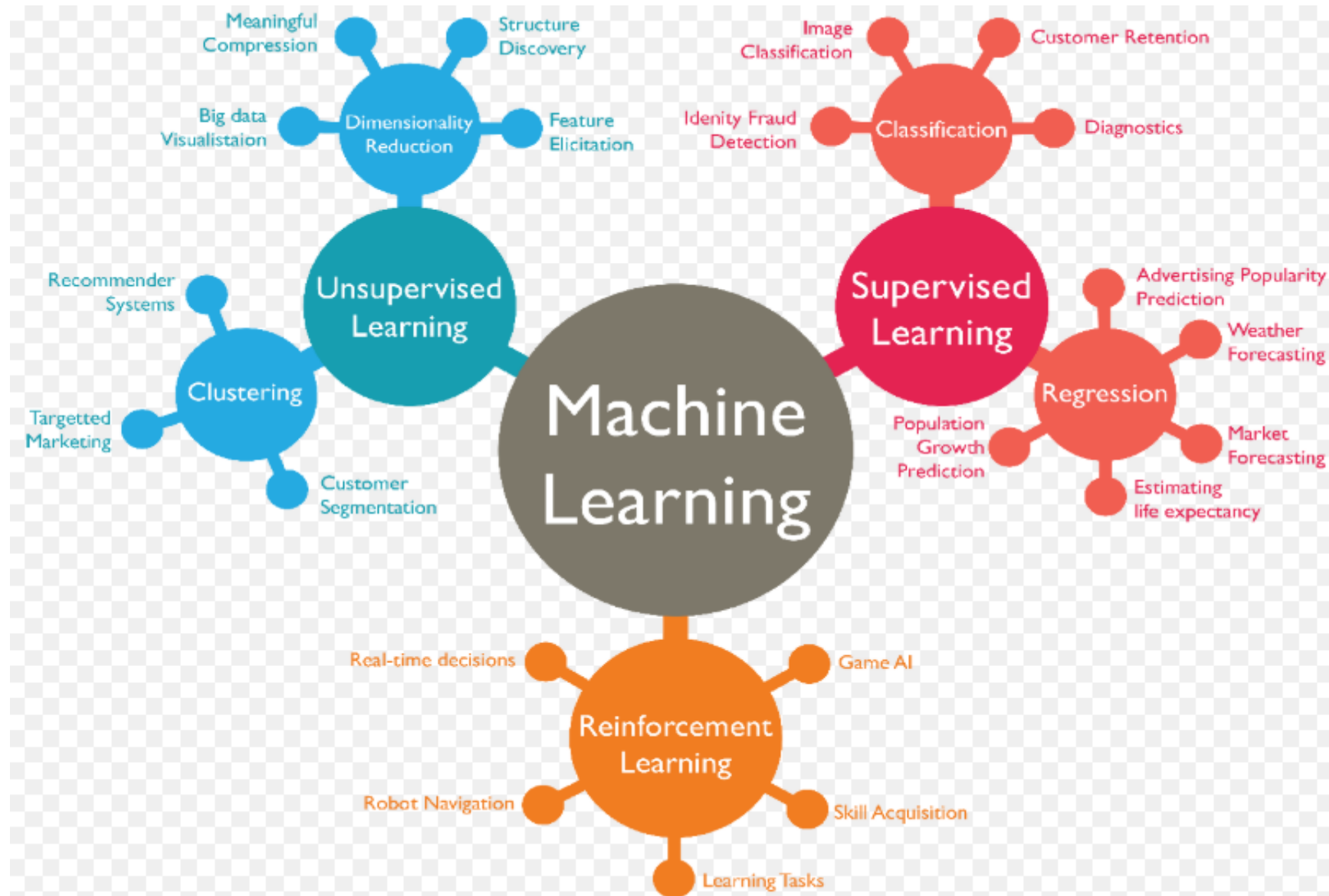
# Inteligencia Artificial, Machine Learning y Deep Learning



Since an early flush of optimism in the 1950's, smaller subsets of artificial intelligence - first machine learning, then deep learning, a subset of machine learning - have created ever larger disruptions.

Pic Credit: Linked IN | Machine Learning vs Deep Learning

Tomado de : Xenostack



# Enfoques de Machine Learning

**1. Aprendizaje No Supervisado:** el sistema intenta descubrir la estructura oculta de los datos o asociaciones entre variables. En ese caso, los datos de entrenamiento consisten en instancias sin ninguna etiqueta correspondiente. Ejemplos: Reglas de Asociación (e.g. Análisis de secuencias biológicas), Clustering.

**2. Aprendizaje por Refuerzo:** El sistema intenta aprender a través de la interacción directa con el entorno para maximizar alguna noción de recompensa acumulada. El sistema **no tiene conocimiento** del comportamiento del entorno y la única forma de averiguarlo es a través de **prueba y error**. El sistema toma decisiones en tiempo real. Fruto de este aprendizaje el sistema es capaz de adquirir habilidades. Ejemplos de aplicaciones: Aprendizaje con Robots Autónomos, Navegación con robots, Videojuegos, etc.

# Enfoques de Machine Learning

**Aprendizaje por Refuerzo:** El sistema aprende a través de la interacción directa con el entorno teniendo en cuenta la noción de recompensa acumulada. Fruto de este aprendizaje el sistema es capaz de adquirir habilidades. Ejemplo: Realizar tareas de clasificación y ordenamiento de objetos.



# Enfoques de Machine Learning

**3. Aprendizaje Supervisado:** El sistema debe "aprender" inductivamente una función llamada **función objetivo**, la cual es una expresión de un modelo que describe los datos. Se realiza a través de un subconjunto de casos de entrada o datos de entrenamiento (ejemplos) donde se conoce su valor de salida (label o etiqueta).

Ejemplos: **Clasificación** y **Regresión**. Los modelos de clasificación intentan predecir distintas clases, como por ejemplo grupos sanguíneos, mientras que los modelos de regresión predicen valores numéricos. Técnicas: Los árboles de decisión (DT), la clasificación bayesiana (BC), el aprendizaje de reglas y el aprendizaje basado en instancias (IBL), los k-vecinos más cercanos (k-NN), los algoritmos genéticos (GA), Redes Neuronales (ANN) y las máquinas de soporte vectorial (SVM).



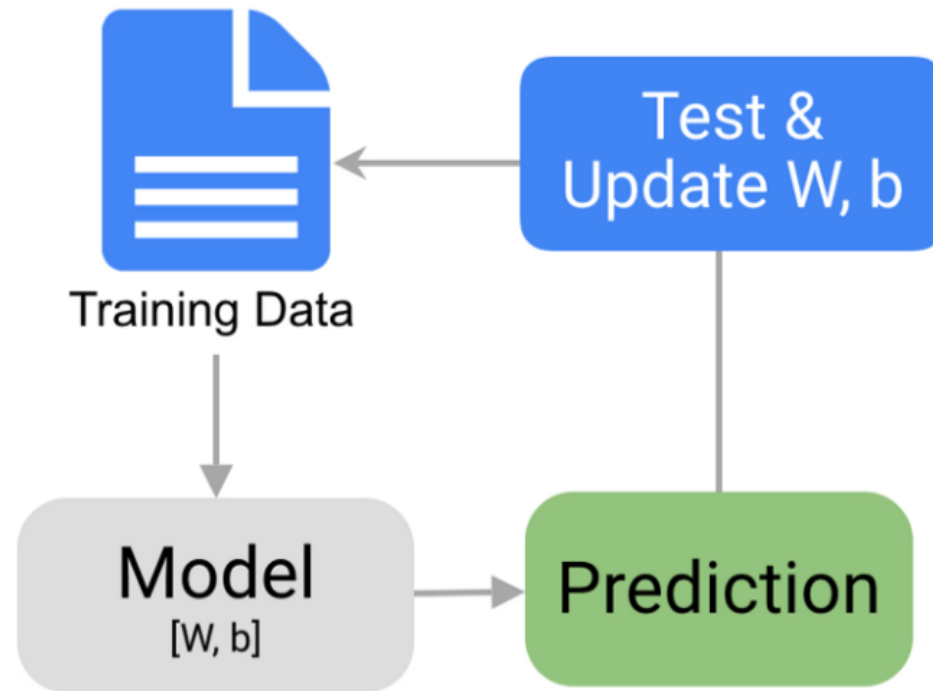
# Conceptos

- **Label** (etiqueta): dato de salida o lo que estamos intentando predecir con un modelo de Aprendizaje Supervisado.
- **Feature** (característica, variable): dato de entrada.
- **Modelo**: relación entre features y labels.
- **Función de activación**: relaciona entradas, salidas, pesos y segos.
- **Fase de training** (entrenamiento o aprendizaje) aprende con ejemplos de entradas etiquetadas.
- **Epochs**: cantidad de iteraciones de entrenamiento
- **Fase de predicción** (ó inferencia) aplicación del modelo entrenado con datos no etiquetados.

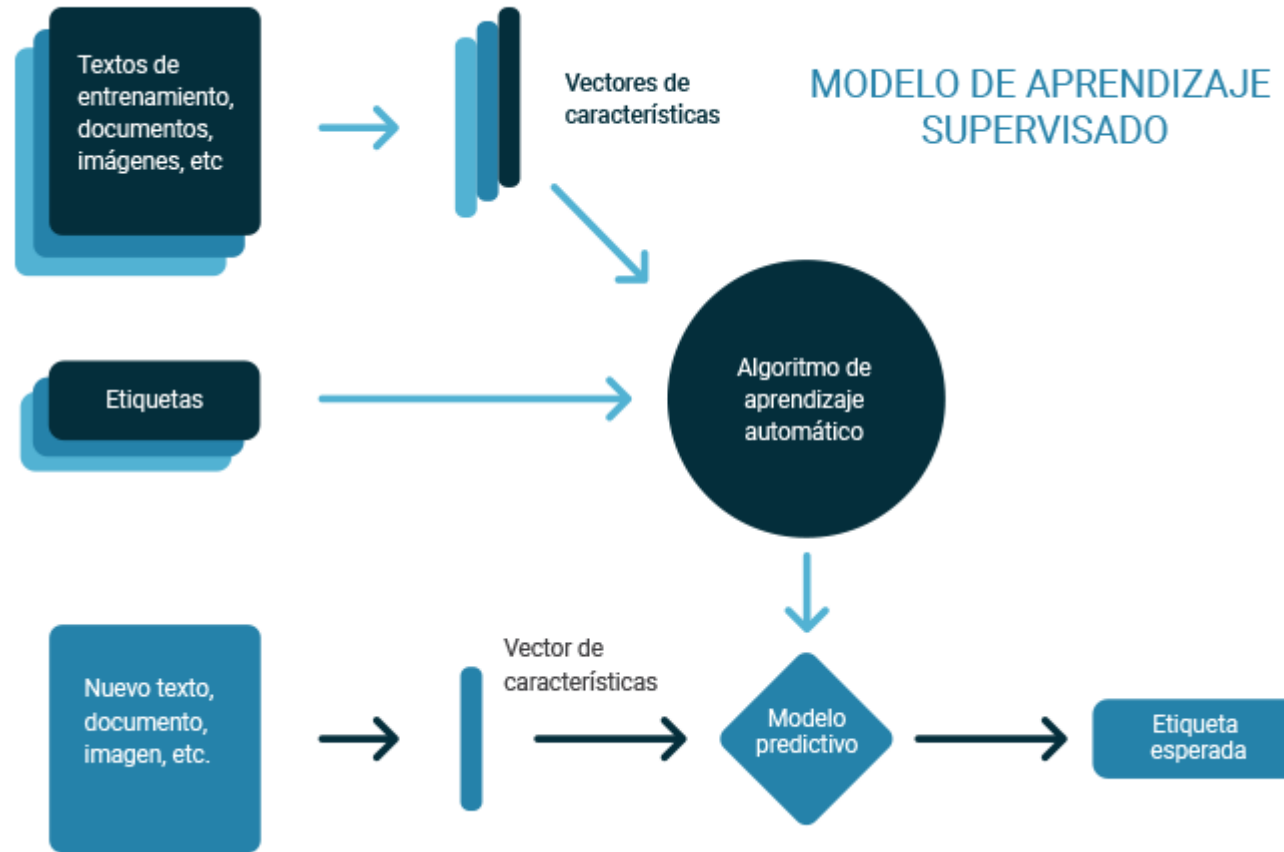
## Aprendizaje Supervisado

- ✓ El sistema debe "**aprender**" inductivamente una función llamada **función objetivo**, la cual es una expresión de un modelo que describe los datos.
- ✓ Se realiza un **entrenamiento** a través de un subconjunto de casos de entrada (ejemplos) donde se conoce su valor de salida (**datos etiquetados**).
- ✓ En este caso, los datos que usamos para el entrenamiento incluyen la solución deseada, llamada "**etiqueta**" (label).

## Ciclo del Aprendizaje



# Ejemplo de aprendizaje supervisado para textos y documentos

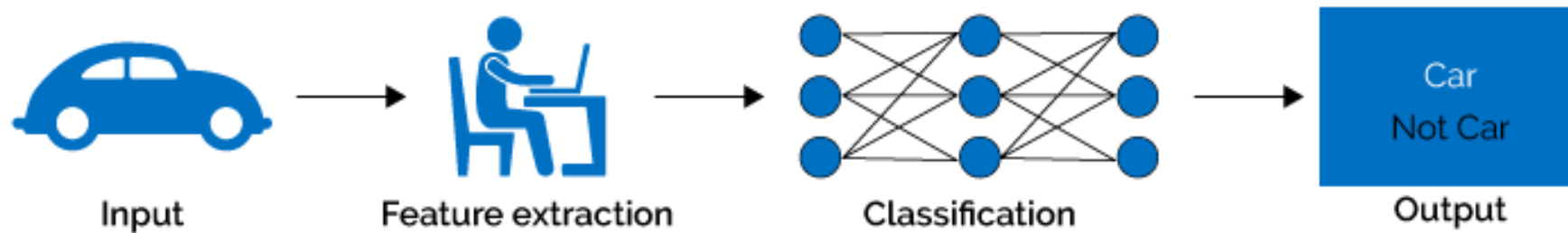


# Fases de Entrenamiento y Predicción

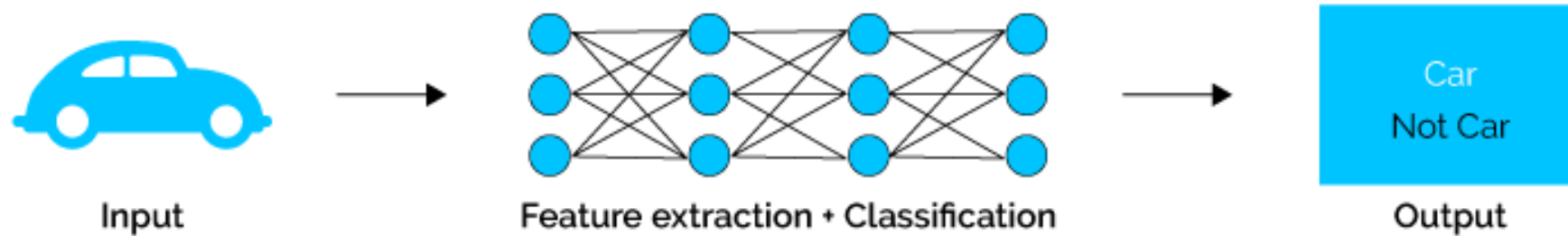
Las fases de entrenamiento y predicción del modelo de aprendizaje supervisado requiere de los siguientes pasos:

- Alimentar los datos de entrenamiento al modelo.
- El modelo aprende a asociar features (características) y etiquetas.
- Se le pide al modelo que haga predicciones sobre un conjunto de muestras no etiquetadas.
- Verificamos que las predicciones coincidan con las etiquetas definidas con datasets de prueba.

## Machine Learning



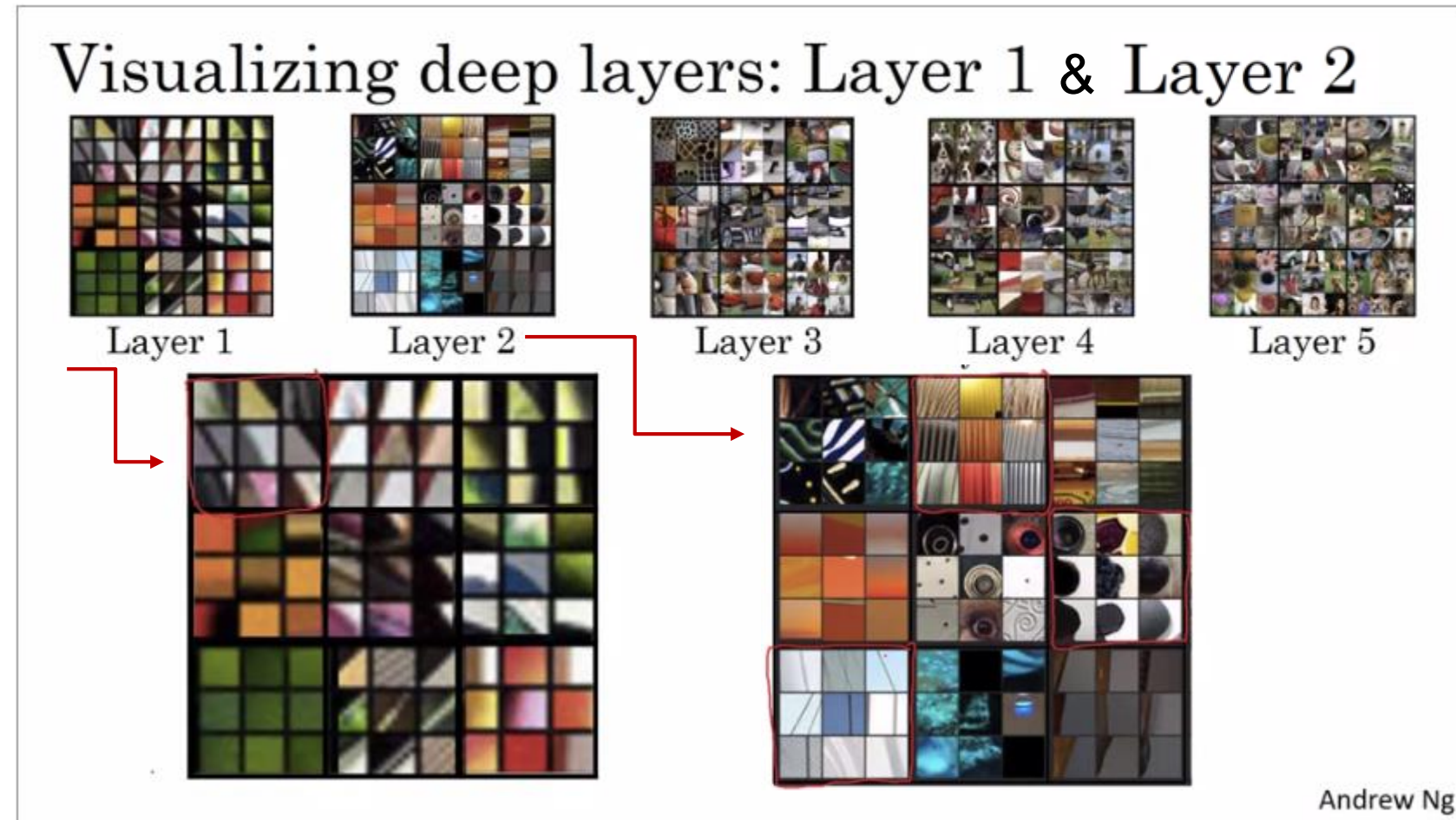
## Deep Learning



Pic Credit: Xenonstack | Machine Learning vs Deep Learning

# Deep Learning

What are deep ConvNets learning?



Tomado de : deeplearning.ai

# Deep Learning

What are deep ConvNets learning?

## Visualizing deep layers: Layer 3



Layer 1



L



4



Layer 5

Andrew Ng

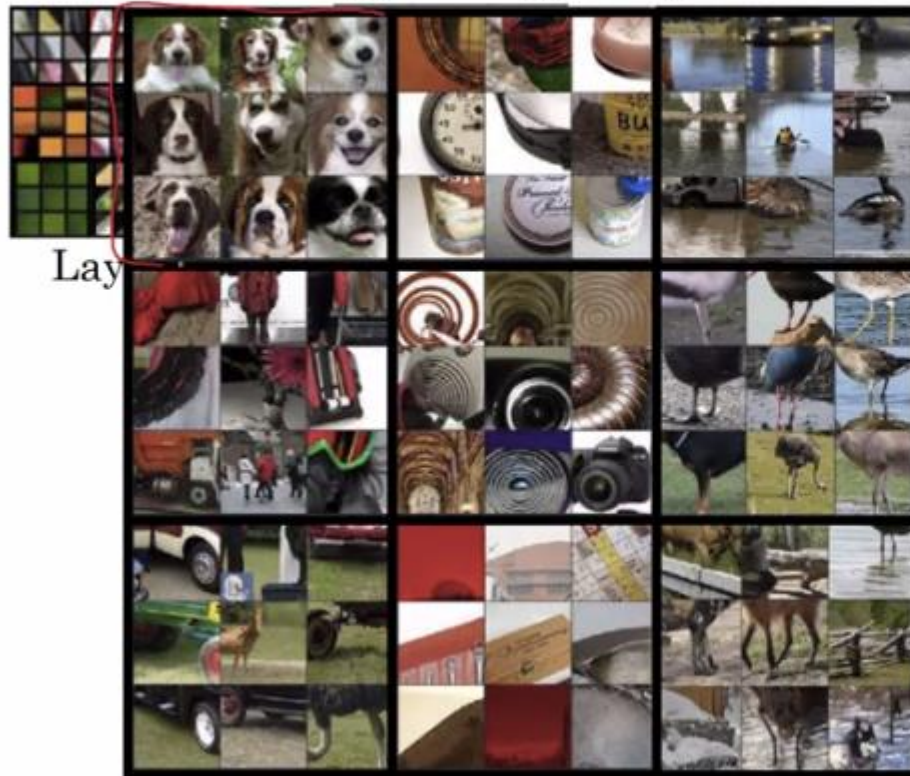
Tomado de : deeplearning.ai



# Deep Learning

What are deep ConvNets learning?

## Visualizing deep layers: Layer 4



Layer 4



Layer 4



Layer 5

Andrew Ng

Tomado de : deeplearning.ai

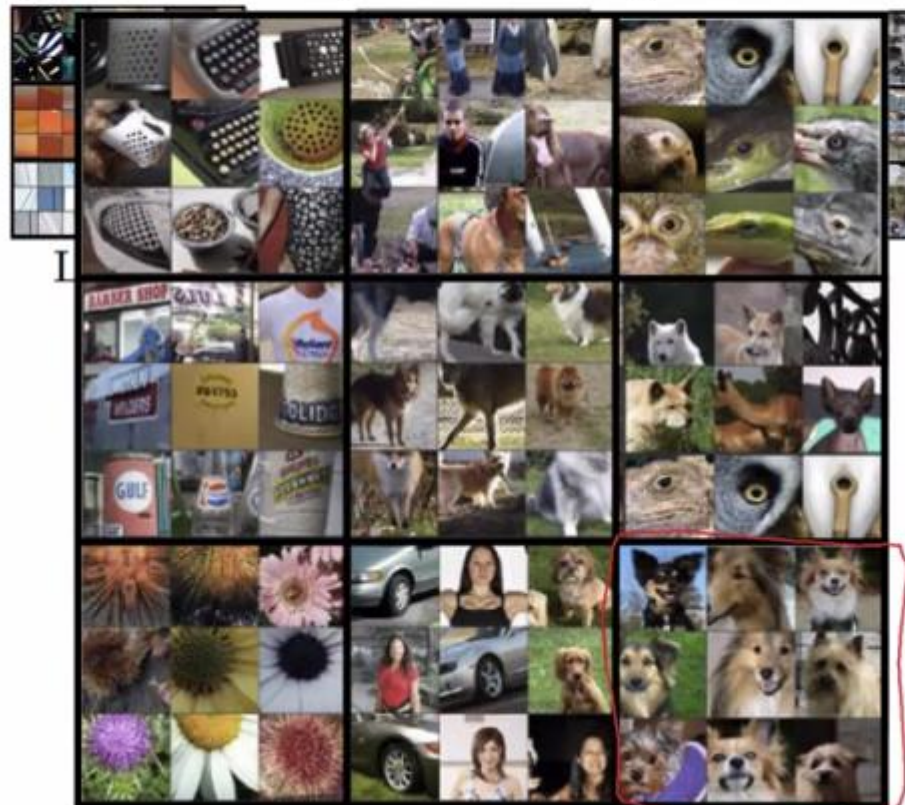
# Deep Learning

What are deep ConvNets learning?

## Visualizing deep layers: Layer 5



Layer 1



Layer 5

Andrew Ng

Tomado de : [deeplearning.ai](http://deeplearning.ai)

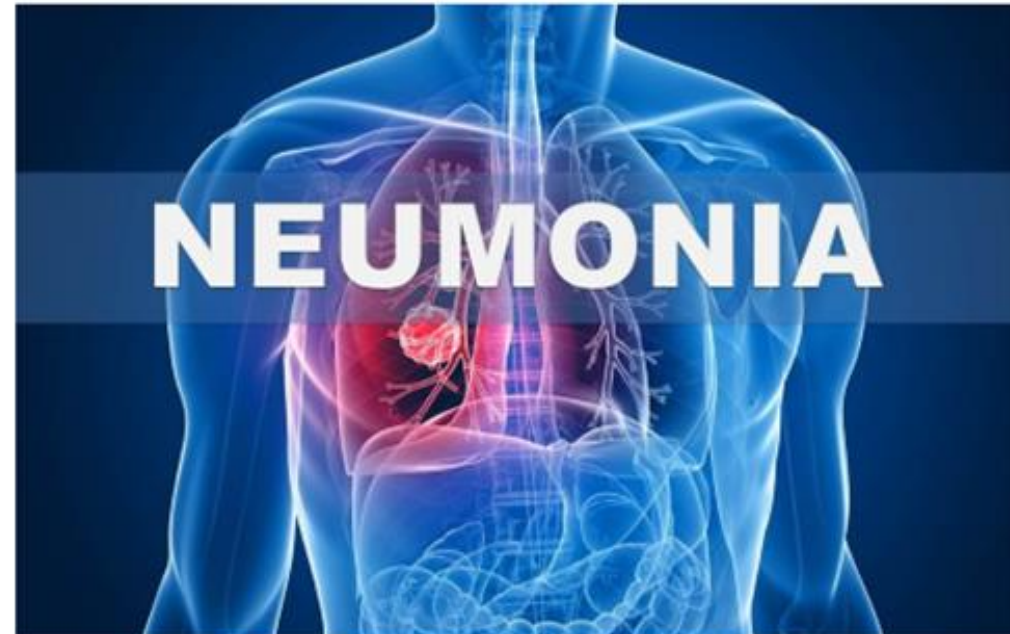
# CNN – Red Neuronal Convolucional

Video: **How do convolutions work**

Fuente: Andrew Ng, *CNN - Convolutionals  
Neural Networks*

## Sistema CheXNet: Aprendizaje Supervisado

CheXNet: Detección de Neumonía a Nivel de Radiólogo en Rayos X de Tórax con Aprendizaje Profundo



CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning (Rajpurkar et al., 2017)

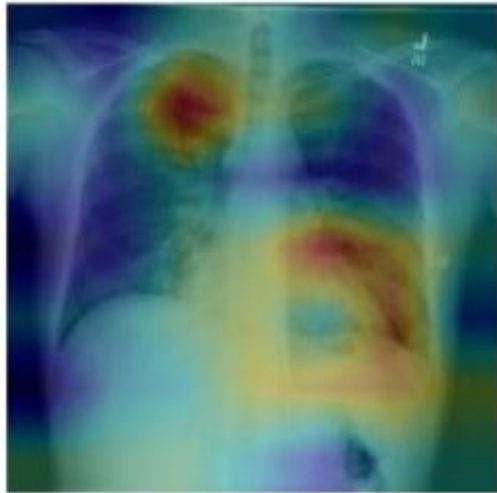


## Problemas actuales de la detección de la neumonía

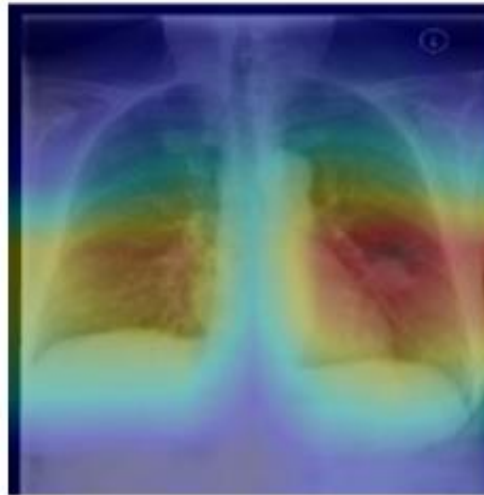
- Las imágenes de rayos X a menudo son vagas
- Diagnóstico puede solaparse con otros diagnósticos.
- Inclusive realmente pueden ser anomalías benignas y no una enfermedad o patología.

Estas discrepancias causan una considerable variabilidad entre los radiólogos en el diagnóstico de la neumonía.

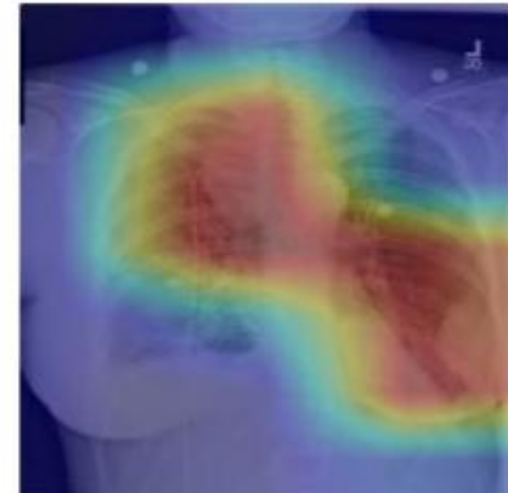
## Sistema CheXNet: Algunas Patologías Detectadas



Neumonía

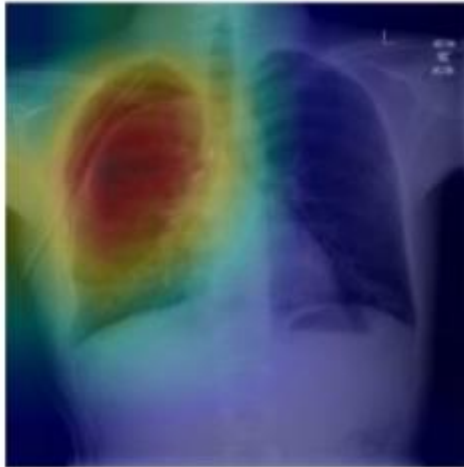


Nódulo en el pulmón izquierdo



Malignidad pulmonar  
primaria

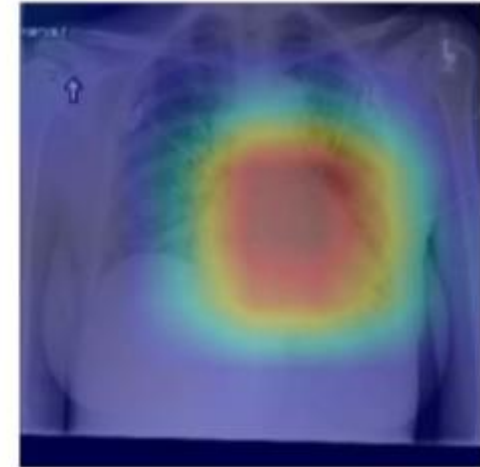
## Sistema CheXNet: Algunas Patologías Detectadas (cont)



Neumotórax

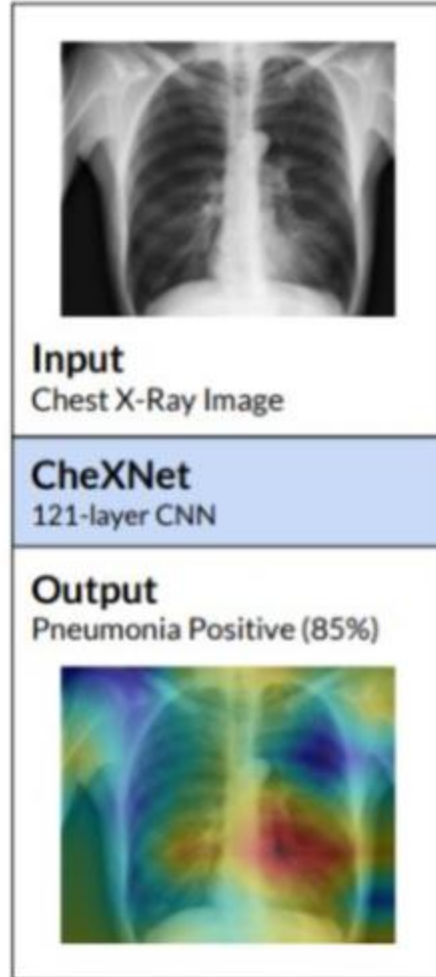


Derrame pleural



Insuficiencia cardíaca  
congestiva y  
cardiomegalia

# Sistema CheXNet: Aprendizaje Supervisado



## Entrada:

Radiografía tomada del tórax

## CheXNet:

Red convolucional de 121 capas, entrenada con un conjunto de radiografías (ChestX-ray14), cada una con hasta 14 etiquetas de patologías torácicas diferentes utilizando métodos de extracción automática en los informes de radiologías.

## Salida:

Probabilidad de neumonía junto con un mapa térmico que localiza las áreas de la imagen más indicativas de neumonía.



## Sistema CheXNet-14: algunas características

### Características conjunto de datos:

- 112.120 vistas frontales de radiografías del tórax
- Radiografías provenientes de 30.805 pacientes
- Cada radiografía etiquetada con 14 posibles enfermedades

### Entrenamiento red neuronal:





- (28.744 pacientes, 98.637 imágenes) para entrenamiento
- (1.672 pacientes, 6.531 imágenes) para validación
- (389 pacientes, 420 imágenes) para tests.
- Algunas imágenes también se giraron aleatoriamente.

# Redes Neuronales Convolucionales (CNN) para Reconocimiento de Rostros



Face Verification and Binary Classification

## Face verification supervised learning

$x$	$y$	
	1	"Same"
	0	"Different"
	0	
	1	

[Taigman et. al., 2014. DeepFace closing the gap to human level performance]

Andrew Ng

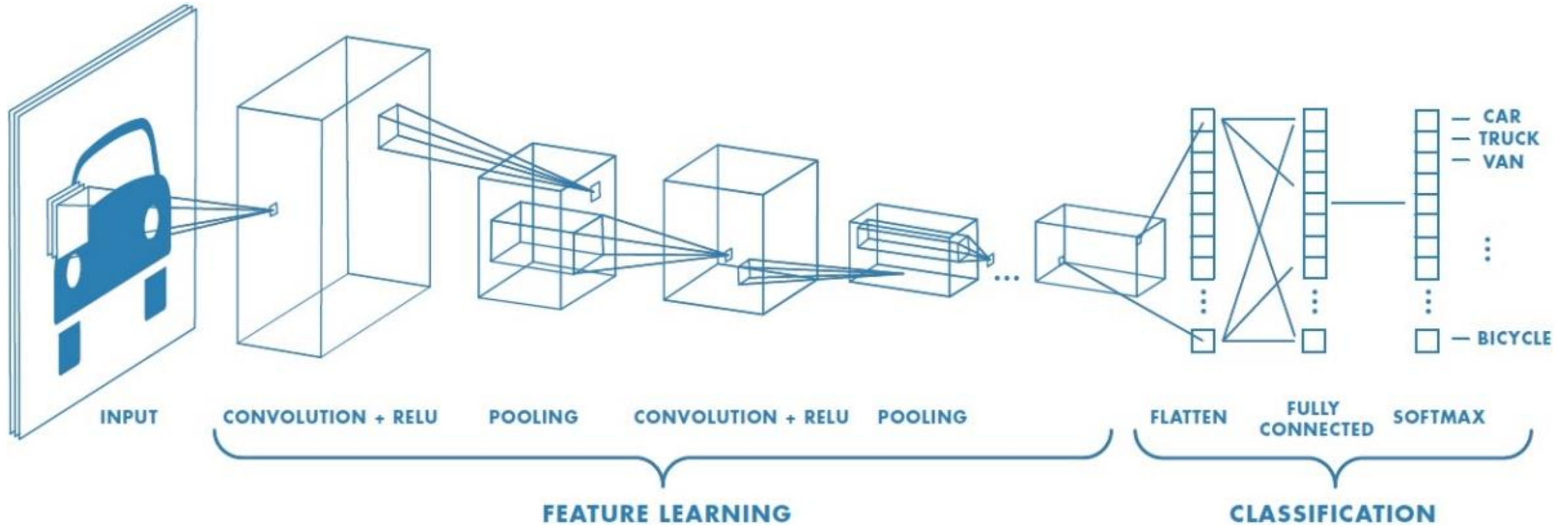
Figura tomada de COURSERA: *deeplearning.ai*

# Redes Neuronales Convolucionales (CNN) para Neural Style Transfer

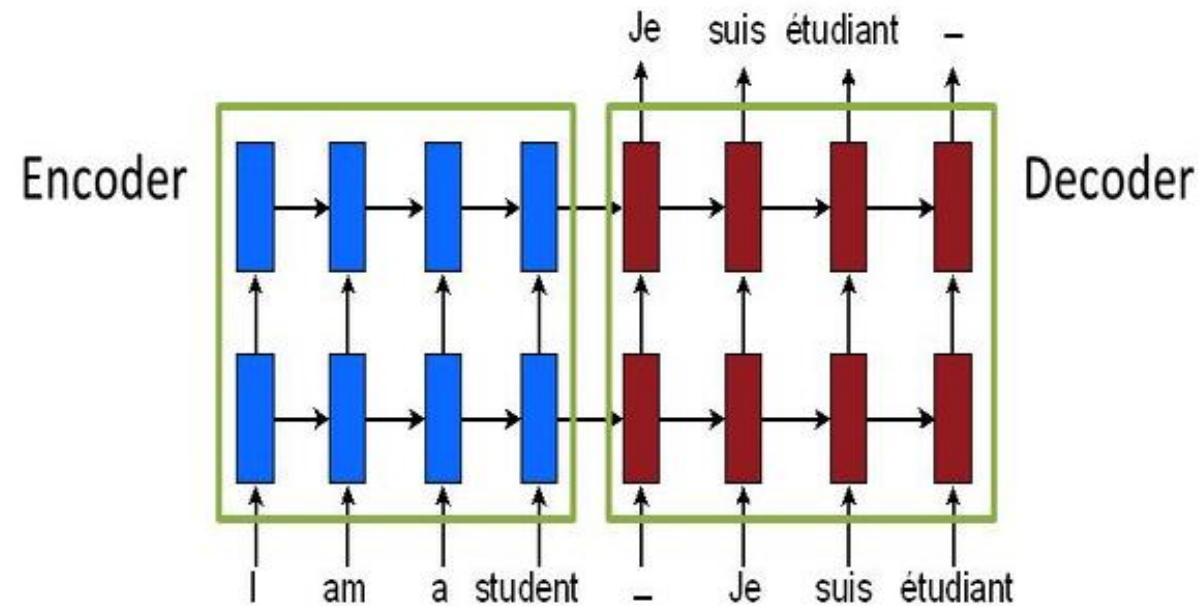


Figura tomada de COURSERA: *deeplearning.ai - Generated\_Image\_Neural\_Style\_Transfer\_Monet*

# CNN – Red Neuronal Convolucional para Reconocimiento de Vehículos



# Redes Neuronales Recurrentes (RNN) para Traducción automática





# Métrica de Desempeño: F1-Score

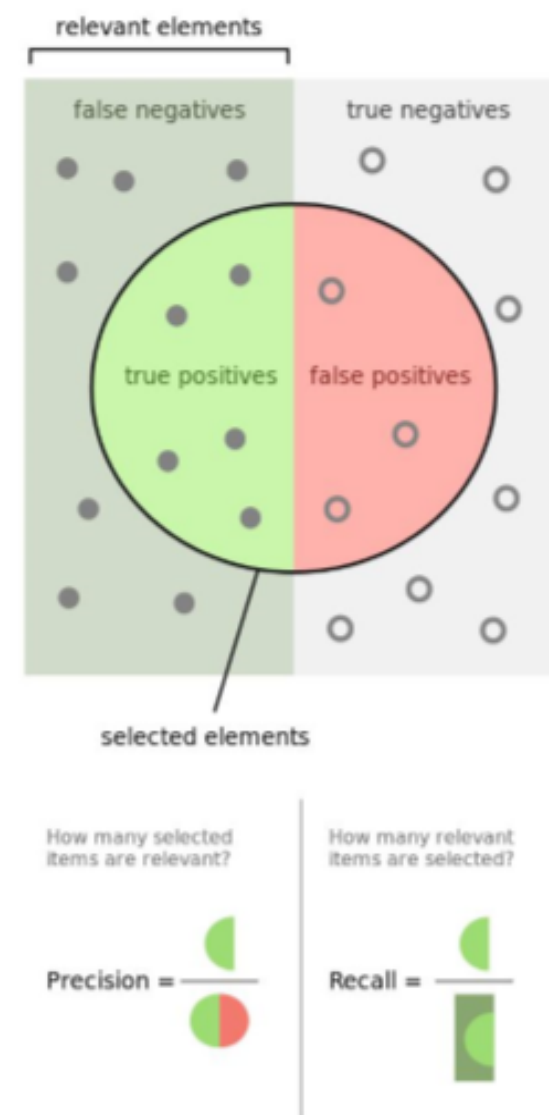
- Es una métrica para el desempeño de una prueba que toma en cuenta la precisión y la sensibilidad (recall) de la prueba, ignora completamente los valores verdaderos negativos

$$F_{\beta} = \frac{(1 + \beta^2) * precision * recall}{(\beta^2 * precision) + recall}$$

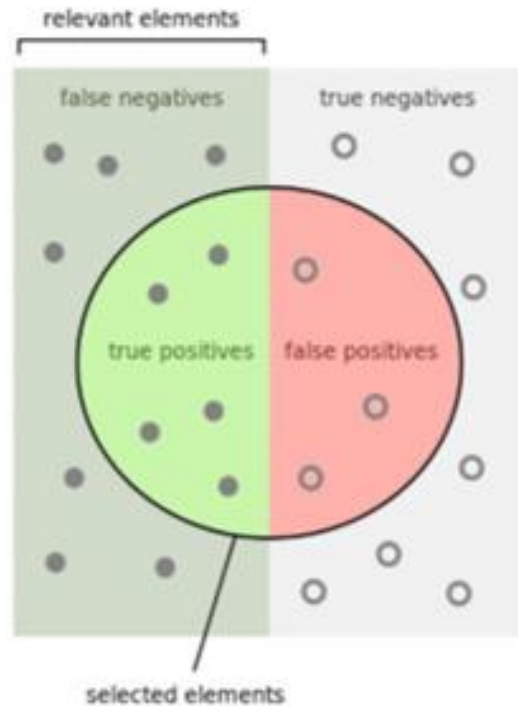
$$F_{\beta} = \frac{(1 + \beta^2) * truePositive}{(1 + \beta^2) * truePositive + \beta^2 falseNegative + falsePositive}$$

- El Beta más utilizado es 1 dado que le da el mismo peso a la precisión y a la sensibilidad

$$F_1 = \frac{2 * precision * recall}{precision + recall} \quad precision = \frac{tp}{tp + fp} \quad recall = \frac{tp}{tp + fn}$$



# Matriz de confusión



How many selected items are relevant?

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}}$$

How many relevant items are selected?

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}$$

		Classifier Prediction	
		Positive	Negative
Actual Value	Positive	True Positive	False Negative
	Negative	False Positive	True Negative

# Ejemplos no reales para cálculo de F1-Score

Prueba para una muestra de 100 radiografías, 10 de ellas de pacientes con neumonía:

- Caso 1: un algoritmo clasifica a los 100 pacientes con neumonía.  
**10 verdaderos positivos, 90 falsos positivos.**  
precisión = 0.1, recall = 1, F1 = 0.18
- Caso 2: un algoritmo clasifica exitosamente un paciente con neumonía.  
**1 verdadero positivo, 9 falsos negativos, 90 verdaderos negativos.**  
precisión = 1, recall = 0.1, F1 = 0.18
- Caso 3: un algoritmo clasifica a los 10 pacientes con neumonía.  
**10 verdaderos positivos, 90 verdaderos negativos.**  
precision = 1, recall = 1, F1 = 1
- Caso 4: un algoritmo clasifica exitosamente 8 pacientes con neumonía y 3 erróneamente. **8 verdaderos positivos, 3 falsos positivos, 2 falsos negativos, 87 verdaderos negativos.**  
precision = 0.727, recall = 0.8, F1 = 0.76

$$F_1 = \frac{2 * precision * recall}{precision + recall} \quad precision = \frac{tp}{tp + fp} \quad recall = \frac{tp}{tp + fn}$$



## Desempeño de los Radiólogos vs CheXNet

- Se usó el conjunto para pruebas de 420 radiografías
- Los radiólogos tenían 4, 7, 25 y 28 años de experiencia.
- Radiólogos y red neuronal etiquetaron radiografías con neumonía.

	F1 Score (95% CI)
Radiologist 1	0.383 (0.309, 0.453)
Radiologist 2	0.356 (0.282, 0.428)
Radiologist 3	0.365 (0.291, 0.435)
Radiologist 4	0.442 (0.390, 0.492)
Radiologist Avg.	0.387 (0.330, 0.442)
CheXNet	0.435 (0.387, 0.481)

# Watson for Oncology : Big Data + Machine Learning



- Sistema de IA en la nube que brinda soporte médico a 11 tipos de cáncer, entre ellos a los de mayor incidencia en Colombia: seno, tiroides, próstata y gastrointestinal (y otros como pulmón, colon, recto, cervical, ovario, vejiga, hígado, tiroides).
- 72.000 pacientes nuevos con cáncer cada año - solo 200 especialistas en el tema, es decir un oncólogo para cada 360 casos nuevos, la capacidad de estos para mantenerse al día con todos los estudios y avances es prácticamente nula.
- Clasificar las opciones de tratamiento, vinculando estudios que han sido investigados y filtrados en 13 países en el mundo para que el Oncólogo considere tratamientos alternativos que están ligados a la ficha médica del paciente.
- Son más de 15'000.000 de páginas de casos médicos puestos a disposición del profesional de la salud para que este, y no la máquina, complemente su conocimiento y practique de mejor manera su profesión.
- Watson no reemplaza al médico, lo apoya. Las decisiones siempre estarán en cabeza del experto.