

- **Inteligencia Artificial**  
Se refiere a la habilidad de una máquina para realizar tareas que necesitan el nivel de inteligencia de un humano, realizandolas al mismo nivel o mejor que una persona.
- **Machine Learning**
  - Basándose en la Inteligencia Artificial, la idea es que un programa pueda utilizar data para entrenarse y aprender.
- **Deep Learning**
  - Basándose en Machine Learning, la idea es que un programa aprenda de forma no supervisada de data no estructurada y no etiquetada.
- **Artificial Neural Network**
  - Es la forma en la cual un sistema de computadoras se inspira en la red neuronal biológica para poder aprender al tener una serie de nodos conectadas.
- **Supervised Learning**
  - 
  - Entrenamiento de un **modelo** a partir de datos de entrada y sus **etiquetas** correspondientes.  
Labeled  
Unlabeled
- **Unsupervised Learning**
  - Entrenamiento de un **modelo** para encontrar patrones en un conjunto de datos, generalmente sin etiqueta.
- **Reinforcement Learning**
  - Consiste de algoritmos con metas, las cuales al ser cumplidas, indican que la red ha aprendido.
- **Use Cases**
  - los casos que se usan para aprender o entrenar nuevas redes
- **Ejemplos**

Text	Speech	Image	Audio	Content
Classification Prediction	Recognition Text to Speech Speech to Text	Object detection Object localization OCR Gesture Facial modeling Segmentation Clustering Compression Super resolution	Translation Voice Synthesis	Video Generation Text Generation Audio Generation

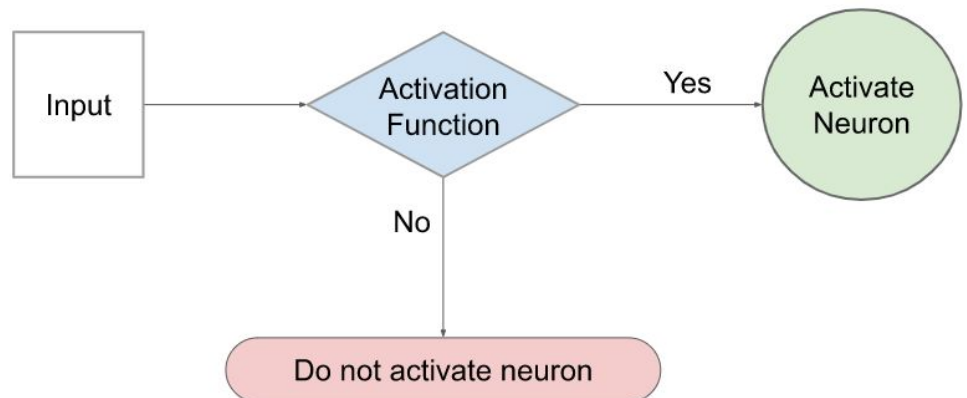
- Idea de implementación



- Features - Input
  - Características específicas de la data, las cuales son clave al momento de procesar la información a través de la red neuronal.
- Labels - Output
  - Es el etiquetado final al momento de dar el output.
  - En el aprendizaje supervisado, parte de "respuesta" o "resultado" de un **ejemplo**. Cada ejemplo de un conjunto de datos etiquetado consiste en uno o más atributos y una etiqueta.
- Modelos
  - Regresión Lineal
    - Tipo de modelo de regresión que da como resultado un valor continuo a partir de una combinación lineal de atributos de entrada.
- Pérdida - Loss
  - Es la diferencia entre el resultado obtenido y el resultado deseado, mientras más cercano a 0, mejor ha aprendido la red neuronal.
  - Medición de la distancia entre las **predicciones** de un modelo y su **etiqueta**.
- Optimizador
  - Implementación específica del algoritmo de **descenso de gradiente**.
- Epoch
  - El ciclo en el cual se recorre la red neuronal, una época indica que s
  - Recorrido de entrenamiento completo por todo el conjunto de datos, de manera que cada ejemplo se observe una vez. Por lo tanto, las repeticiones representan  $N/\text{iteraciones}$  de entrenamiento del **tamaño del lote**, donde  $N$  es el número total de ejemplos.
  - 
  -
- Stochastic Gradient Descent (SGD)
  - Es una forma de optimizar la red neuronal para obtener el mejor gradient descent, conforme cada iteración, se ingresa data distinta, pero se da un ajuste del gradiente.
  - En otras palabras, el SGD se basa en un solo ejemplo elegido al azar de un conjunto de datos de maera uniforme para calcular una estimación del gradiente en cada paso.
- Training
  - Proceso de determinar los **parámetros** ideales que conforman un modelo.
- Inference

- En aprendizaje automático, suele hacer referencia al proceso de realizar predicciones mediante la aplicación del modelo entrenado a [ejemplos sin etiqueta](#).
- Transfer Learning
  - Transferencia de información de una tarea de aprendizaje automático a otra.
  - El aprendizaje por transferencia puede implicar la transferencia de conocimientos desde la solución de una tarea más simple hasta otra más compleja, o la transferencia de conocimientos de una tarea en la que hay más datos a otra en la que hay menos.
- TensorFlow - Keras
  - API de aprendizaje automático muy utilizada de Python. [Keras](#) se ejecuta en diversos entornos de aprendizaje profundo, incluido TensorFlow, donde está disponible como [tf.keras](#).
- Dataset
  - Conjunto de datos, los cuales son utilizados para poder entrenar a la red neuronal.
- Validation
  - Subconjunto del conjunto de datos, separado del conjunto de entrenamiento, que se usa para ajustar [hiperparámetros](#).
- Test
  - Subconjunto dentro del conjunto de datos que se usa para probar un [modelo](#) después de que este pasó por la evaluación inicial a través del conjunto de validación.
- Activation Functions
  - sigmoide
    - Atributo que asigna resultados de regresión multinomial o logística (probabilidades logísticas) a probabilidades, y devuelve un valor entre 0 y 1.
    - 
    - 
    - 
    -

## Perception of Artificial Neuron



- 
- softmax
  - Función que proporciona probabilidades para cada clase posible en un **modelo de clasificación de clases múltiples**.
- relu

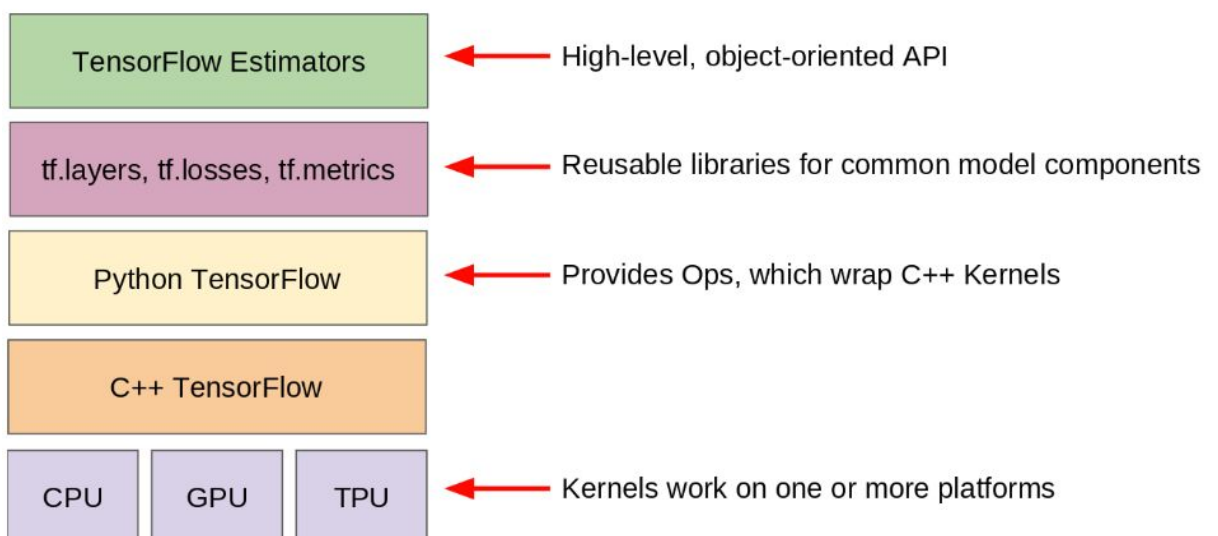
**Atributo de activación** con las siguientes reglas:

- Si la entrada es negativa o cero, el resultado es 0.
- Si la entrada es positiva, el resultado es igual a la entrada.
- 

## Perceptron

- Es la neurona, la cual contiene una entrada con pesos, la función de activación y sus salidas.
- Densely Connected
  - Cuando todas las neuronas de cada capa están conectadas unas con otras.
- Sparsely Connected
  - Cuando ciertas neuronas de cada capa están conectadas unas con otras.
- Cross Entropy
  - Una generalización de pérdida logística en problemas de clasificación de clases múltiples. La entropía cruzada cuantifica la diferencia entre dos distribuciones de probabilidad. Consulta también perplejidad.
- Feature Extraction
  - Proceso en el que se determina qué **atributos** podrían ser útiles para entrenar un modelo y luego convertir los datos sin procesar de los archivos de registro y otras fuentes en dichos atributos.
- Fine Tuning

- Utilizar un modelo ya entrenado para un propósito, y hacer que haga otro propósito similar al cual fue entrenado.
- Overfitting
  - Cuando la red neuronal se “aprende” las respuestas del test, por lo que al darle data distinta, no va a poder operarla correctamente.
- Scaling
  - Llevar un dato de un tamaño, a uno de menor o mayor tamaño para poder facilitar su procesamiento.
- Binning
  - Conversión de un atributo (generalmente **continuo**) en varios atributos binarios denominados agrupamientos o discretizaciones, que en general se basan en un rango de valores.
- Scrubbing
  - Datasets can be unreliable because
    - omitted values
    - duplicate examples
    - bad labels
    - bad feature values
- TF Estimators



Toolkit(s)	Description
Estimator (tf.estimator)	High-level, OOP API.
tf.layers/tf.losses/tf.metrics	Libraries for common model components.
TensorFlow	Lower-level APIs

- Multilabel Classification
- Text Representation
- One-hot encoding
- Encoding

- Embeddings
- Convnet
  - Es capaz de “recordar” información espacial como color y forma, puede hacerlo en una imagen completa o en subregiones de píxeles
  - Los features son jerárquicos, se componen los high-complexity features de los low-complexity features en lugar de intentar aprender high complexity de un solo
  - Los features son traslacionales e invariantes, si funcionan en (x,y) podrán computar en sus primas
- Vanishing Gradients
  - Sucede cuando el vector gradiente decrece exponencialmente, por lo que toma mucho tiempo de procesamiento y nunca llega al valor deseado.
- Exploding Gradients
  - Sucede cuando el vector gradiente crece exponencialmente, por lo que ya no es útil ya se pasa de nuestro valor deseado.
- Dead RELU Units
  - Cuando un nodo con función relu siempre recibe un valor negativo y se queda trabado en 0.
- Dropout
  - forma de regulizacion oara evitar overfitting
- Data Augmentation
  - Utilizar un dataset y aplicar diferentes modificaciones a la data para aumentar la cantidad de datos total con un solo dataset.
- Convolutional Neural Networks
- Fully Connected Net
- Convolution
  - Consiste de realizar filtros, donde se obtienen valores característicos de la data respecto al tipo de filtro.
- Padding
  - Ajustar la data para que toda sea igual, en imágenes, consiste de llevarlas a un mismo tamaño, en texto, consiste de llevar el string a un mismo tamaño.
- Receptive Field
  - subregion of image is sent to the neurons as input instead of whole image
- Conv Layer
- Pooling
  - Reducir una matriz (o matrices) creada por una **capa convolucional** anterior a una matriz más pequeña.
- Sequences
- Tanh activation
  - valores van de [-1,1]
- Long short-term memory
- sigmoid
  - [0,1]
- forget gate
  - el gate decide que se queda y que olvida. La información incluye la información previa escondida y el input actual

- input gate
  - regula usando la funcion de activacion de tanh y decide que es importante usando la función de activacion sigmoide
- cell state
  - usando el forget gate para botar información y generar un nuevo estado de cell
- output gate
  - generamos un output que sera el input escondido de la siguiente cell
- LSTM
  - forget gate decide que es relevante del estado previo
  - input gate decide que es relevante del input actual
  - output gate determina que sera el proximo hidden state
- Gated recurrent unit (GRU)
  - reset gate se usa para decidir cuanta información pasada guardar
  - update gate funciona como el forget y input gate del LSTM
- Manual feature extraction
  - Domain Knowledge
  - Define features
  - Detect features to classify
- Maximum likelihood
- Filters to extract features
- CNNs for classification
  - convolucion aplica filtros con pesos predefinidos para generar mapas de features
  - no linealidad - Relu
  - Pooling obtener los valores mas relevantes de la imagen
- beyond classification
  - semantic segmentation
  - object detection
  - image captioning
- Visualize data
  - preprocess
    - normalize
    - transform
  - define a model
  - train your model
    - define loss
    - optimization
  - save the best model
  - test your model
- backpropagation
  - un input y un output
  - los hidden nodes conectados por pesos
  - se usa una funcion de activacion no lineal para evitar que red neural solo aprenda modelos lineales, generalmente se usa sigmoide

- funcion de error para medir que tan lejos estamos de obtener los pesos correctos
- forward propagation tenemos un ejemplo inicial y actualizamos la capa de input de la red. consideramos a esta capa ser otro nodo pero sin funcion de activacion
  - actualizamos la primera hidden layer y usamos los pesos para computar la siguiente capa
  - actualizamos el output de los nodos en la primera hidden layer, aqui usamos la funcion de activacion
  - propagamos al resto de la red
- Error derivativo, el algoritmo de backpropagation decide cuando actualizar cada peso despues de comparar el output predicho con el deseado.
- regresamos chequeando este error desde el valor predicho hasta el input
- que es bias
  - Una intersección o un desplazamiento del origen.
- que son las hidden layers
  - Capa sintética en una **red neuronal** entre la **capa de entrada** (es decir, los atributos) y la **capa de salida** (la predicción). Una red neuronal contiene una o más capas ocultas.
- que es una deep neural network
- que parametros aprende
- conv = filtros, kernel size que es el tamaño del filtro, padding y activation
- que es empirical loss
  - Elección de la función del modelo que minimiza la pérdida en el conjunto de entrenamiento.
- que es loss optimization
- que es clipping
- que son conexiones residuales
- que hacen los seeds
- que es el forward pass
- full batch gradient descent
- mini batch
- Earlystopping no pasarme de epochs - codigo usamos call backs
  - Método de **regularización** que implica finalizar el entrenamiento del modelo *antes* de que la pérdida de entrenamiento deje de disminuir.
- data prep
  - limpiar data
  - normalizar la data
  - padding
  -