



Machine Learning: de la teoría a la práctica

Andrés Ramírez-Morales

e-mail: andres.ramirez.morales@cern.ch

Universidad de Londres

08-12 de Julio 2019

Lo que este curso es

- Es una introducción a la programación científica
 - Pretende dar las herramientas necesarias para que el participante interesado siga aprendiendo
 - Introducir conceptos de la teoría de Machine Learning
 - Trabajar con ejemplos simples de Machine Learning

Lo que este curso **NO** es

- Un curso completo de computación (se requieren un par de años para dominar el arte)
 - Un curso de Inteligencia Artificial (solo se toca los puntos claves relacionados a ML)
 - Un curso de álgebra y estadística avanzada (se explica la teoría brevemente pero se usarán librerías de Python ya establecidas)

Al finalizar el curso

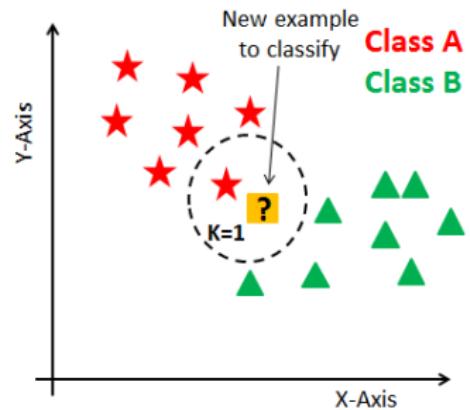
- El participante tendrá una noción consistente de la importancia de la computación en la ciencia y la tecnología
- El participante que no ha usado software libre, y su ideología, apreciará la importancia del esfuerzo global que existe para crear y mantener software libre y útil para todos
- Se busca que se tenga una idea clara del poder de la inteligencia artificial en los avances de la tecnología y el conocimiento

¿Qué es Machine Learning?

- Sabemos que los humanos aprenden de experiencias pasadas
 - Las máquinas siguen las instrucciones dadas por los humanos
 - Imagina que los humanos pueden entrenar máquinas a aprender de experiencias pasadas (datos) y hacer lo que los humanos hacen mucho más rápido
 - Hablar de Machine Learning is hablar de Inteligencia Artificial

¿Qué es Machine Learning?

- ### ■ Ejemplo: algoritmo de los vecinos más cercanos



Instalación

- El uso de Linux debe ser una herramienta básica para cualquier científico/ingeniero
- El uso de Python ha venido mejorando significativamente el desempeño de la computación



Instalación: Hola mundo (Hello world)

- Uso de python: lenguaje de programación orientada a objetos
- Después de Hello world, veamos el poder de Python
- Celebremos



Machine Learning

¿Para qué usar Machine Learning?

- PayPal, detener estafas: fácil de usar, automático y confiable



Aplicaciones de Machine Learning

- Buscadores, reconocimiento de voz, reconocimiento de placas, lector de sueños
- Machine learning nos facilita la vida, hace los procesos más consistentes y confiables



Search Engine
Result



Voice
Recognition



Number Plate
Recognition



Dream Reader

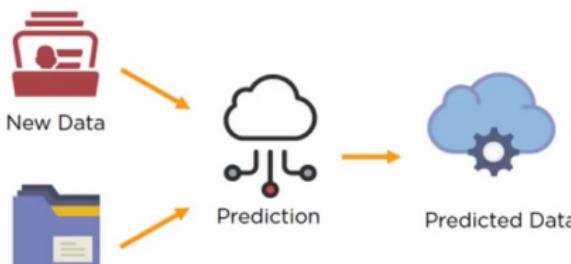
¿Cómo funciona Machine Learning?

■ Fase 1: Aprendizaje

- Pre-procesamiento: limpiar los datos, darle formato a los datos
- Aprender, tomar esos datos y aprender de ellos (supervisado y no supervisado)
- Después de entrenar tu algoritmo, hay que verificar que da las respuestas correctas

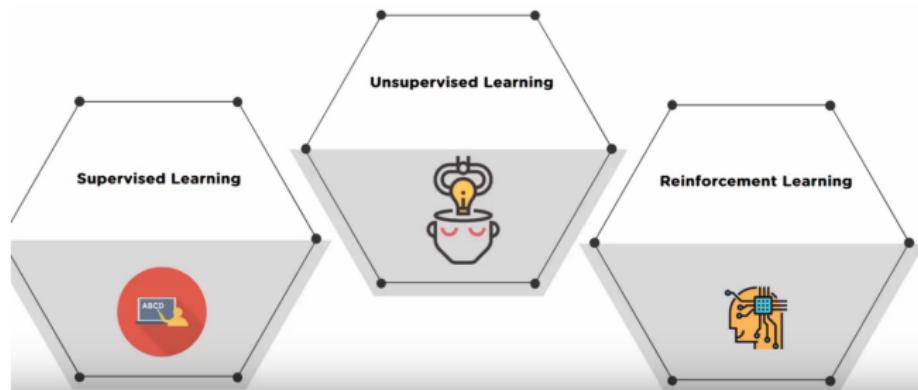
■ Fase 2: Predicciones

- Nuevos datos, model entrenado, predicción, datos predichos las respuestas correctas



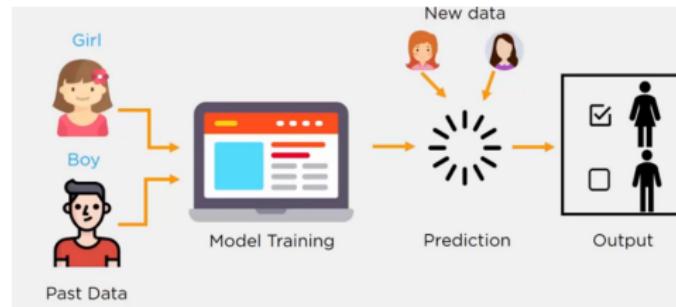
Tipos de Machine Learning

- Aprendizaje supervisado
- Aprendizaje NO supervisado
- Aprendizaje reforzado



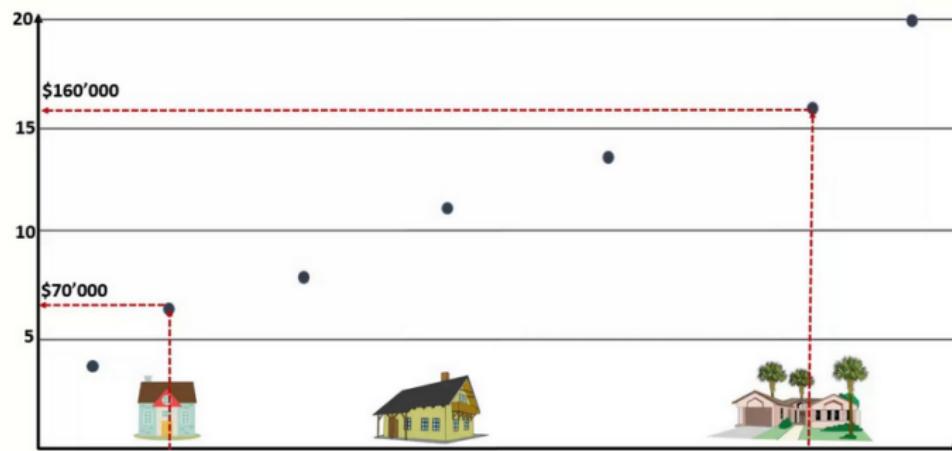
Aprendizaje supervisado

- El modelo de Machine Learning aprende de datos insertados en el pasado y hace futuras predicciones como salida
- **Tipos de Aprendizaje supervisado**
 - Clasificación: separar datos en diferentes tipos, (árboles de decisión o support vector machine)
 - Regresión: basado en datos anteriores, la máquina predice un continuo de valores (regresión lineal, regresión polinomial)



Aprendizaje supervisado: Regresión lineal

- El precio de una casa depende del tamaño
- Usar los datos anteriores para predecir el precio de una casa de cualquier tamaño

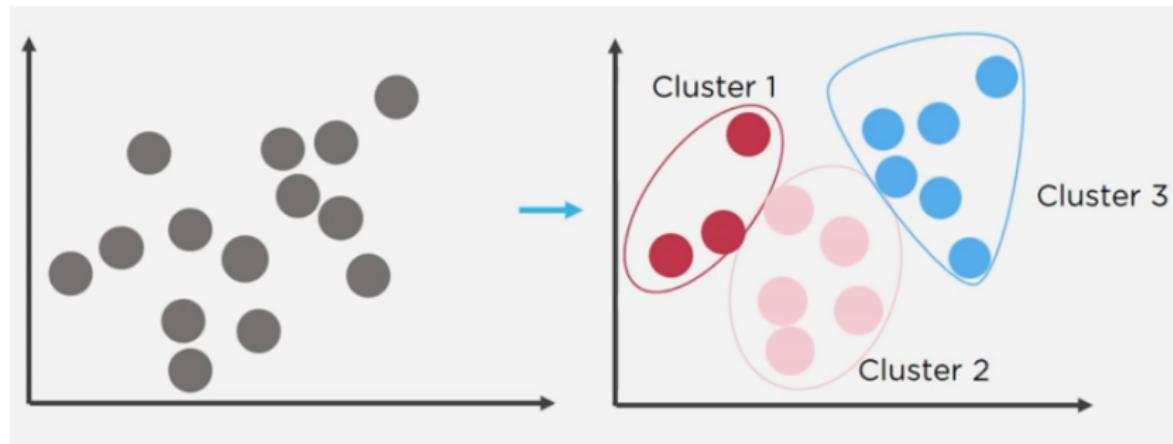


Aprendizaje NO supervisado

- Significa que la máquina aprende utilizando datos insertados sin etiquetar
- Permite que el algoritmo actue sobre los datos sin ninguna guía
- Recordemos que para el caso de aprendizaje supervisado, teníamos una respuesta, aquí no la tenemos, es decir permitimos que el algoritmo obtenga las respuestas para nosotros
- **Tipos de Aprendizaje NO supervisado**
 - Clustering: Es usado para analizar y agrupar datos (k-means, clustering jerárquico, modelo escondido de Markov)
 - Asociación: descubre la posibilidad de la ocurrencia asociada en una colección de objetos (Algoritmo Apriori, y crecimiento-FP)

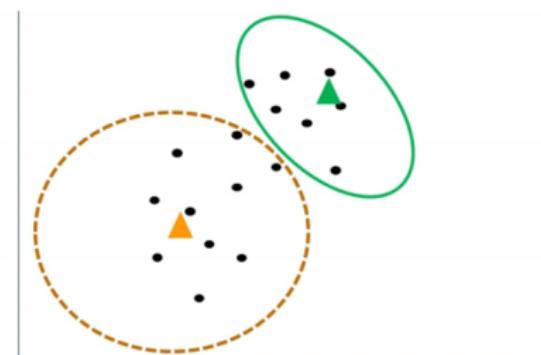
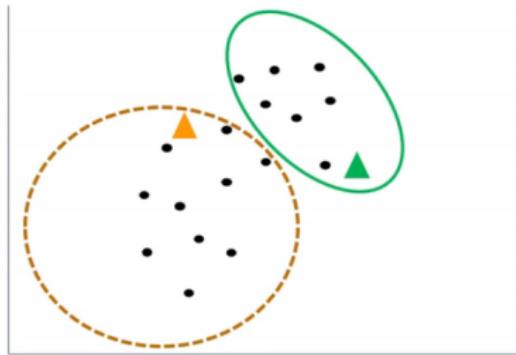
Aprendizaje NO supervisado: k-means clustering

- La idea es tomar varios similares objetos y clasificarlos en clusters



Aprendizaje NO supervisado: k-means clustering

- **Paso 1:** Iniciar aleatoriamente dos puntos llamados centroides
- **Paso 2:** Basado en la distancia de los centros se hará un grupo
- **Paso 3:** Mover los centroides un poco para ajustarlos para optimizar, repetir esto hasta que los centroides dejen de cambiar de posición, es decir que los grupos estén bien definidos (convergencia)

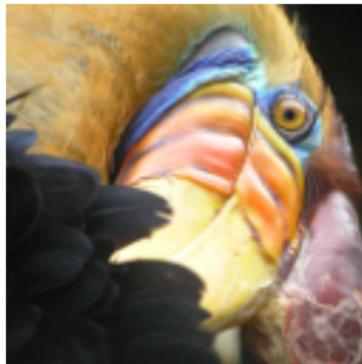


k-means clustering: Análisis del clima

- **Problema:** predecir el clima usando Machine Learning
- **Posibles retos:**
 - Identificar las áreas donde hay lluvia
 - Considerar la temperatura y la presión como parámetros: el algoritmo k-means indetificará las regiones que están expuestas a fuerte lluvia

k-means clustering: Comprimir una imagen

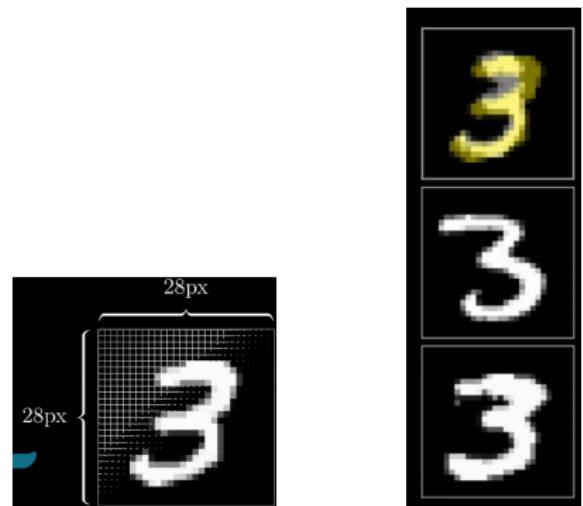
- **Problema:** comprimir una imagen usando Machine Learning
- **Posibles retos:**
 - Identificar las colores principales que representen a la imagen
 - Mapear los colores 24-bit a una dimensión más baja usando asignación de clusters



Neural Networks

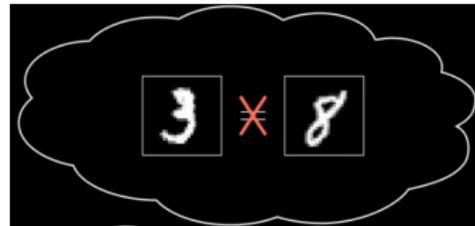
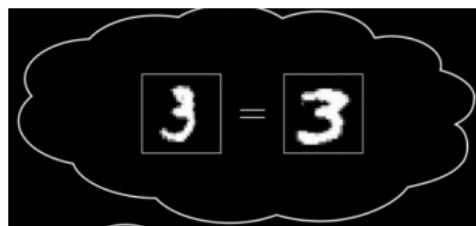
Neural Networks

- **Problema:** identificar un numero tres manuscrito con una computadora
- **Possibles retos**
 - Tu cerebro no tiene problema alguno para reconocerlo como un tres
 - La luz que llega de estos tres's es muy diferente pero algo en nuestra inteligente corteza cerebral los identifica como (la misma idea) tres's apesar de que el acomodo de sus pixeles es muy diferente



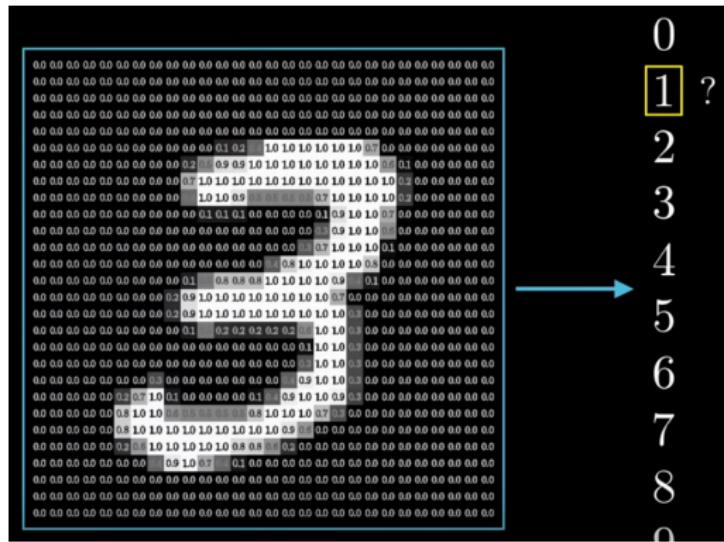
Neural Networks

- Al mismo tiempo reconoce otras imágenes como otras ideas



Neural Networks

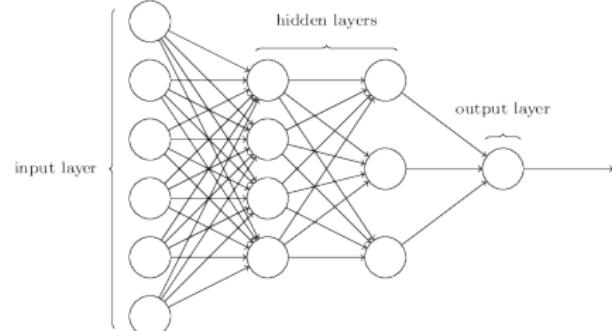
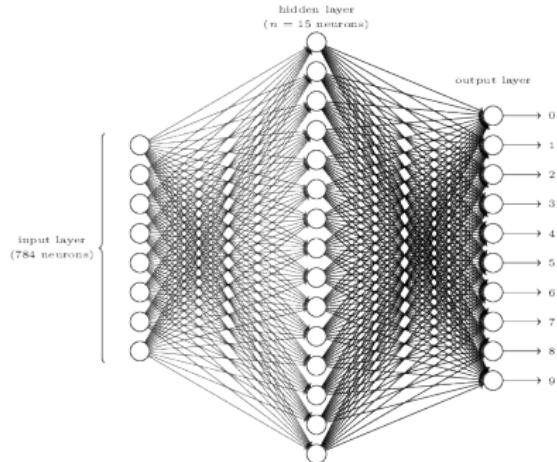
- Escribir un programa que tome una red de 28×28 pixeles y que de una buena suposición de que número se trata
- **Solución** Neural Networks



¿Qué es una Neural Network?

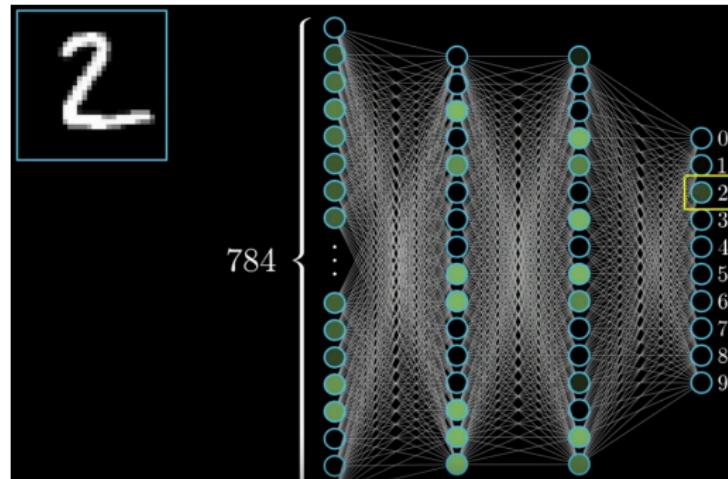
- Neural Networks se expresan matemáticamente:

$$a_{l+1} = \sigma(W_l a_l + b_l)$$



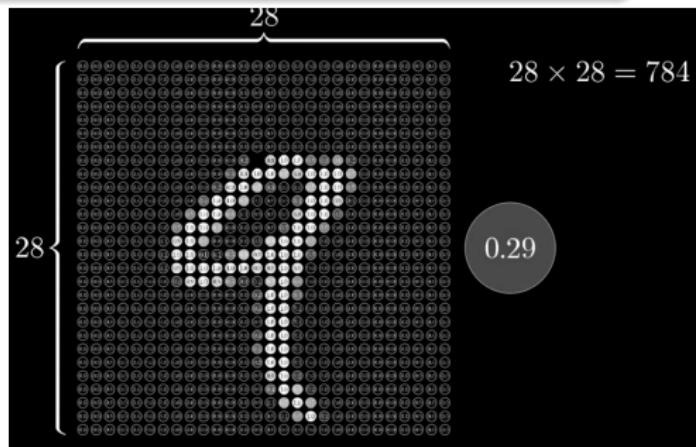
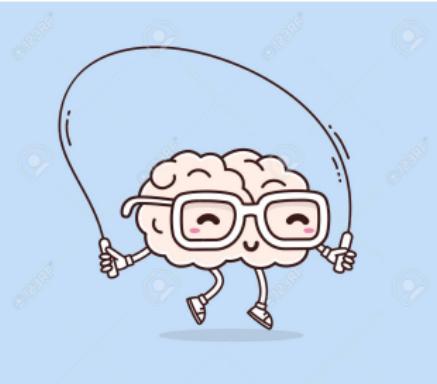
¿Qué es una Neural Network?

- Percepción multicapa, esta es la forma más simple de NN, pero aún así puede aprender a reconocer números manuscritos



¿Qué es una Neural Network?

- Está inspirada por el cerebro:
 - **Neural:** ¿Cuáles son las neuronas?
Neurona: algo que guarda un número entre 0 y 1:
número de activación
 - **Networks (redes):** ¿En qué sentido están conectadas?



¿Qué es una Neural Network?

- Cada pixel de la imagen anterior representa una neurona de la primera capa de la red de neuronas
- La última capa tiene 10 neuronas que representan cada dígito, la activación en estas neuronas representa que tanto el sistema piensa que una imagen dada corresponde a un dígito dado
- La forma en que la red opera es que los número de activación en una capa determina la activación de la siguiente capa: el corazón del método es el mecanismo de procesamiento de como estos números de activación activan a los de la siguiente capa (parecido a nuestro cerebro)
- La red que vamos a trabajar ya ha sido entrenada de como reconocer dígitos: si se inserta una imagen (784 neuronas) y siguiendo los pixeles en la imagen ese patrón de activaciones causa activaciones muy específicas en las siguientes capa hasta llegar a la capa final y la neurona más brillante nos dice que dígito esta imagen representa

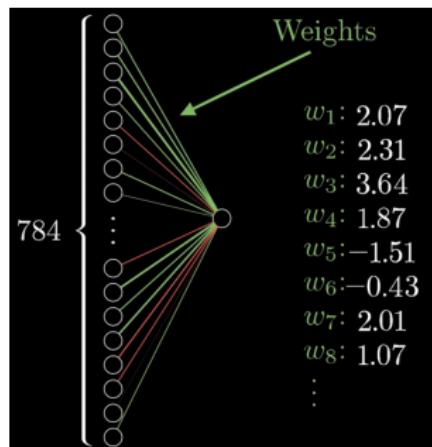
¿Qué es una Neural Network?

- ¿Por qué estructura de capas y por qué se comporta inteligente?
- Nuestro ojo reconoce óvalos y palitos, esperaríamos que cada neurona de la segunda a la última capa, corresponda con cada una de estas subcomponentes
- Requiere aprender que combinación de subcomponentes corresponde a que dígito
- Reconocer un óvalo, reconocer los pequeños bordes
- No solo reconocimiento de imagen, si no discurso de personas: entra audio crudo y el sonido se descompone en palabras y las palabras en letras

¿Qué es una Neural Network?

- Como las activaciones en una capa podrían determinar la activacion de las siguientes un mecanismo que:

pixels → bordes → patrones → dígitos

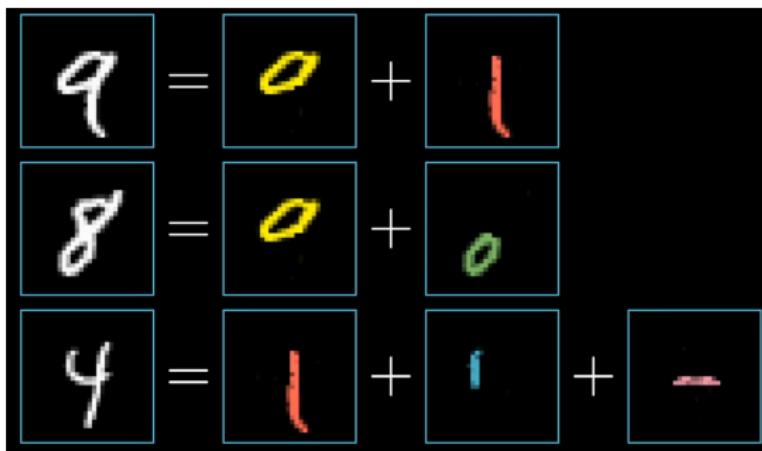


Obtener la suma ponderada de los números de activación

$$w_1a_1 + w_2a_2 + w_3a_3 + \dots + w_na_n$$

Queremos que las activaciones [0,1], apretar la suma un número no cualquiera

¿Qué es una Neural Network?



¿Qué es una Neural Network?

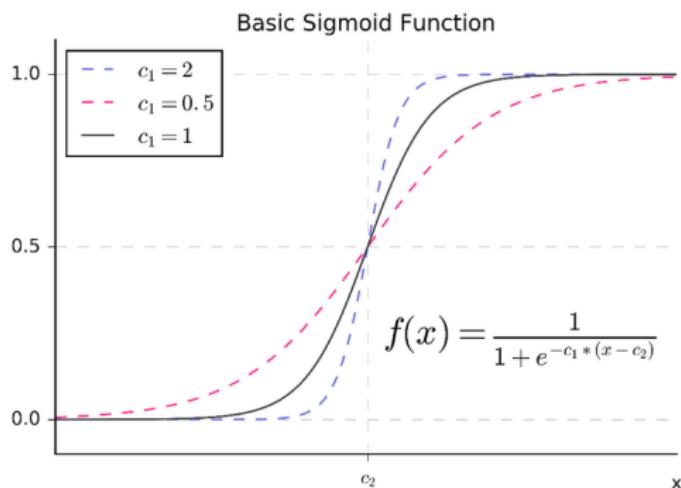
- Usamos la función Sigmoid:

$$\sigma(w_1a_1 + w_2a_2 + w_3a_3 + \dots + w_na_n)$$

- La activación de la neurona es la medida de que tan positiva es la suma ponderada
- Se necesita una adición para la inactividad, *i.e.* quizás no interesa la neurona activa arriba de cero, > 10

$$\sigma(w_1a_1 + w_2a_2 + w_3a_3 + \dots + w_na_n - 10)$$

¿Qué es una Neural Network?



¿Qué es una Neural Network?

- Los pesos (weights) dicen que patrón de pixeles esta neurona en la segunda capa está siguiendo y la adición dice que tan grande la suma ponderada necesita ser antes de que las neuronas empiecen a ser significativamente activas
- Cuando hablamos de aprendizaje estamos hablando de encontrar los pesos y las adiciones para esta gran cantidad de números para que resuelvan el problema

$$a_0^1 = \sigma(w_{0,0}a_0^{(0)} + w_{0,1}a_1^{(0)} + \dots + w_{0,n}a_n^{(0)} + b_0)$$

¿Qué es una Neural Network?

- Representación matricial

$$\sigma \left(\begin{pmatrix} w_{0,0} & w_{0,1} & \cdot & \cdot & \cdot & & w_{0,n} \\ w_{1,0} & w_{1,1} & \cdot & \cdot & \cdot & & w_{1,n} \\ \cdot & & \cdot & & & & \\ \cdot & & & & & & \\ \cdot & & & & & & \\ w_{k,0} & w_{k,1} & \cdot & \cdot & \cdot & & w_{k,n} \end{pmatrix} \begin{pmatrix} a_0^{(0)} \\ a_1^{(0)} \\ \cdot \\ \cdot \\ \cdot \\ a_n^{(0)} \end{pmatrix} + \begin{pmatrix} b_0^{(0)} \\ b_1^{(0)} \\ \cdot \\ \cdot \\ \cdot \\ b_n^{(0)} \end{pmatrix} \right)$$

- En forma compacta la transición de una capa a la que sigue:

$$\mathbf{a}^{(1)} = \sigma(\mathbf{W}\mathbf{a}^{(0)} + \mathbf{b})$$

- Mejor pensamos que una neurona es más bien una función (la red entera es una función, muy complicada con muchos parámetros1)

Recapitulación

- Se introdujo el concepto de Machine Learning
- Se revisó la importancia del uso de Machine Learning en la ciencia e ingeniería

PREGUNTAS