

## RESUMEN

### ¿Qué es una red neuronal?

Se trata de una familia de algoritmos muy potentes con los que se puede modelar comportamientos inteligentes, Las neuronas están formadas por conexiones de entrada que reciben señales externas, que luego se procesan mediante una operación numérica internamente para generar una señal de salida, el procesamiento es un modelo matemático el cual es esencialmente una suma ponderada de los valores de entrada, donde los pesos se asignan a cada conexión y determinan la intensidad con la que cada variable de entrada afecta a la neurona. Entonces dichos parámetros se pueden ajustar para entrenar a la red para que realice predicciones o clasificaciones más precisas. También se observa que dicha expresión matemática es parecida a un regresor lineal.

Las redes neuronales tienen una trascendencia muy alta y se está empleando en muchas áreas como lo son: Reconocimiento de (caracteres, imagen y voz), predicción bursátil, generación de texto, traducción de idiomas, conducción autónoma, clasificación de cosas, etc.

Se representa un ejemplo donde la finalidad es encontrar la combinación perfecta de parámetros para el modelo nocturno ideal, donde ambas variables (realidad virtual y nachos) deben estar activadas para que el resultado sea 1, al final se encuentra un modelo donde se le asigna unos valores numéricos que después se representan en forma binaria y puede verse como una puerta AND lógica o una línea de regresión que separa los puntos en dos grupos, pero no debemos limitarnos a usar una sola neurona ya que como se ve al final del video no basta con una sola ya que se necesita realizar dos divisiones simular a una compuerta XOR.

### Parte 2:

Las neuronas se pueden representar mediante capas las cuales pueden ser tres (capa de entrada, capas ocultas y capa de salida), si colocamos dos neuronas de forma secuencial la red puede aprender conocimiento jerarquizado, se muestra el mismo ejemplo, pero complementando con otra neurona que establece la motivación por un examen, por ello cuando se agregan más capas más complejo puede ser el conocimiento que elaboremos y eso da origen al aprendizaje profundo. Sin embargo, el uso de muchas operaciones de regresión lineal puede hacer que la estructura colapse y se vuelva equivalente a una red de una sola neurona.

Entonces para impedir esto necesitaríamos realizar manipulaciones no lineales, que se logran mediante funciones de activación. La función de activación escalonada no es muy interesante para emplear, pero la sigmoide tiene una forma más suave la cual permite establecer que los valores grandes se saturan en 1 y los valores pequeños en cero y sirve para representar probabilidades; de igual manera se explican las funciones de activación y su importancia en la generación de patrones no lineales para la clasificación de datos.

En el ejemplo se muestra una combinación de cuatro neuronas puede resolver un problema de clasificación al producir una superficie plana con una protuberancia en el medio, lo que da como resultado el límite circular necesario para el problema en cuestión que es separar los

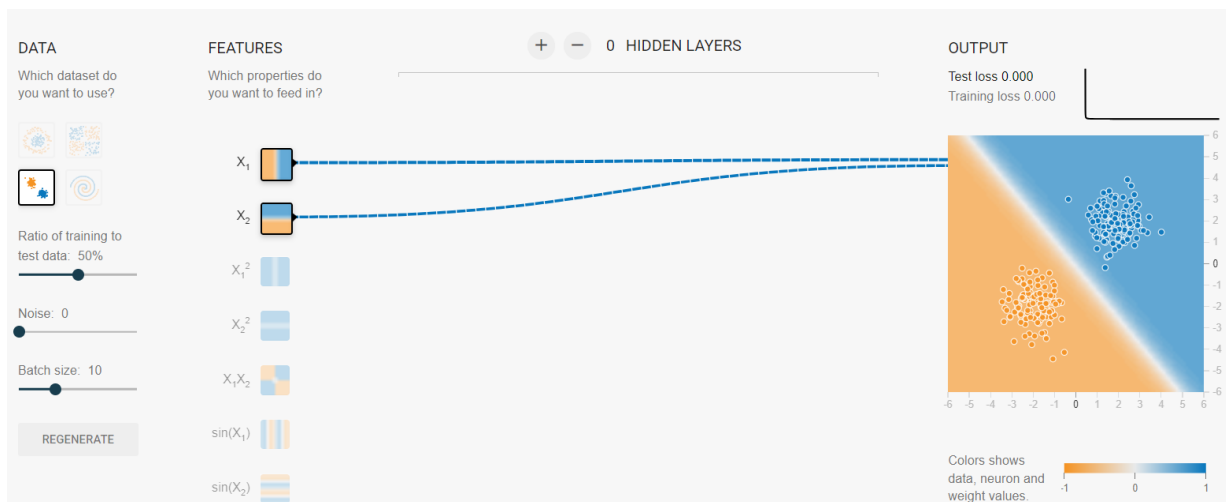
“O” de las “X”. El ejemplo demuestra que las redes neuronales son capaces de desarrollar soluciones complejas combinando muchas neuronas.

### Jugando con una Red Neuronal:

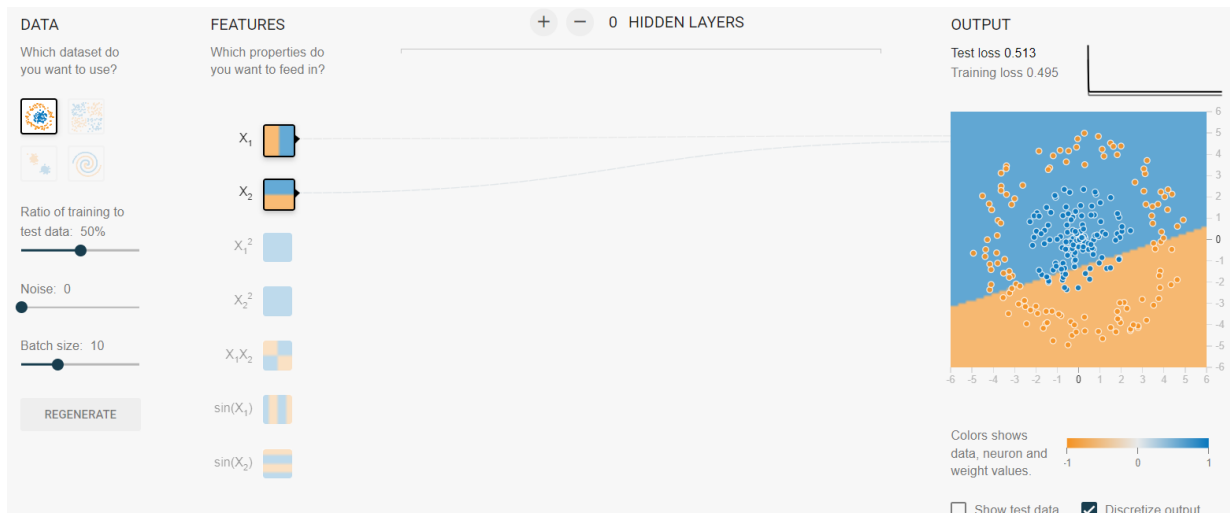
En la primera parte se presenta un flujo tensor de juegos, es una herramienta en línea donde los usuarios pueden jugar con una red neuronal y se puede ir ajustando diferentes parámetros mientras se visualiza la salida, esta herramienta emplea cuatro ejemplos de puntos de datos para demostrar problemas de clasificación a través de diferentes arquitecturas de redes neuronales, en la herramienta web se tiene la libertad de modificar variables para poder probar cómo se comporta la red neuronal.

Se analiza las variables de entrada, salida de las redes neuronales y las limitaciones que se presentan si se desea visualizar datos de alta dimensión ya que, en el ejemplo, se analiza en dos dimensiones es decir tenemos dos variables de entrada, pero se puede trasladar a dimensiones más altas. Si bien los ejemplos mostrados presentan dos variables de entrada, en realidad, las redes neuronales pueden funcionar con un número mucho mayor de variables, a veces millones.

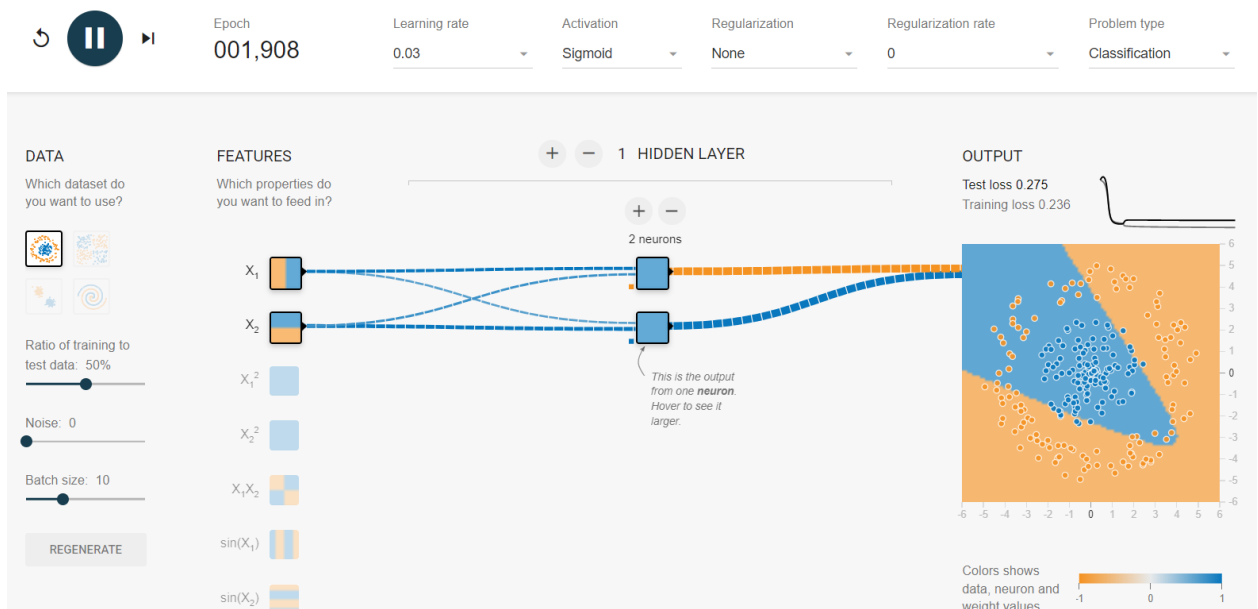
Se realiza una práctica con la nube de puntos gaussiana ya que puede ser separada linealmente lo cual no tiene gran complejidad para entender el comportamiento de agregar capas y más neuronas.



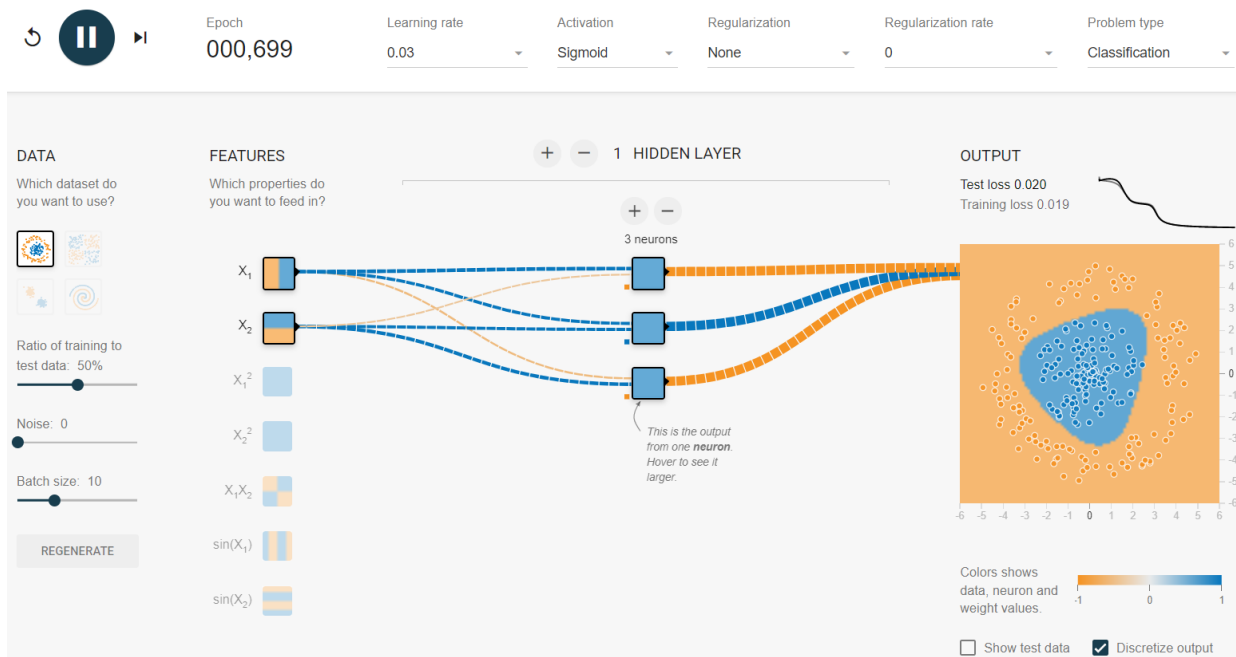
Como se observa en este primer punto se realiza bien la clasificación sin agregar ninguna capa intermedia o capas ocultas como se les llama comúnmente.



Para este caso se observa que para la otra nube de puntos no se realiza bien la clasificación ya que es necesario colocar capas que permitan procesar los datos y tener una salida más acertada.



En este caso ya usamos dos neuronas con una función de activación sigmoide, pero se observa que no se ajusta a la nube de puntos y que esto se soluciona agregando otra neurona más como se observa en la siguiente imagen:



De esa manera se puede ir agregando cantidad de neuronas y capas convenientes y tiene una forma suavizada gracias a la función de activación.

### Descenso del gradiente:

El gradiente se presenta como una parte clave del campo del aprendizaje automático y la mayoría de los sistemas de IA que existen hoy en día utilizan este algoritmo; dicho concepto del algoritmo de descenso de gradiente consiste en modificar los parámetros de un modelo, y si alteramos estos parámetros, podemos variar el error de nuestro modelo. Entonces, la función de costo nos dice el error para cada combinación de parámetros; el mínimo de una función convexa siempre será un mínimo global, pero con funciones no convexas, podemos tener múltiples mínimos locales y puntos silla. en conclusión, el descenso de gradiente es utilizado para movernos en la dirección donde el error es menor, donde se evalúa la pendiente y recorriendo el gradiente hasta alcanzar el mínimo.

### ¿Qué es una Red Neuronal? Parte 3

En 1986 permitió que las redes neuronales aprendieran de forma autónoma. Antes de esto, la neurona única utilizada en el modelo de perceptrón sólo podía resolver problemas lineales, lo que limitaba el potencial de las redes neuronales.

La retropropagación permitió a las redes neuronales ajustar sus parámetros y aprender de los datos que estaban procesando. La retropropagación se utiliza para calcular los gradientes de redes neuronales complejas esto implica rastrear la responsabilidad de la contribución de cada neurona a la salida final hasta la primera capa de la red, pero cuando no tenemos capas es más sencillo ya que se asemeja a un regresor lineal. Entonces se debe calcular cuánto debe ajustarse cada parámetro de cada neurona para optimizar la función de costos de la red.

Se realiza un ejemplo con producción y distribución de café para ilustrar el concepto de cadena de responsabilidades, que puede considerarse como una red neuronal donde cada nodo (o neurona) es responsable de una tarea específica.

## ¿Qué es una Red Neuronal? Parte 4

Se explicó el algoritmo de retropropagación para entrenar redes neuronales. Este algoritmo se utiliza para calcular las derivadas parciales de los parámetros de la red con respecto a la función de costo, esencial para optimizar la red con el descenso de gradiente. Se usó una analogía empresarial para ilustrar cómo se propaga el error en la red. Aunque las matemáticas pueden parecer complicadas, el presentador aseguró que son comprensibles con estudio y práctica.

La retropropagación en redes neuronales implica calcular cómo cambian los pesos de la red para minimizar el error. Esto se hace calculando tres partes clave:

- La diferencia entre la salida deseada y la salida real de la red (error).
- Cómo cambian las sumas ponderadas de las entradas a las neuronas.
- Cómo cambian los pesos mismos.

Estas partes se multiplican juntas para encontrar la dirección en la que deben ajustarse los pesos para reducir el error. Luego, se utilizan para actualizar los pesos y mejorar el rendimiento de la red durante el entrenamiento. El concepto matemático de retropropagación en redes neuronales implica calcular cómo se propagan los errores a través de la red para ajustar sus parámetros y mejorar el rendimiento. Para hacerlo, se calcula la derivada de la función de activación en relación con la salida de la neurona, así como las derivadas de la suma ponderada y los parámetros de polarización.

El concepto de retropropagación se basa en cuatro expresiones clave que permiten recorrer una red neuronal de atrás hacia adelante, calculando derivadas parciales. Dos de estas expresiones se utilizan para calcular el error en la capa actual en relación con la capa anterior, mientras que las otras dos se emplean para calcular las derivadas parciales de los parámetros de sesgo y ponderaciones.