

# Mod1: Quick review in Deep Learning

Tuesday, April 21, 2020 8:44 AM

## Intro to deep learning

### Why deep learning?

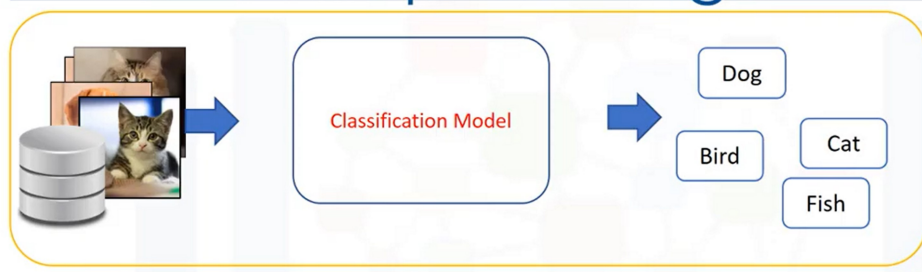


### POR QUE DEEP LEARNING?

La popularidad de Deep Learning se debe a:

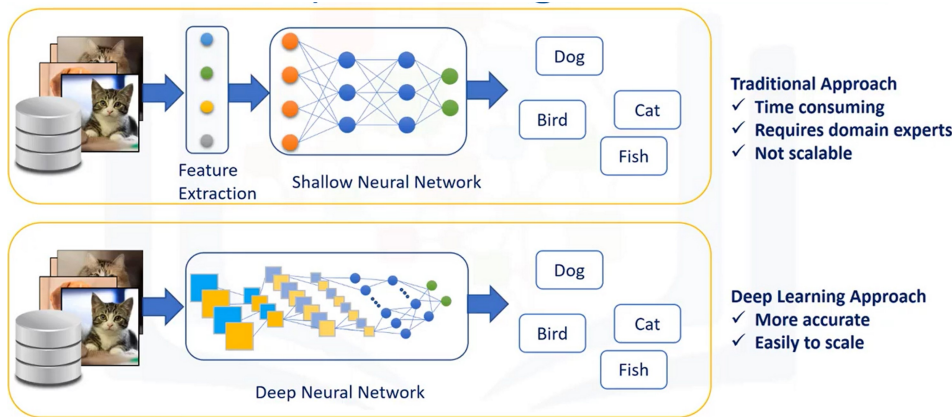
- Incremento en la capacidad de procesamiento de las computadoras
- Disponibilidad de mucha cantidad de datos para entrenar sistemas de computadoras
- Avances en la investigación de algoritmos Machine Learning

### What is deep learning?



Este modelo de clasificación quiere diferenciar gatos de perros. Esto es muy simple para los humanos, pero no para las computadoras.

1. Extracción y selección de atributos: elegir los mejores atributos de cada clase (gatos y perros) para que el algoritmo sea más óptimo. Ej: color, ejes, ubicación de pixels, etc
2. Dar esos atributos a un algoritmo de clasificación (red neuronal)
3. Clasificar clases: perros de gatos



### ENFOQUE TRADICIONAL

Desventajas de este procedimiento:

- Empleo de mucho tiempo
- Se necesitan expertos de un área
- No es escalable

En el proceso de elegir atributos se emplea mucho tiempo porque generalmente tiene que ser el trabajo de un experto para tener buenos resultados en la predicción/clasificación. Además los atributos no pueden generalizarse para otras situaciones. Ej: reconocer números escritos a mano.

### ENFOQUE DEEP LEARNING

Deep learning es una técnica en la que los algoritmos aprenden directamente de las muestras (imágenes de gatos y perros) de manera más efectiva que el enfoque tradicional.

Con este tipo de enfoque se encuentran los mejores atributos en lugar de necesitar a un humano experto.

Para la clasificación de imágenes se usa Redes Neuronales Convolucionales (Convolutional Neural Networks).

## Deep Learning Pipeline

### Deep learning in action

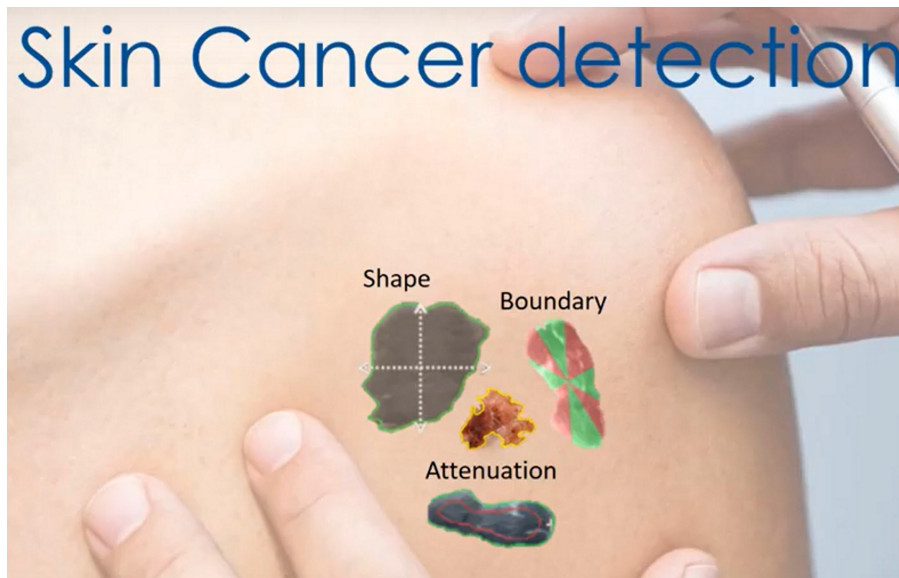
- 5 million cases are diagnosed each year
- 100,000 of these cases involve **melanoma**
- **Melanoma** is the **deadliest form of skin cancer**
  - 9,000 deaths a year in US
- **Catching melanoma early is the key to a patient's survival**
  - Short of specialized physicians
  - Hard to interpret
  - Time consuming



### DEEP LEARNING EN ACCION

Es importante la detección temprana de esta enfermedad. Sin embargo algunos de los problemas actuales son:

- No hay muchos médicos especializados
- Es difícil interpretar los resultados incluso para los especialistas
- Toma tiempo: hay sacar la biopsia, analizar

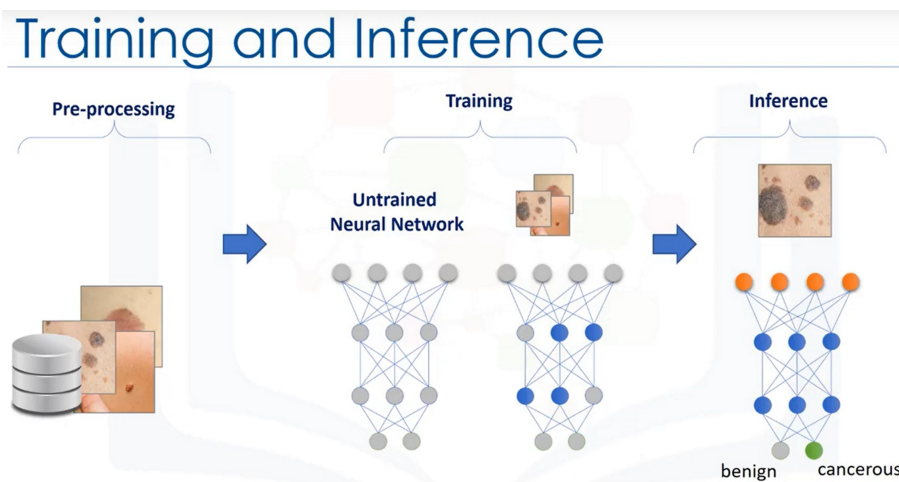


### DETECCION DE CANCER DE PIEL

La solución es Deep Learning para detectar el cancer de piel. Mucho más rápido que un médico con un microscopio.

Un algoritmo capaz de discriminar un tumor benigno de cáncer a partir de una imagen. Este problema es muy similar al anterior: perro vs gato = cáncer vs benigno. Es un problema de clasificación.

El mejor algoritmo para este caso es CNN (Convolutional Neural Network). Tiene que extraer atributos como: ejes, formas, atenuación de color, etc para la clasificación.



### ENTRENAMIENTO E INFERENCIA

El procedimiento para tener un modelo DL listo para el mundo real es:

- Pre-procesamiento  
Las imágenes tienen que convertirse a un formato comprensible para nuestra red.
- Entrenamiento  
La red es alimentada con muchas imágenes para su aprendizaje.
- Inferencia  
El modelo entrenado es usado para clasificar imágenes nuevas. Y por lo tanto puede emplearse para la detección de melanoma en la vida real.

Este procedimiento es muy lento por las siguientes razones:

- El entrenamiento es lento: por la gran cantidad de imágenes y operaciones complejas
- Construir una red DL es un proceso iterativo. Hay que optimizar y afinar. Los científicos de datos tienen que probarlo varias veces antes de que este listo para ser utilizado (inferencia).
- El modelo tiene que actualizarse periódicamente ya que nueva data tiene que agregarse para el entrenamiento



Con una simple arquitectura intel x86, entrenar el modelo puede tomar varios días. El problema está en que el procedimiento DL es iterativo. Si se pudiese acelerarlo se podrían hacer más cosas como:

- Entrenar varios modelos y luego juntarlos en uno para aumentar la precisión ("ensemble model")
- Tener más cuidado en la optimización y afinamiento del modelo
- Entrenar con más frecuencia para que el modelo sea más preciso