

## Los comienzos (décadas de 1950 a 1980)

cubriremos el nacimiento de las redes neuronales con el **Perceptron en 1958**, el invierno de la IA de los años 70 y el regreso a la popularidad de las redes neuronales con la retropropagación en 1986.

na función es diferenciable si es una línea suave y agradable: el Perceptrón de Rosenblatt calculó la salida de tal manera que la salida saltó abruptamente de 0 a 1 si la entrada excedía algún número, mientras que Adaline simplemente emitió la entrada, lo cual fue un buen no. -línea de salto.

## El deshielo del invierno de la IA

Minsky y Papert no solo mostró la imposibilidad de calcular XOR con un solo perceptrón, sino que argumentó específicamente que tenía que hacerse con múltiples capas de perceptrones, lo que ahora llamamos redes neuronales multicapa

La retropropagación en redes neuronales Paul Werbos fue el primero en EE. UU. en proponer que podría usarse para redes neuronales después de analizarlo en profundidad en su tesis doctoral de **1974**

1970 hubo falta de interés académico que no fue hasta más de una década después, en **1986** "[Learning internal representations by error propagation](#)" de David Rumelhart, Geoffrey Hinton y Ronald. Williams destaca por la forma concisa y clara en que se plantea la idea entrenar las redes neuronales multicapa para abordar problemas de aprendizaje complejos.

## Parte 2: Florecimiento de las redes neuronales (décadas de 1980 a 2000)

### Las redes neuronales ganan visión

En **1989** Las redes feedforward multicapa son aproximadores universales". Esencialmente, probó matemáticamente que múltiples capas permiten que las redes neuronales implementen teóricamente cualquier función, y ciertamente XOR.

n 1989, Yann LeCun et al. en AT&T Bell Labs demostró una aplicación real muy significativa de la retropropagación en [""Backpropagation Applied to Handwritten Zip Code Recognition](#)

a idea de convolución se denominó 'peso compartido', y en realidad se discutió en el análisis ampliado de **1986** de propagación hacia atrás por Rumelhart, Hinton y Williams

LeCun continuó siendo un gran defensor de las CNN en Bell Labs, y su trabajo en ellas resultó en un uso comercial importante para la lectura de cheques a a fines de la década de **1990**, uno de estos sistemas estaba leyendo del 10 al 20% de todos los cheques en los Estados Unidos"

### Las redes neuronales no están supervisadas

Aprendizaje supervisado y no supervisado

### Creencias de ganancia de redes neuronales

aprendizaje basado en la energía , "Algoritmo de aprendizaje para máquinas de Boltzmann" de **1985** por Hinton , red de creencias ,tener conjuntos separados de pesos para inferir variables ocultas a partir de variables visibles ( **pesos de reconocimiento** ) y viceversa ( **pesos generativos** )

### Las redes neuronales toman decisiones

el aprendizaje por refuerzo, nforme de **1990** "Identificación y control de sistemas dinámicos mediante el uso de neuronas". redes"

. Tal vez como era de esperar, había otro campo separado del aprendizaje automático en el que las redes neuronales eran útiles: la robótica. Un ejemplo importante del uso temprano de redes neuronales para robótica provino del NavLab de CMU con ["Alvinn: un vehículo terrestre autónomo en una red neuronal" de 1989.](#)

**1993** ["Aprendizaje por refuerzo para robots que utilizan redes neuronales"](#)

red neuronal que aprendió a ser un jugador de backgammon de clase mundial. Apodada [TD-Gammon](#)

### **Las redes neuronales se vuelven locas**

las redes neuronales recurrentes (RNN)

Bengio escribió en 1993 ["Un enfoque conexionista para el reconocimiento de voz".](#)

Yann LeCun estaba trabajando con CNN allí. De hecho, en 1995 coescribieron el artículo resumen ["Redes convolucionales para imágenes, voz y series temporales"](#)

### **Las Redes Neuronales Empiezan a Hablar**

Bengio de 2003 [" Un modelo de lenguaje probabilístico neural "](#). Como su título lo indica, este trabajo tuvo que ver con el uso de redes neuronales para hacer modelos de lenguaje,

### **Un nuevo invierno amanece**

Schmidhuber y Hochreiter introdujeron un concepto muy importante en 1997 que esencialmente resolvió el problema de cómo entrenar redes neuronales recurrentes, al igual que lo hicieron las CNN para las redes neuronales de avance: [memoria a largo plazo](#) (LSTM)

## **Parte 3: aprendizaje profundo (décadas de 2000 a 2010)**

### **La financiación de más capas**

Hinton, Simon Osindero y Yee-Whye Teh publicaron un artículo en 2006 que fue visto como un gran avance, un avance lo suficientemente significativo como para reavivar el interés en las redes neuronales: [un algoritmo de aprendizaje rápido para la investigación profunda . redes de creencias](#)"aprendizaje profundo

redes neuronales con muchas capas realmente podrían entrenarse bien, si los pesos se inicializan de una manera inteligente en lugar de aleatoriamente

La idea básica es entrenar cada capa una por una con un entrenamiento no supervisado, que comienza con los pesos mucho mejor que solo darles valores aleatorios y luego terminar con una ronda de aprendizaje supervisado, como es normal en las redes neuronales.

### **El desarrollo de los grandes datos**

conjunto de datos tenía 3,2 millones de imágenes para 5247 conceptos, todavía muy lejos del objetivo final de 50 millones de imágenes,

### **La importancia de la fuerza bruta**

to ["Aprendizaje profundo no supervisado a gran escala usando procesadores de gráficos"](#)

del mismo año sugiere un número: 70 veces más rápido. Sí, 70 veces, reduciendo las semanas de trabajo a días

["Deep Big Simple Neural Nets Excel on Handwritten Digit Recognition" de 2010.](#) notablemente coescrito por J. Schmidhuber, uno de los inventores de las redes LSTM recurrentes), que mostró que se podía lograr una enorme tasa de error de %0.35 en el conjunto de datos MNIST sin nada más especial que redes neuronales realmente grandes

Microsoft y Google, así como el laboratorio de IBM y Hinton, dio como resultado el impresionante título ["Redes neuronales profundas para el modelado acústico en el reconocimiento de voz: las opiniones compartidas de cuatro grupos de investigación".](#)

en 2012.

En 2011 Dean y, junto con Ng, formaron Google Brain, un esfuerzo por construir redes neuronales verdaderamente gigantes y explorar lo que podían hacer. El trabajo dio como resultado un aprendizaje de redes neuronales no supervisado de una escala sin precedentes: 16 000 núcleos de CPU

### **La ecuación del aprendizaje profundo**

Xavier Glort y Yoshua Bengio a escribir ["Comprender la dificultad de entrenar redes neuronales de avance profundo"](#) en 2010 .

2006, quedó claro que el entrenamiento previo sin supervisión no es esencial para el aprendizaje profundo.

**Deep Learning =**

**Gran cantidad de datos de entrenamiento + Cómputo paralelo + Algoritmos escalables e inteligentes**

[https://machinelearningknowledge.ai/brief-history-of-deep-learning/#Deep\\_Learning\\_History\\_Timeline](https://machinelearningknowledge.ai/brief-history-of-deep-learning/#Deep_Learning_History_Timeline)