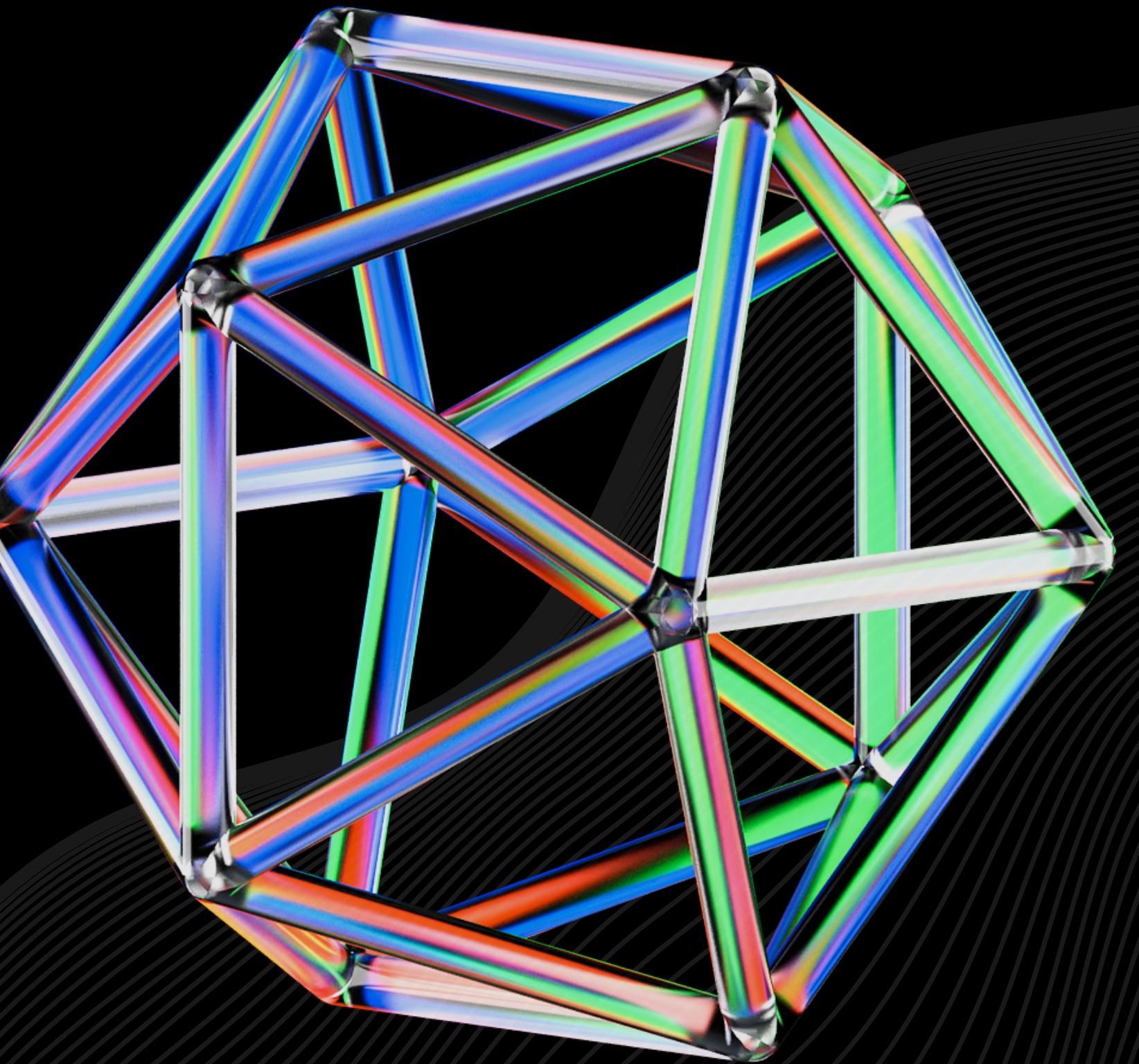


THE UNREASONABLE EFFECTIVENESS OF DEEP LEARNING IN ARTIFICIAL INTELLIGENCE

TERRENCE J. SEJNOWSKI



TERRENCE SEJNOWSKI



PROFESOR DE NEUROBIOLOGÍA EN EL
INSTITUTO SALK.

PROFESOR ADJUNTO EN LA UNIVERSIDAD
DE CALIFORNIA EN SAN DIEGO.

COMPRENDER LOS PRINCIPIOS DEL
CEREBRO Y APLICARLOS EN EL
DESARROLLO DE ALGORITMOS DE
APRENDIZAJE AUTOMÁTICO MÁS
AVANZADOS.

TERRENCE SEJNOWSKI

PREMIO SWARTZ DE INVESTIGACIÓN
TEÓRICA EN NEUROCIENCIA (2001).

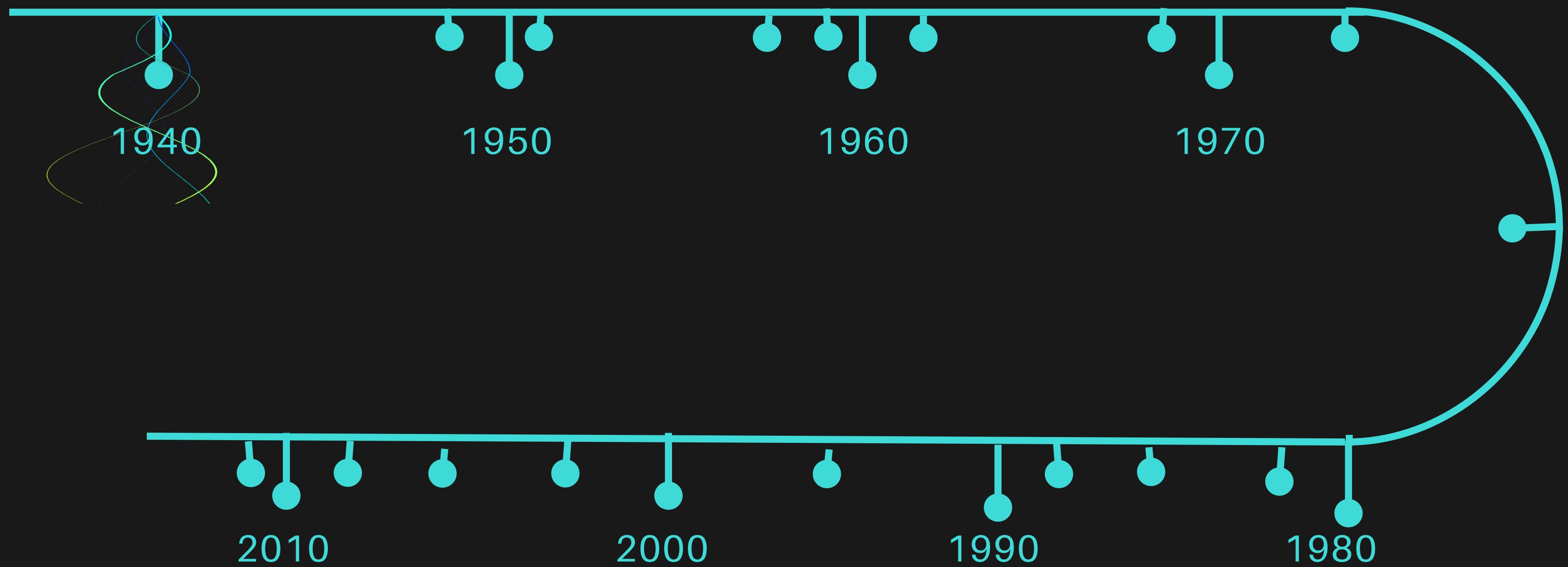
PREMIO HEBB DE LA SOCIEDAD NEURAL
(2005).

PREMIO DE CIENCIA Y TECNOLOGÍA DE LA
FUNDACIÓN PAUL G. ALLEN (2016)

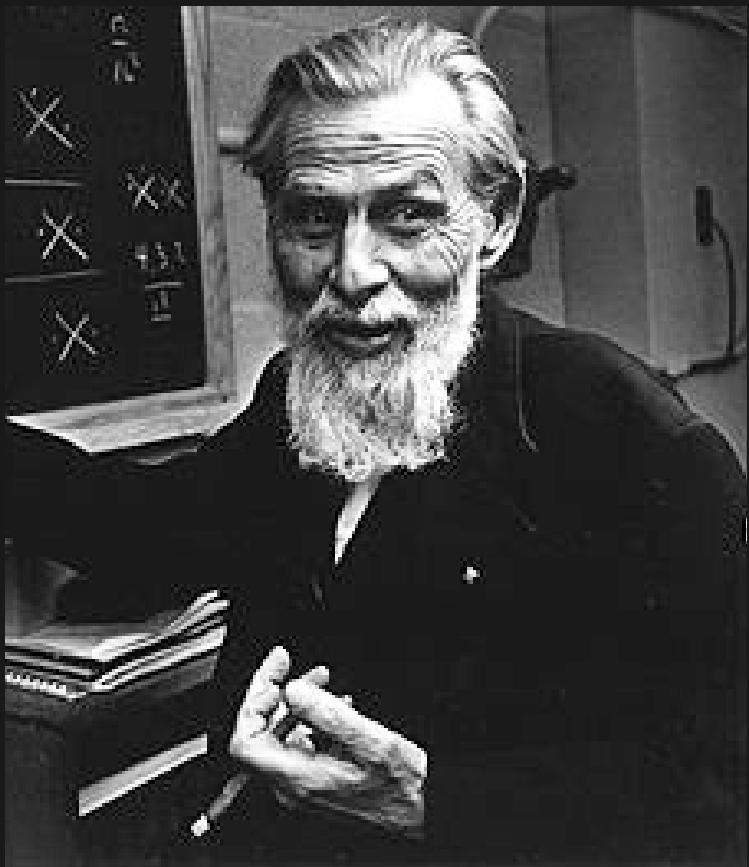
MIEMBRO DE LA ACADEMIA NACIONAL DE
CIENCIAS DE ESTADOS UNIDOS



Evolución del aprendizaje profundo



1940

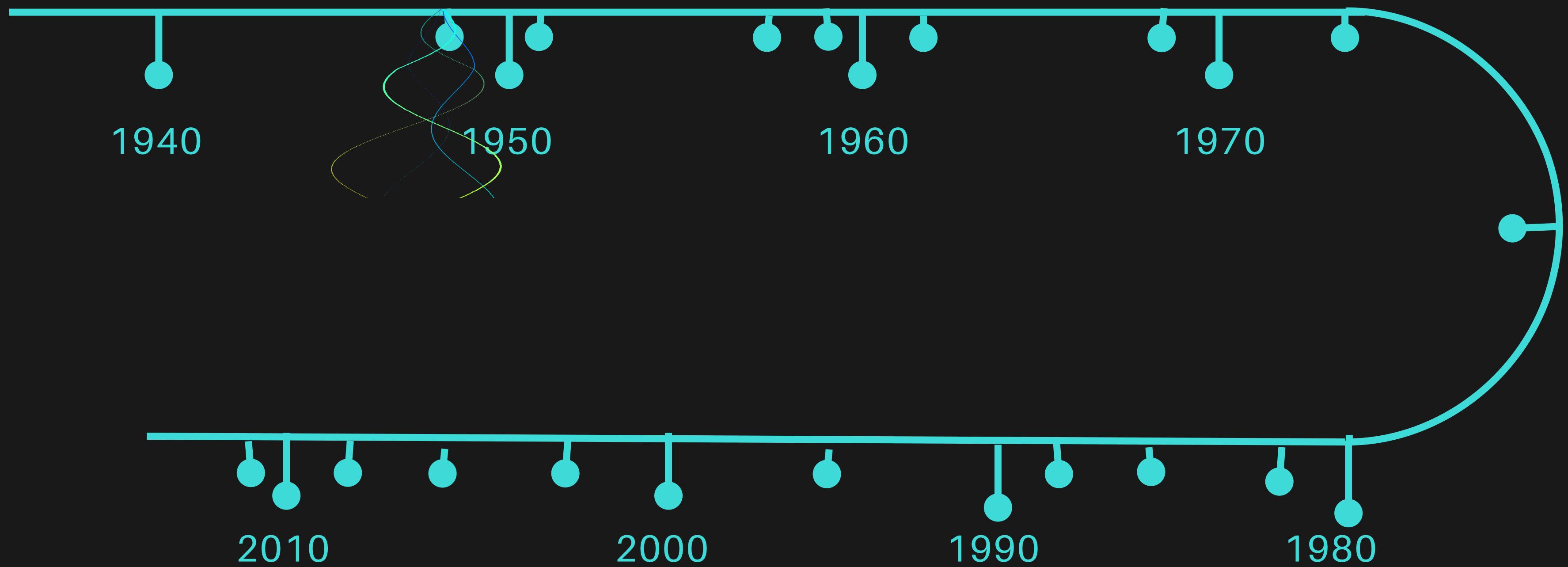


Crearon un modelo computacional para redes neuronales que generó investigaciones no solo en el cerebro sino también en su aplicación a la inteligencia artificial.

El modelo propuesto era constituido por neuronas artificiales, en el que cada una de ellas se caracterizaba por estar activa o desactiva y la activación se daba como respuesta a la estimulación producida por una cantidad suficiente de neuronas vecinas.

**WARREN MCCULLOCH & WALTER
PITTS**

Evolución del aprendizaje profundo



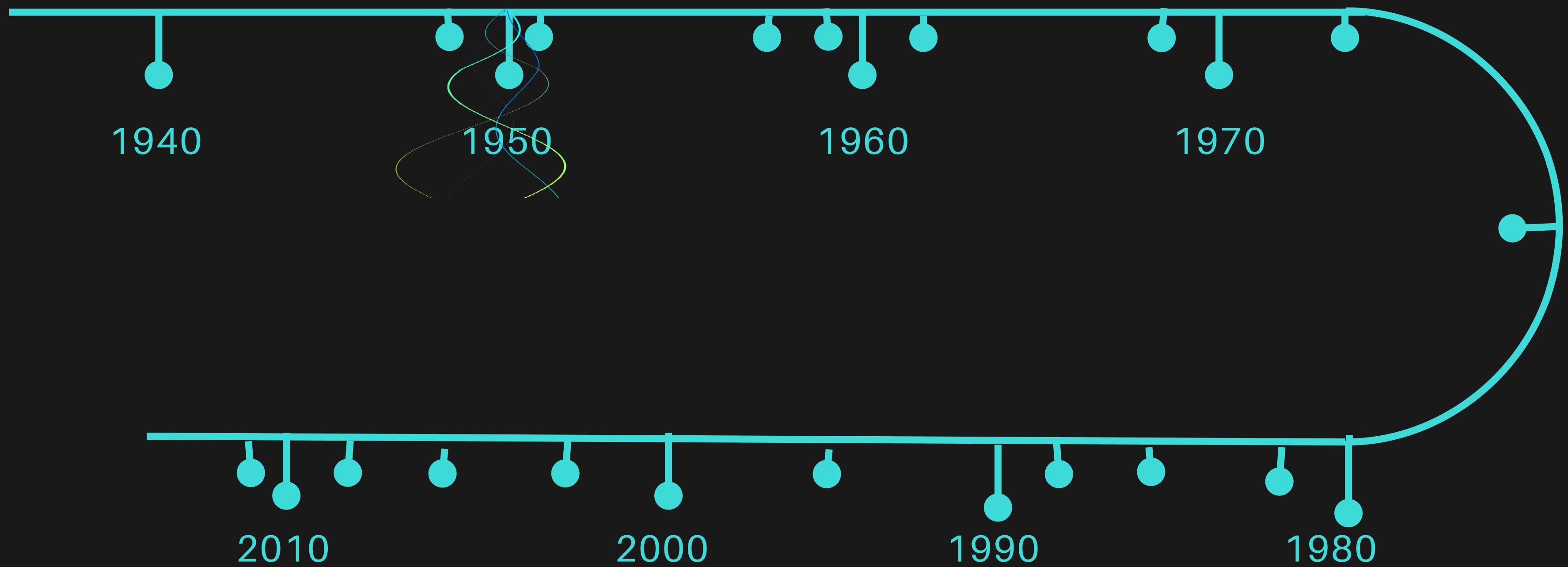
1949



DONALD HEBB

Propuso y demostró una regla de actualización para modificar las intensidades de las conexiones entre neuronas. Creó el aprendizaje de Hebbian , en el que se observa desde la perspectiva de la biología, que la sinapsis entre dos neuronas se fortalece si las dos neuronas están simultáneamente activas.

Evolución del aprendizaje profundo



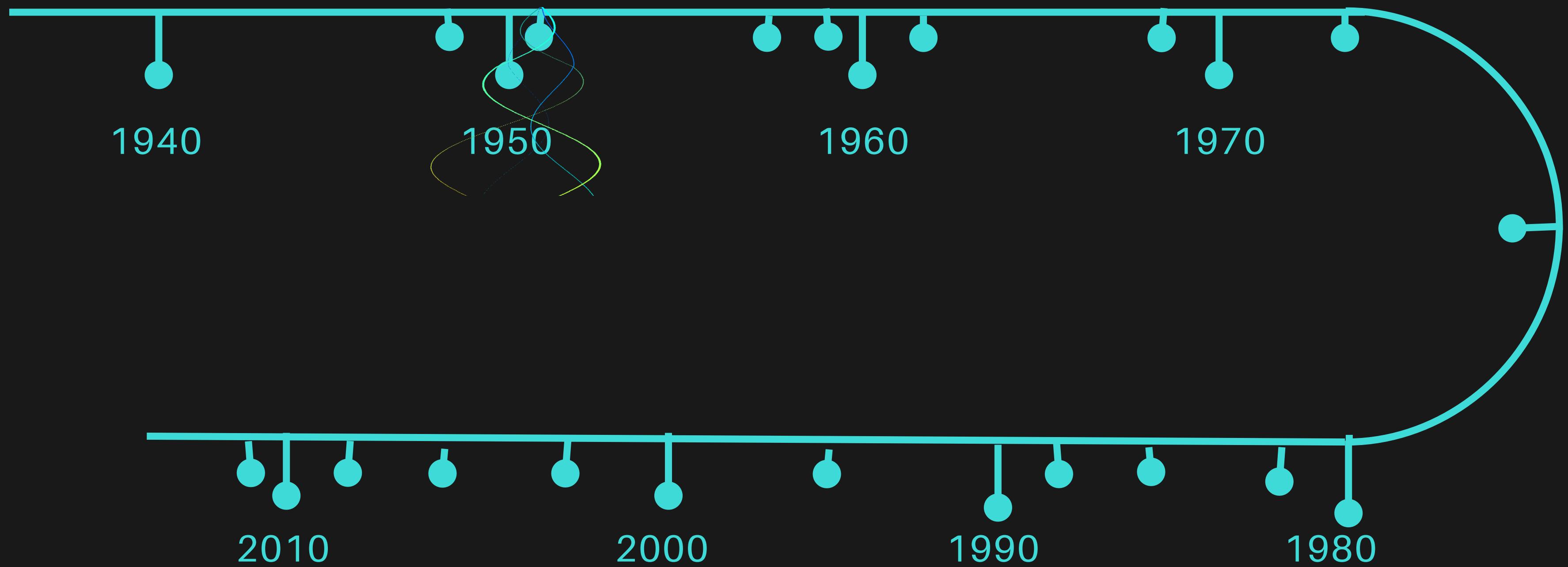
1950



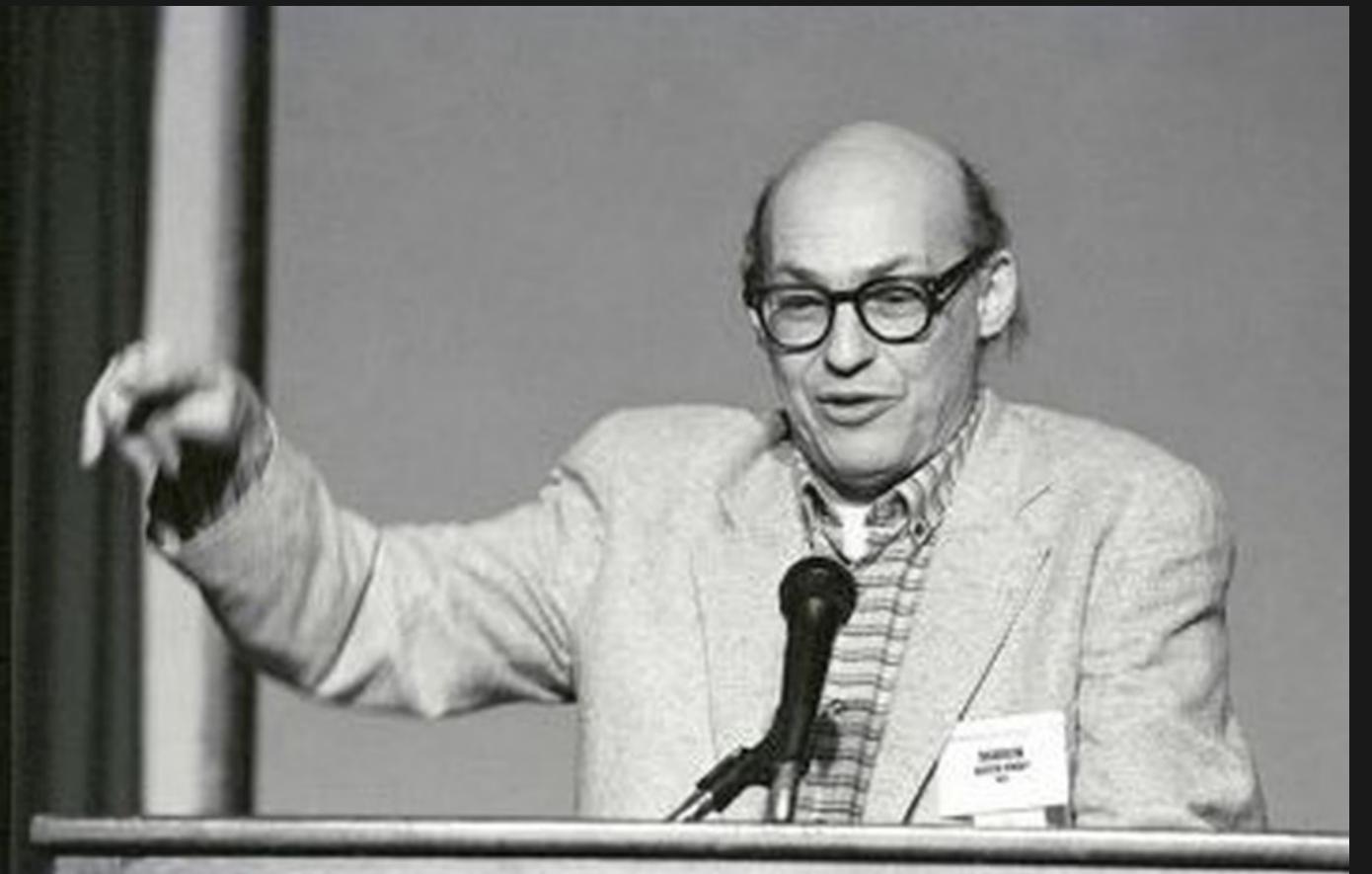
A medida que las computadoras se hicieron más avanzadas en la década de 1950, finalmente fue posible simular una red neuronal. El primer paso hacia esto fue realizado por Nathaniel Rochester de los laboratorios de investigación de IBM. Desafortunadamente para él, el primer intento de hacerlo fracasó.

NATHANIAL ROCHESTER

Evolución del aprendizaje profundo



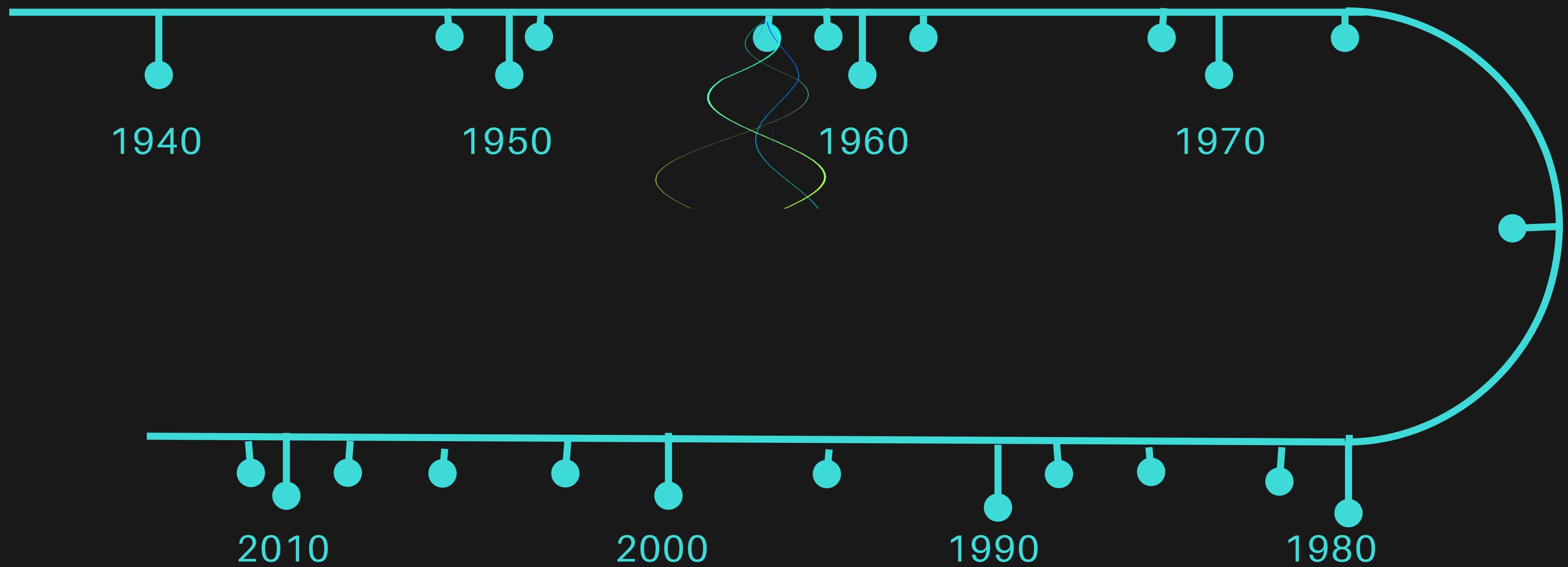
1951



MARVIN MINSKY & DEAN
EDMONDS

Construyeron la primera red neuronal artificial que simulaba una rata encontrando su camino a través de un laberinto. Diseñaron el primer neuroordenador (40 neuronas), SNARC (Ordenador de refuerzo analógico estocástico estocástico). utilizaba 3000 válvulas de vacío y un mecanismo de piloto automático obtenido de los desechos de un avión bombardero B-24 para simular una red con 40 neuronas.

Evolución del aprendizaje profundo



1956



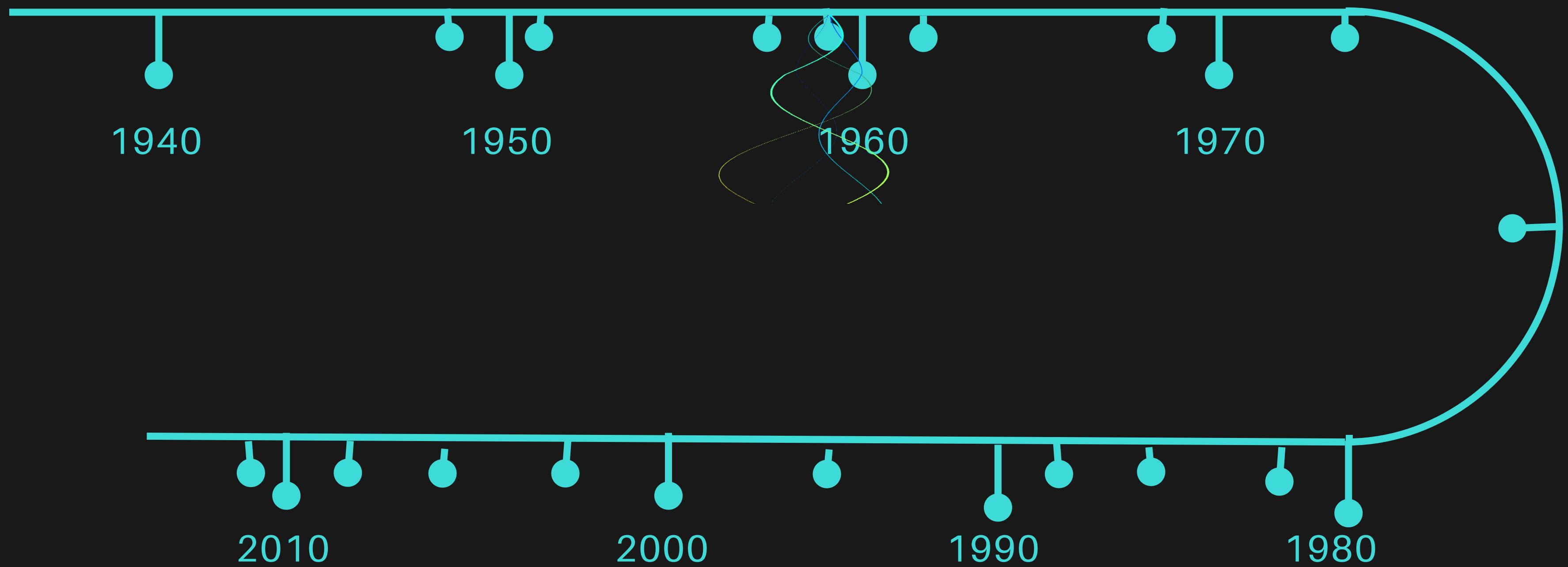
ALLEN NEWELL

Allen Newell, científico informático de la Universidad Carnegie Mellon y uno de los pioneros de la IA que asistió a la conferencia de verano de Dartmouth en 1956, por qué los pioneros de la IA habían ignorado el cerebro, el sustrato de la inteligencia humana.

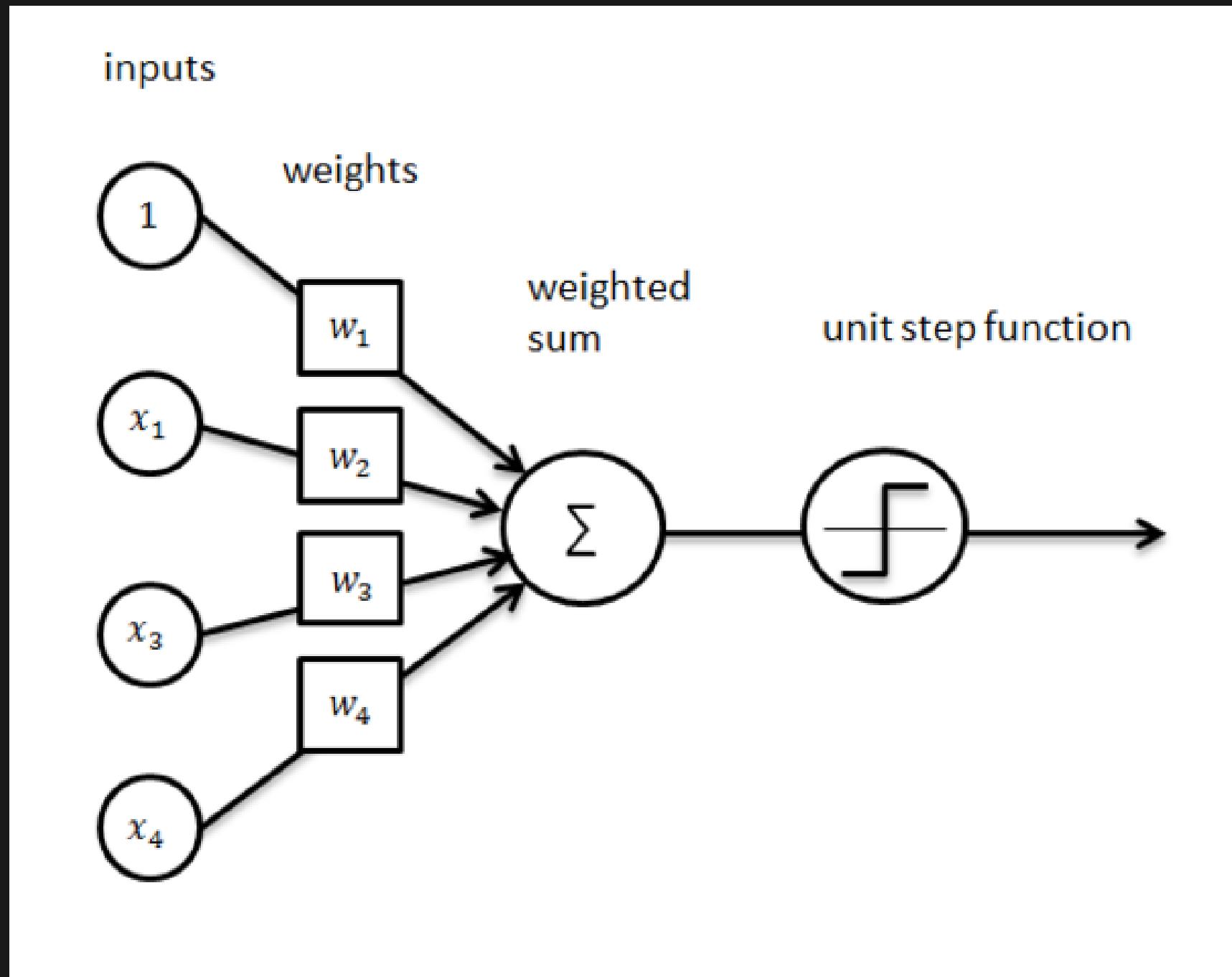
Me dijo que personalmente había estado abierto a las ideas de la investigación del cerebro, pero que simplemente no se sabía lo suficiente sobre los cerebros en ese momento como para ser de mucha ayuda.

No ser relevante ----- No se sabe lo suficiente

Evolución del aprendizaje profundo



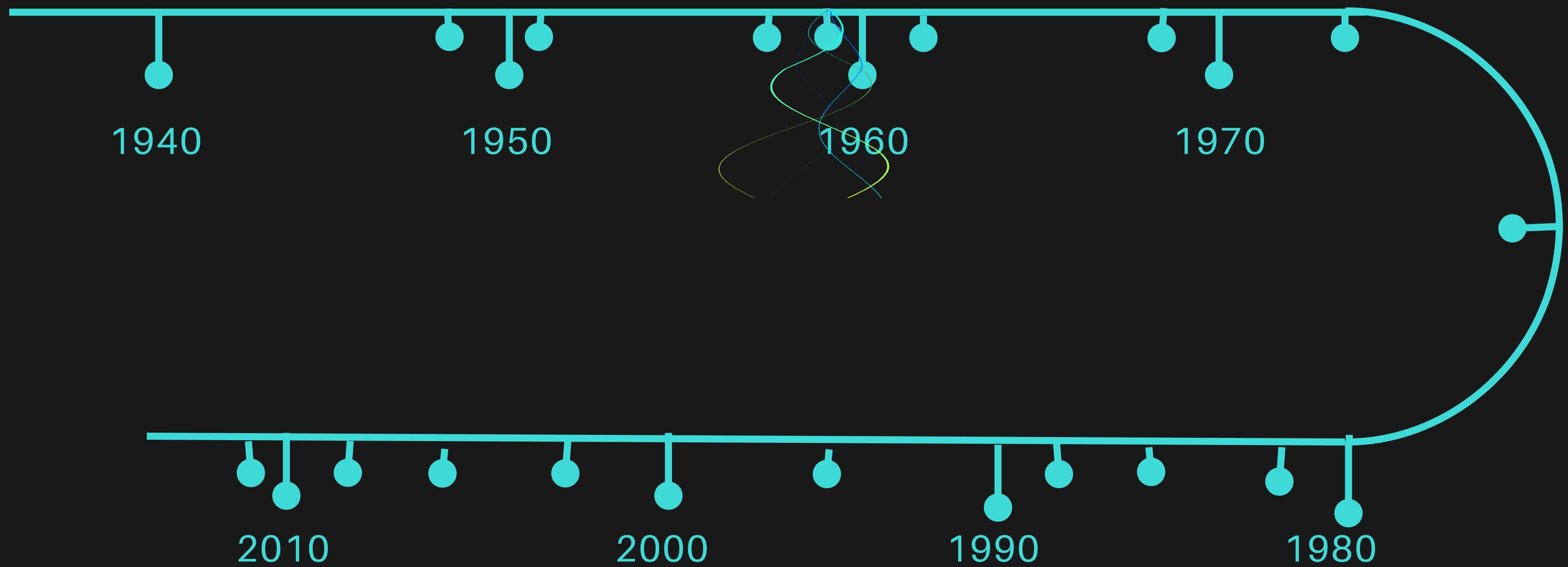
1958



Frank Rosenblatt creó el perceptrón, un modelo neuronal simple que podría usarse para clasificar los datos en dos conjuntos. Sin embargo, este modelo sufrió en que no pudo clasificar correctamente un OR exclusivo.

Los perceptrones se consideran los primeros modelos de redes neuronales artificiales, un perceptrón es un modelo simplificado de una neurona artificial que toma múltiples entradas y produce una única salida

Evolución del aprendizaje profundo



1959



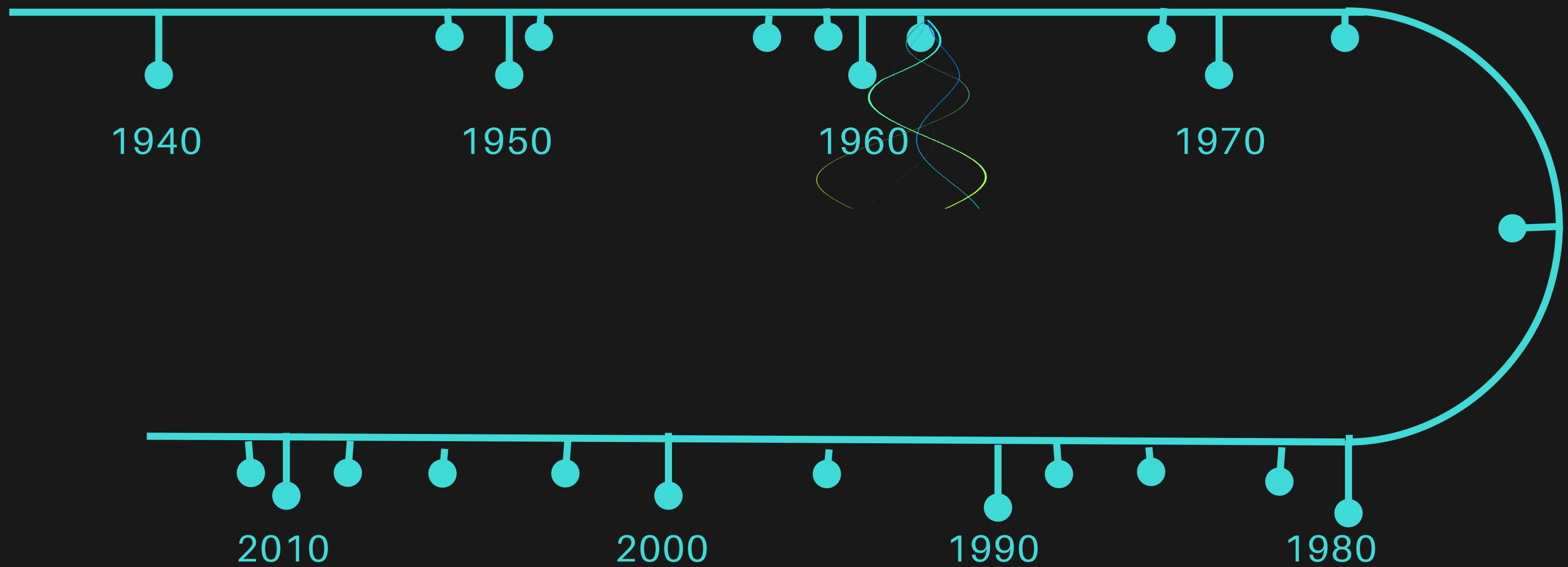
Modelos desarrollados por Widrow y Hoff.

Adeline reconocía patrones binarios para predecir el próximo bit al leer una linea telefónica.

Madeline fue la primera red neuronal aplicada a un problema del mundo real, utilizando un filtro adaptativo que elimina los ecos en las líneas telefónicas

"ADALINE" Y "MADALINE"

Evolución del aprendizaje profundo



1962

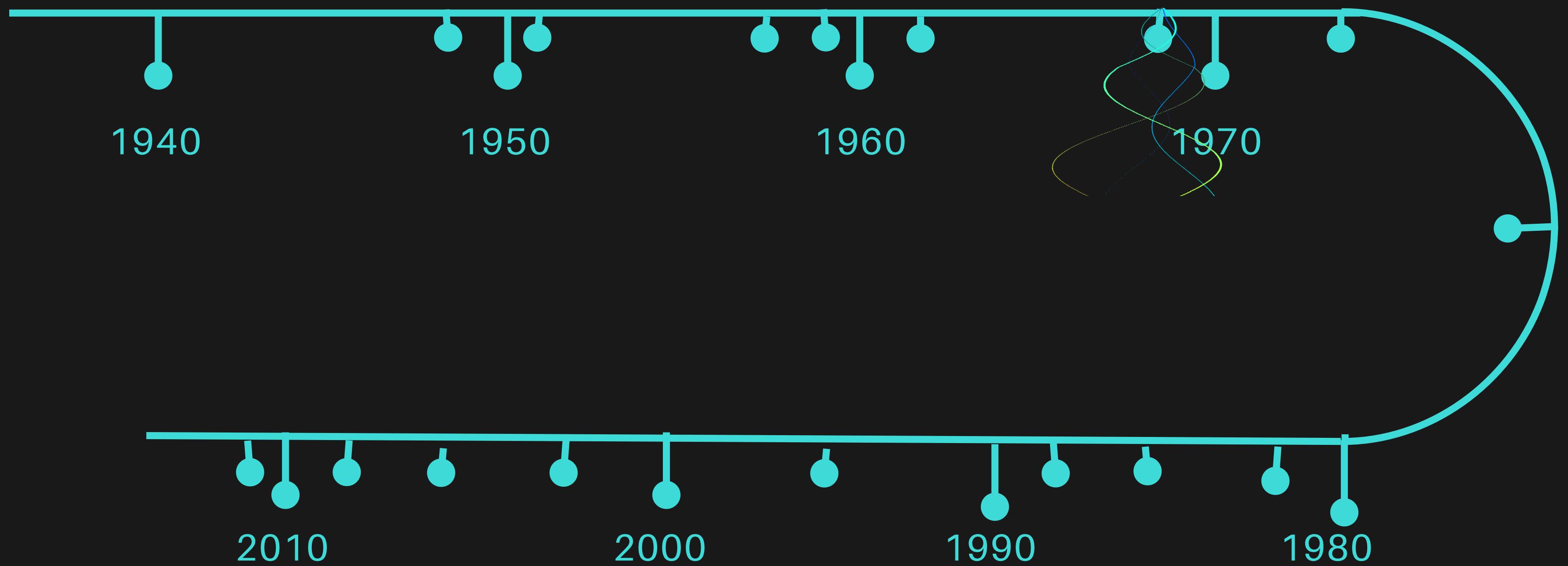


WIDROW & HOFF

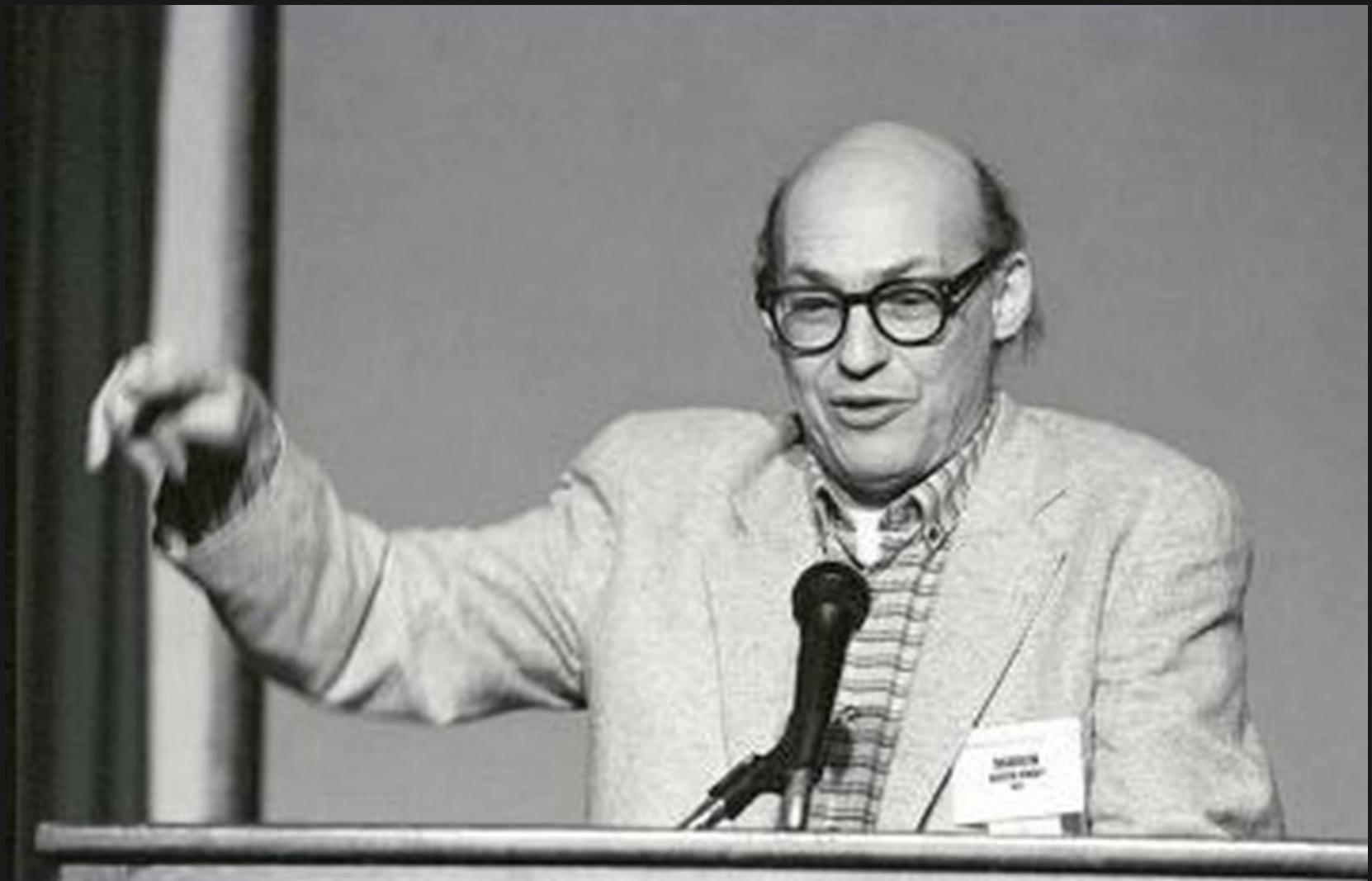
Desarrollaron un procedimiento de aprendizaje que examina el valor antes de que el peso lo ajuste (es decir, 0 o 1) según la regla: Cambio de peso = (Valor de línea anterior al peso) * (Error / (Número de entradas)) .

Son autores del Filtro de Mínimos Cuadrados Medios (LMS)

Evolución del aprendizaje profundo



1969

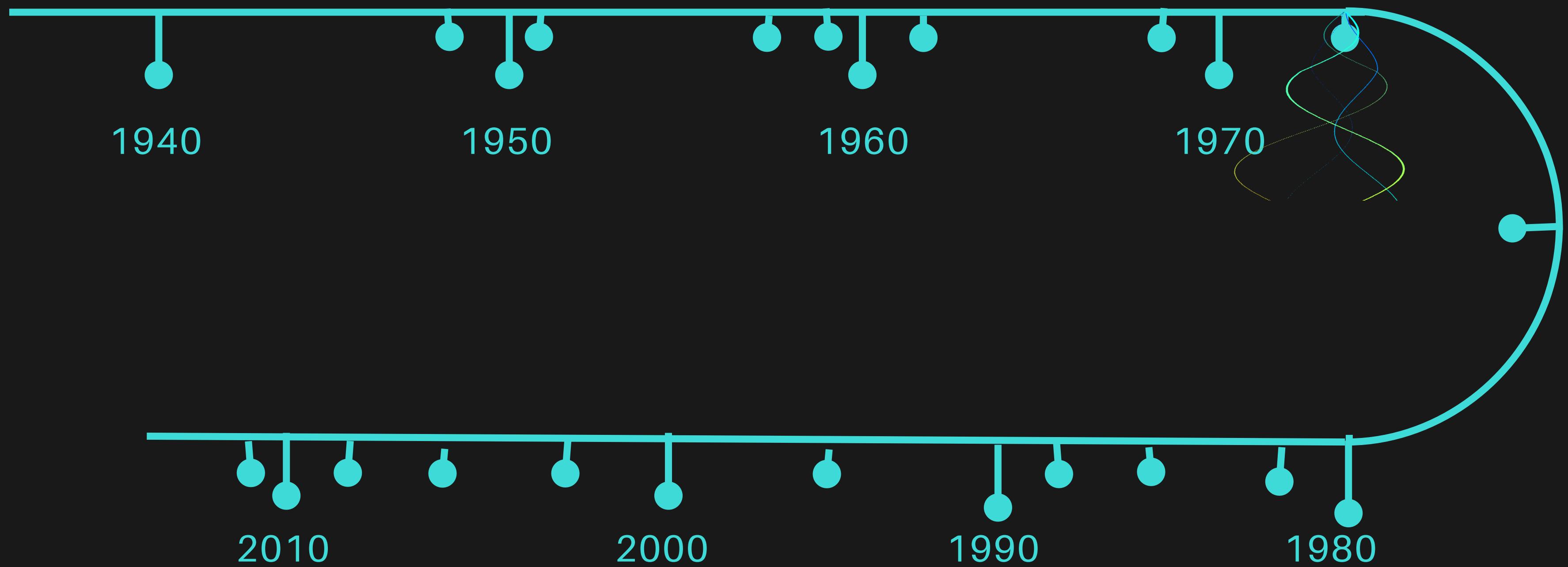


**MARVIN MINSKY Y SEYMOUR
PAPERT**

Su trabajo señaló las limitaciones de los perceptrones y provocó un declive en la investigación y el interés en las redes neuronales en la década de 1970, lo que se conoce como el "invierno de la inteligencia artificial".

Planteó preguntas y desafíos interesantes que inspiraron a futuros investigadores a buscar soluciones y desarrollar nuevas técnicas y modelos de redes neuronales, como las redes neuronales multicapa y la backpropagation

Evolución del aprendizaje profundo



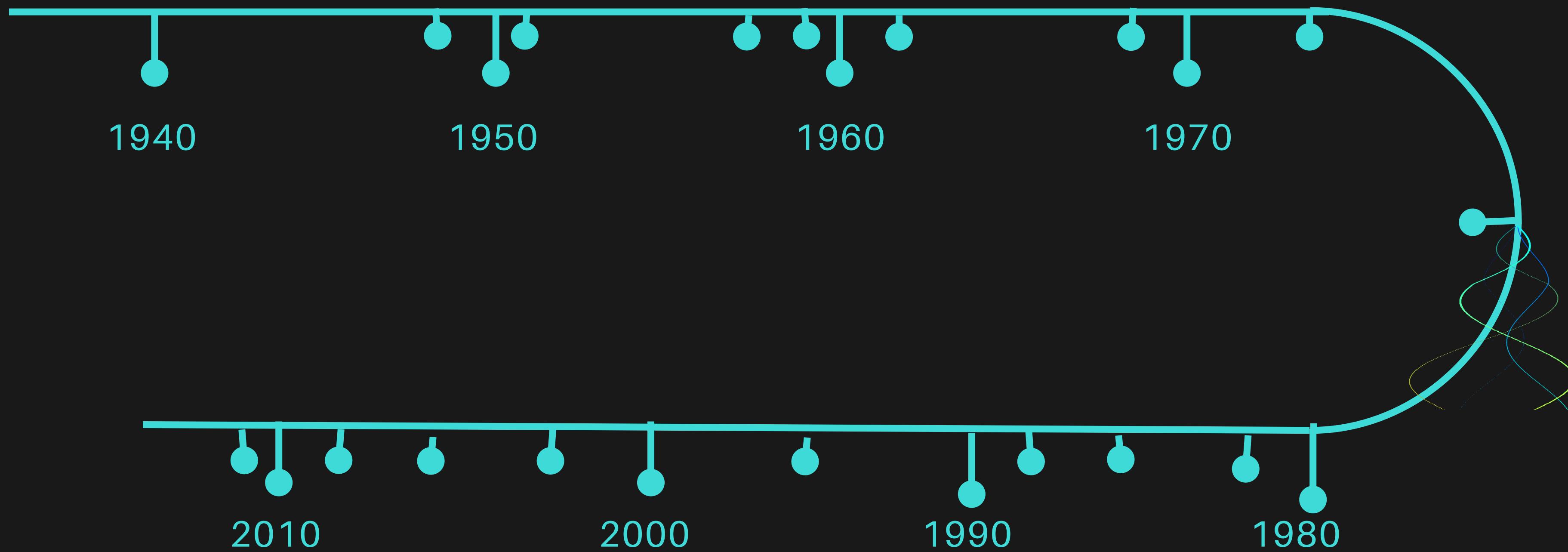
1972



De manera independiente desarrollaron una red similar entre sí usando la misma idea matemática para describir sus ideas. Sin embargo, de manera inconsciente crearon una serie de circuitos analógicos ADALINE. Se supone que las neuronas activan un conjunto de salidas en lugar de solo una.

KOHONEN Y ANDERSON

Evolución del aprendizaje profundo



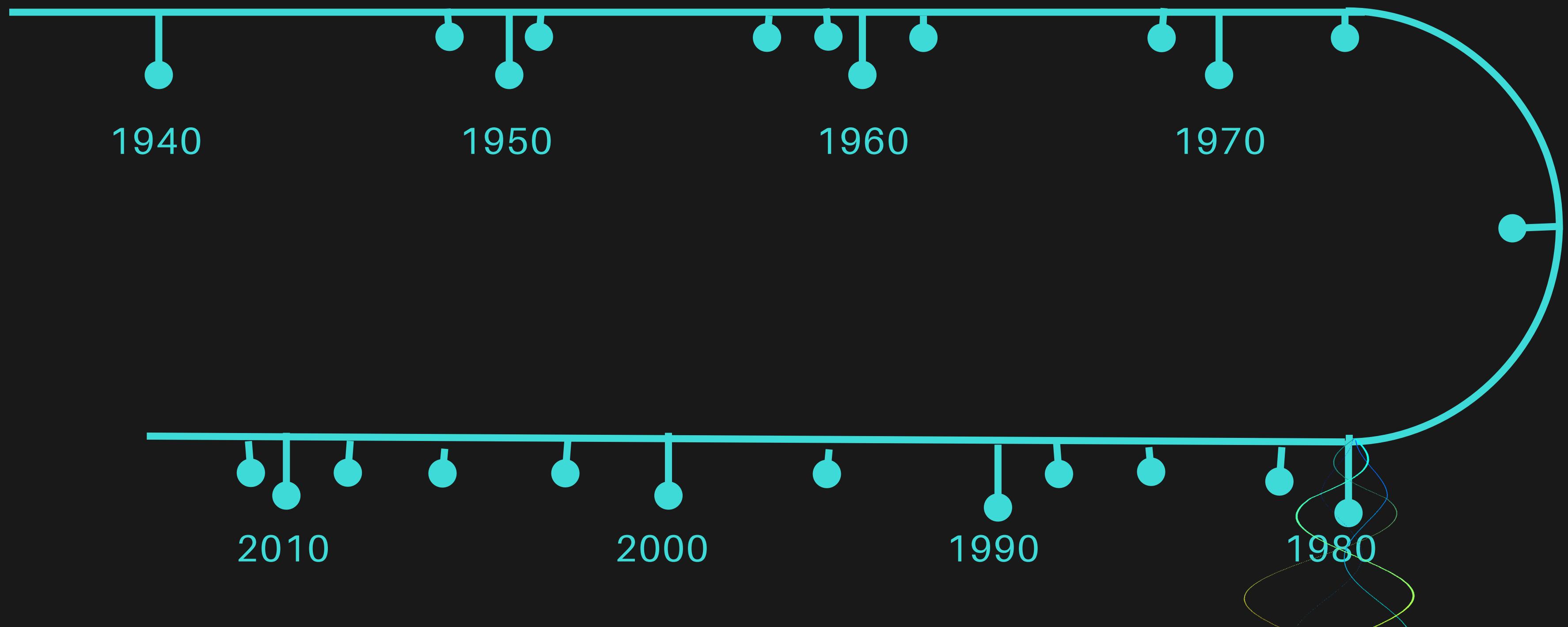
1974



Introdujo el algoritmo de retropropagación del error, el cual permite calcular de manera eficiente los gradientes de error y ajustar los pesos y las conexiones de las neuronas para minimizar el error de salida de la red. Con esto, estableció las bases teóricas de la retropropagación y demostró su utilidad para entrenar redes neuronales multicapa.

PAUL WERBOS

Evolución del aprendizaje profundo



1980



Desarrolló del modelo de red neuronal llamado "Neocognitron". Esta arquitectura se diseñó para el reconocimiento de patrones en imágenes, también incluyó mecanismos de aprendizaje que permitían a la red ajustar sus pesos y conexiones para mejorar su capacidad de reconocimiento de patrones.

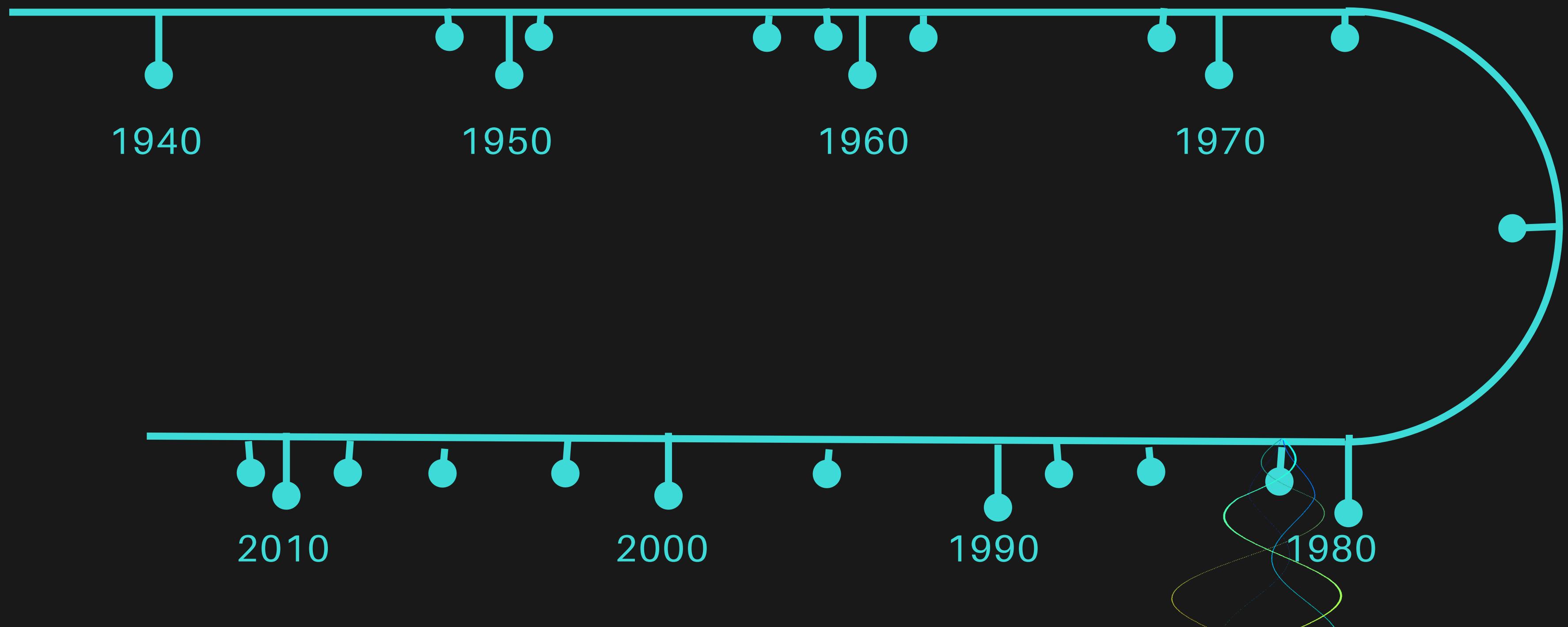
KUNIHIKO FUKUSHIMA

1980

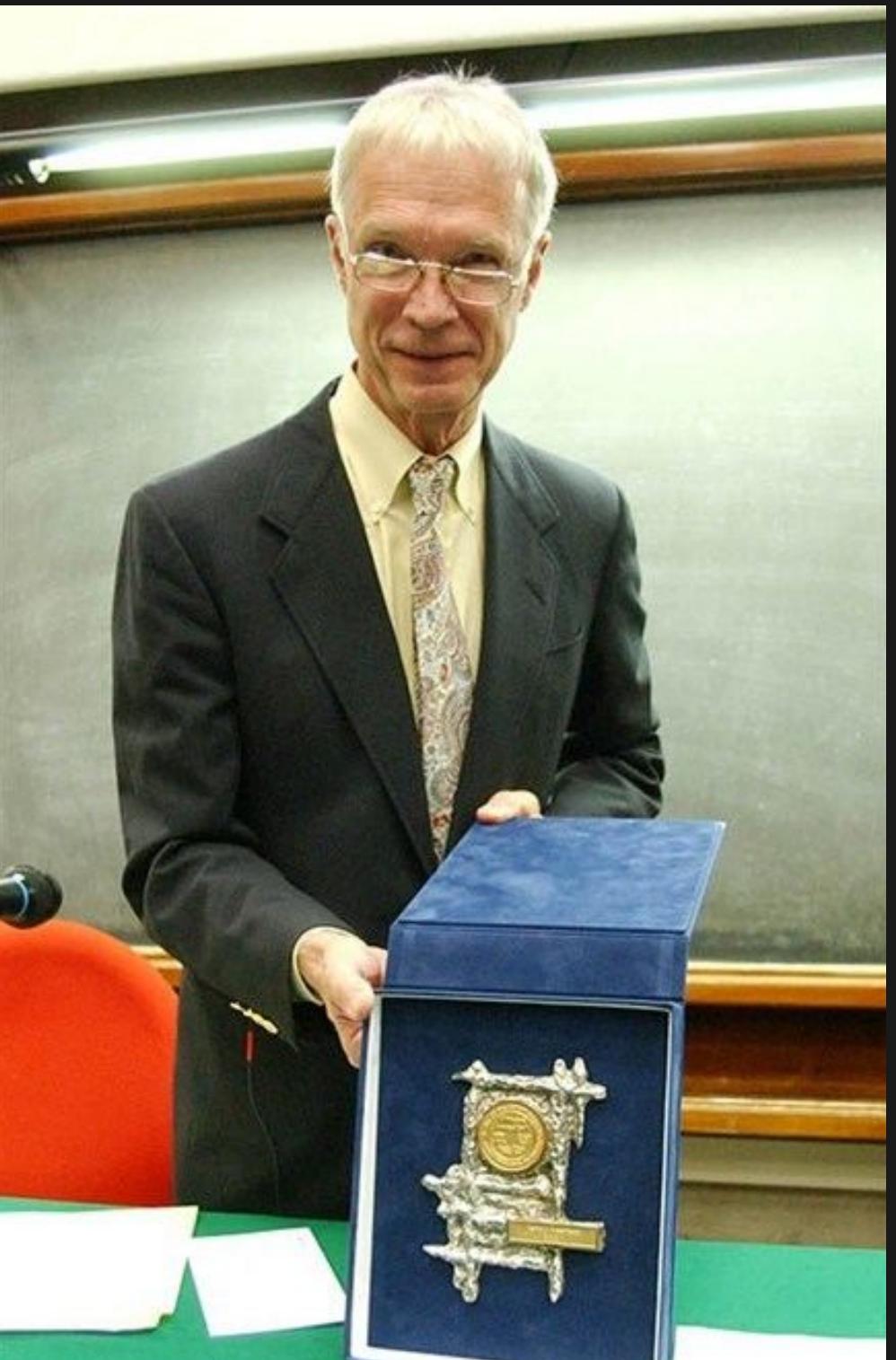
En 1980 rara vez tenían más de una capa de unidades ocultas entre las entradas y las salidas, pero ya estaban sobreparametrizados según los estándares del aprendizaje estadístico.

De acuerdo con la teoría estadística, la generalización no debería ser posible con los conjuntos de entrenamiento pequeños que estaban disponibles. Sin embargo, con métodos de regularización simple se condujeron a modelos con una generalización sorprendentemente buena.

Evolución del aprendizaje profundo



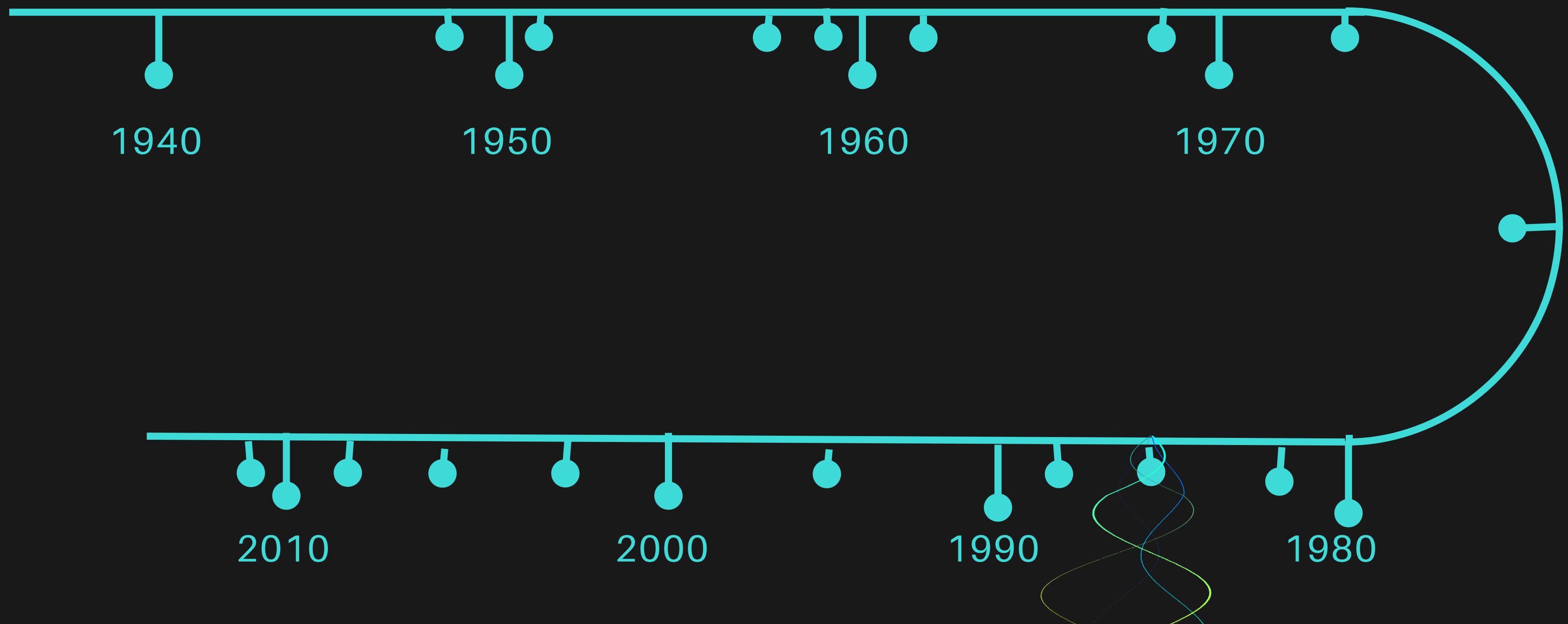
1982



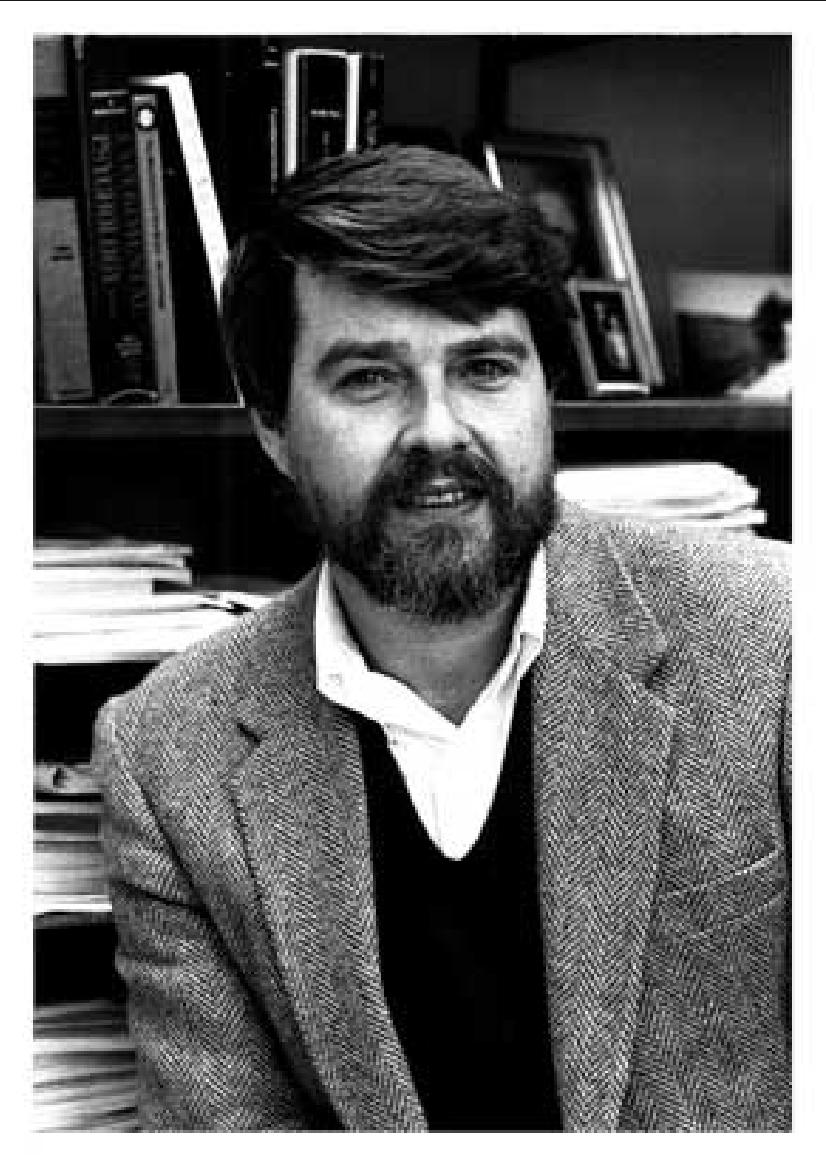
La contribución clave de Hopfield fue su formulación de una función de energía, conocida como función de energía de Hopfield, que permitía a la red alcanzar estados estables y resolver problemas de optimización.

Ayudó a revitalizar el interés en las redes neuronales y abrió nuevas direcciones de investigación en el campo. La Red de Hopfield se ha aplicado con éxito en diversas áreas, como reconocimiento de patrones, optimización, modelado de memoria y procesamiento de información.

Evolución del aprendizaje profundo



1986

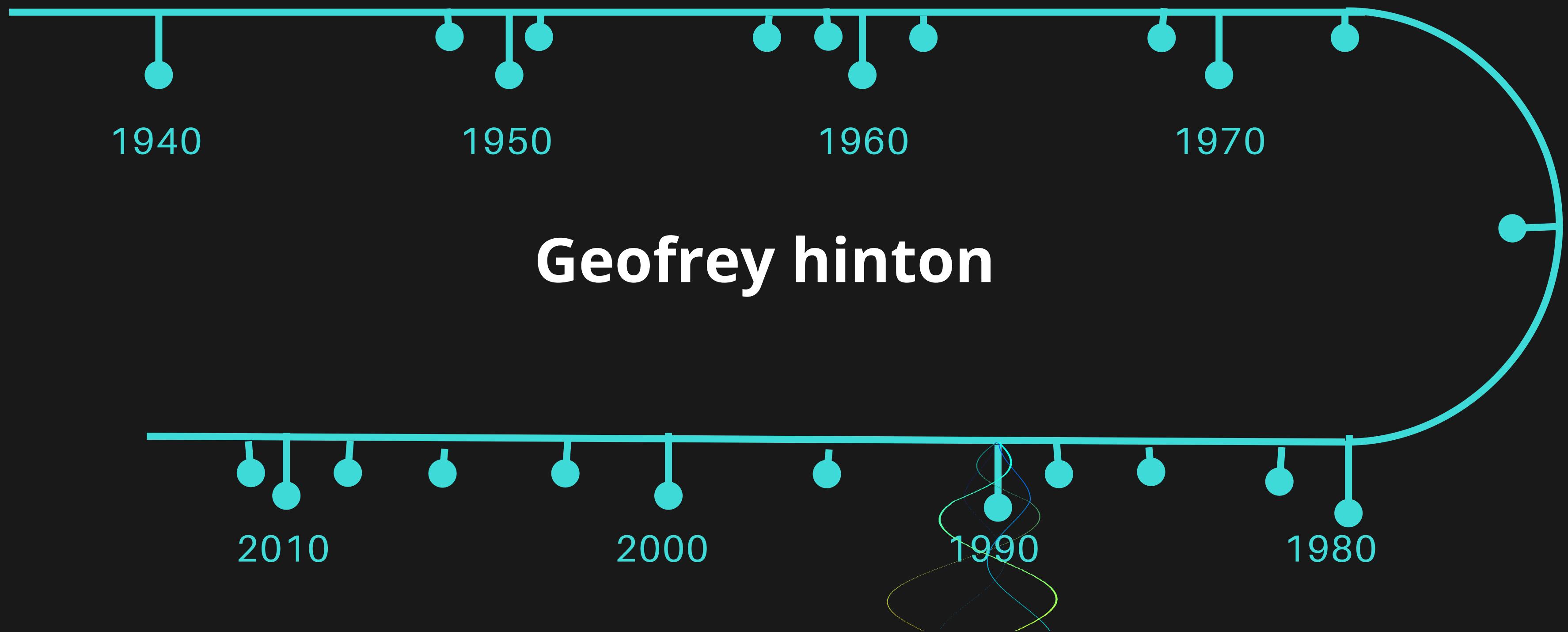


RUMELHART & MCCLELLAND

El algoritmo de aprendizaje de retroalimentación se aplicó a diversos problemas de aprendizaje en los campos de la informática y la psicología, y la gran difusión que conocieron los resultados obtenidos, publicados en la colección "Parallel Distributed Processing".

Adicionalmente, se utilizó una "red híbrida" con múltiples capas, cada capa utilizando una estrategia diferente de resolución de problemas

Evolución del aprendizaje profundo





GEOFREY HINTON

Es uno de los pioneros y líderes en el campo de las redes neuronales y el aprendizaje profundo. Sus contribuciones han sido fundamentales para el avance y la popularización de estas tecnologías.

- Entrenamiento de redes neuronales profundas con la técnica "pre-entrenamiento no supervisado".
- Redes neuronales convolucionales
- Introducción de las Restricted Boltzmann Machines.
- Aprendizaje por refuerzo profundo

**TO BE
CONTINUED...**

PREGUNTAS PARA MEJORAR NUESTRA ARQUITECTURA Y ALGORITMOS DE APRENDIZAJE

- ¿Por qué es posible generalizar a partir de tan pocos ejemplos y tantos parámetros?
- ¿Por qué el descenso de gradiente estocástico es tan efectivo para encontrar funciones útiles en comparación con otros métodos de optimización?
- ¿Qué tan grande es el conjunto de todas las buenas soluciones a un problema?

PREGUNTAS PARA MEJORAR NUESTRA ARQUITECTURA Y ALGORITMOS DE APRENDIZAJE

- ¿Están las buenas soluciones relacionadas entre sí de alguna manera?
- ¿Cuáles son las relaciones entre las características arquitectónicas y el sesgo inductivo que pueden mejorar la generalización?



Tercera generación - IA

- ¿Cómo relacionar el DL con la inteligencia humana?
- ¿Es o no relevante el estudio del cerebro?

Se debe considerar como complemento no como mundos diferentes



Redes Neuronales



Analisis de Vuelo de las
Aves

WHAT'S NEXT?

Futuro de deep learning

Hacia una IA generalizada

