**Machine Learning II**

Actividad 1

# Datos de Referencia

**Tema:** Redes Neuronales

**Autor:** Michel Stivens Larrota Villalba

**Código:** 40098212\_IMPN0007M801

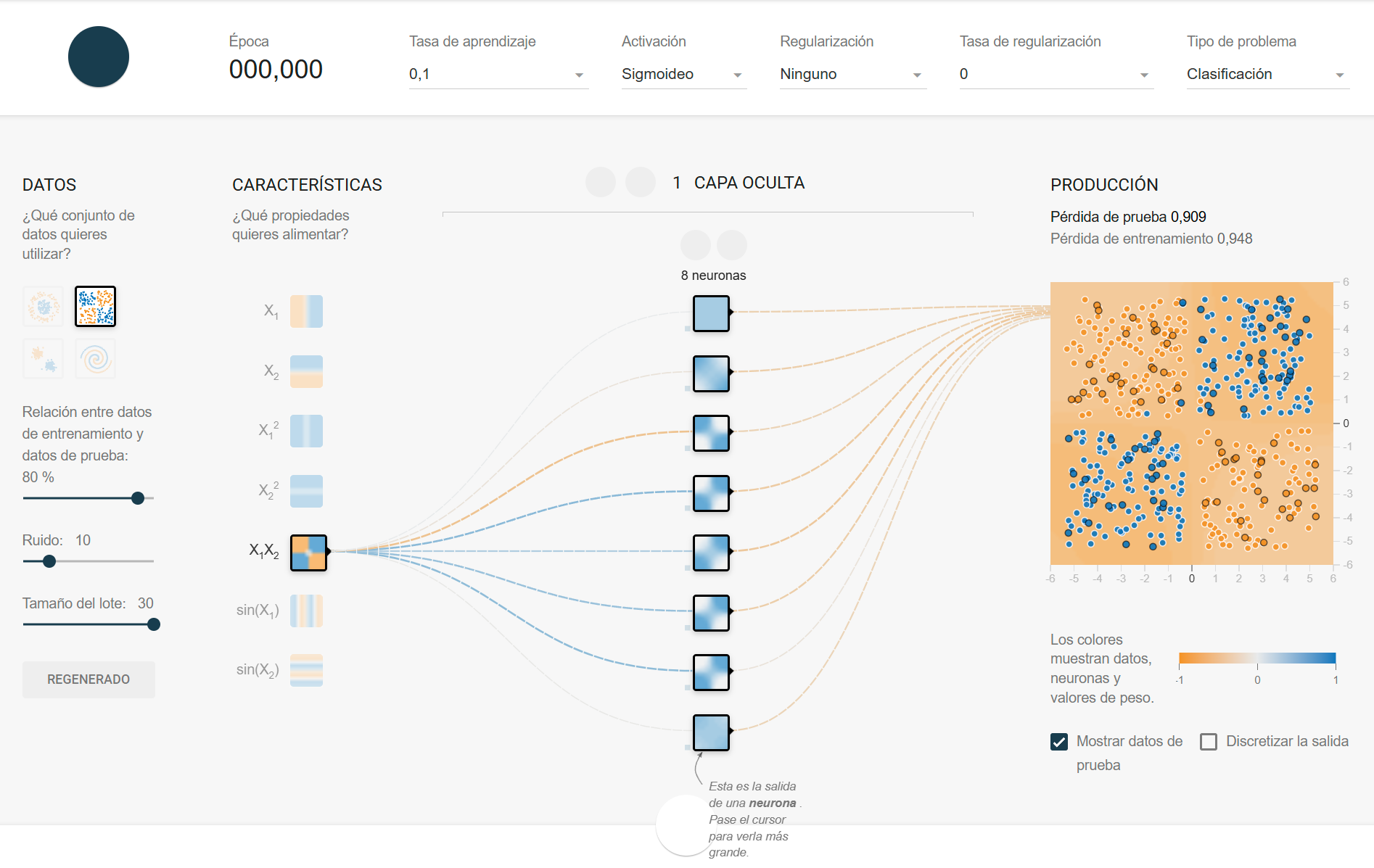
**Fecha:** 20 de noviembre de 2024

# Experimentación

## Generalidad

El desarrollo de la actividad se realiza en 5 fases, las cuales se describen a continuación:

Ajuste general de la herramienta:



Se realizan los ajustes indicados, se establece una relación 80% para datos de entrenamiento y 20% para datos de prueba, se incluye un valor de Ruido de 10, para incluir complejidad al modelo y obtener un resultado más robusto, orientando la actividad a un escenario más real de tratamiento de los datos, se utilizó un tamaño de lote de 30, con el objeto de contar con el mayor número de datos posible en cada interacción de aprendizaje del modelo, con el objeto de contar con una mayor estabilidad y reducir el tiempo de aprendizaje del modelo, en consecuencia al ruido incluido en el modelo; se utiliza 1 capa como fue indicado, pero se utilizan la totalidad de las neuronas disponibles para la capa (8 neuronas), como conjunto de datos para el ejercicio en las primeras 3 fases se usó el conjunto de datos “Exclusivo o”.

De acuerdo al conjunto de datos seleccionado se emplea como entrada la relación X1​X2, la cual permite representar la interacción de las variables originales.​

## Resultados del cambio de la función de activación en relación con la velocidad de aprendizaje y la forma de la frontera de decisión

**Respuesta:** Teniendo en cuenta la configuración antes descrita, el cambio de función de activación entre las 4 disponibles permite evidenciar que con la totalidad de las funciones en tiempo muy corto se llega al modelo con el menor número de prueba, esto se debe a que la distribución de los datos puede ser modelada de manera adecuada por la totalidad de la funciones de activación siendo resultados que se desarrollan entre la pérdida de prueba 0.018 y 0.017 y la pérdida de entrenamiento entre 0.009 y 0.010, se observan variaciones en el peso que cada función entrega a las neuronas, siendo la Sigmoide la que más emplea neuronas para el modelamiento y la lineal la que menos emplea.

## Resultado de la modificación de la tasa de aprendizaje con relación a la convergencia de la red a la solución

**Respuesta:** Frente a este punto es importante resaltar que se debe realizar un punto de equilibrio ya que si la tasa es muy alta, el modelo puede sobrepasar fácilmente el modelo óptimo lo que terminaría en un modelo que no converge o que converge de forma inestable, acercándose al valor óptimo sin lograr llegar a este.

Por otra parte, si la tasa es muy baja, se puede tomar mucho tiempo para llegar al punto adecuado de convergencia, lo que podría terminar en un entrenamiento ineficiente.

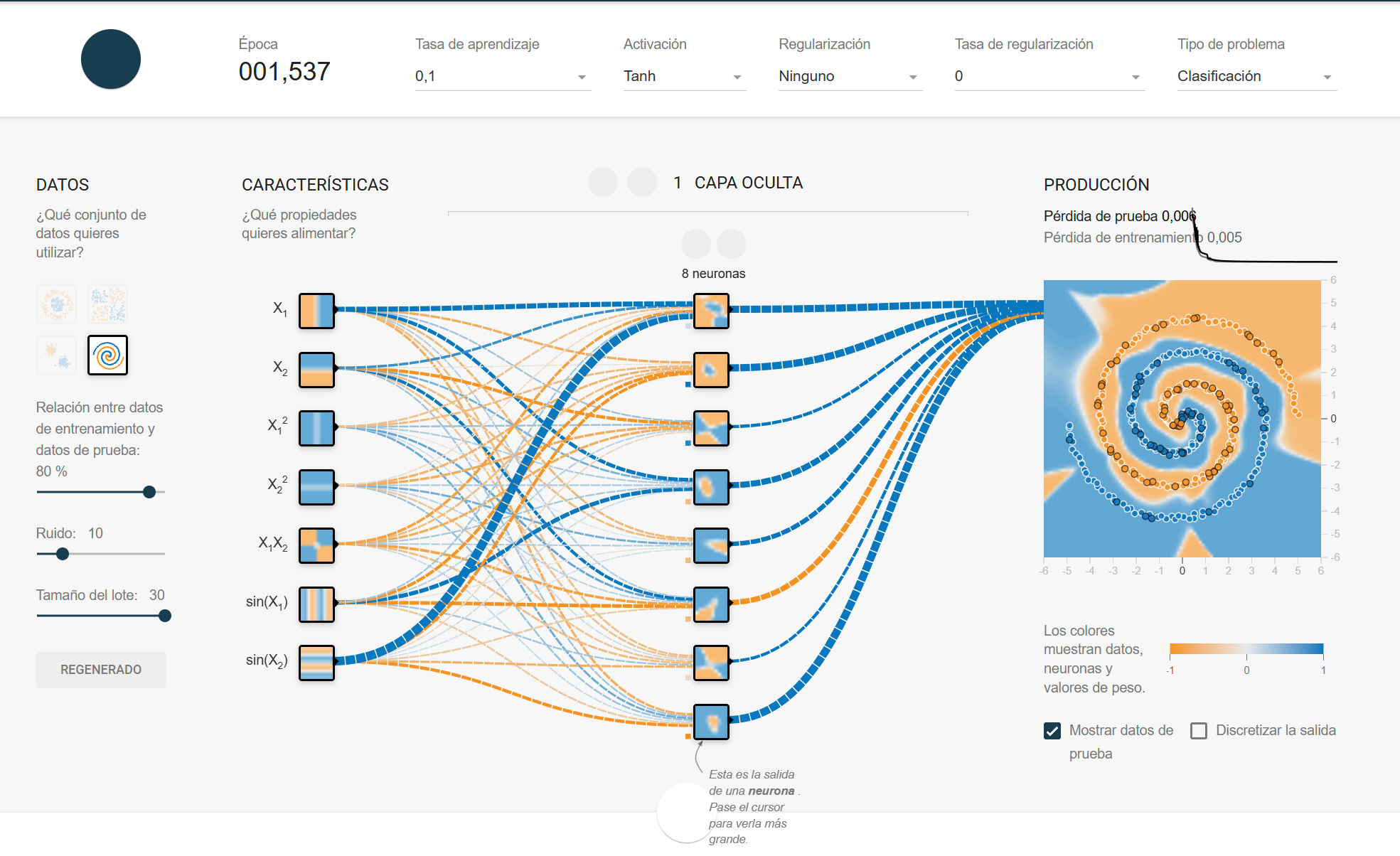
## Resultado del ajuste de las capas y las neuronas con relación a la complejidad de la red y su capacidad de modelar datos

**Respuesta:** Para la configuración antes descrita, el incremento de las capas y del número de neuronas no afecta en gran medida la generación del modelo óptimo, observando el resultado de las 6 capas ocultas, cada una conformada por 8 neuronas genera el mismo resultado que al hacer uso de 1 capa y 1 neurona, la variable que realmente afecta el modelo son las propiedades de alimentación del modelo.

# Exploración

## Resultado de la primera distribución

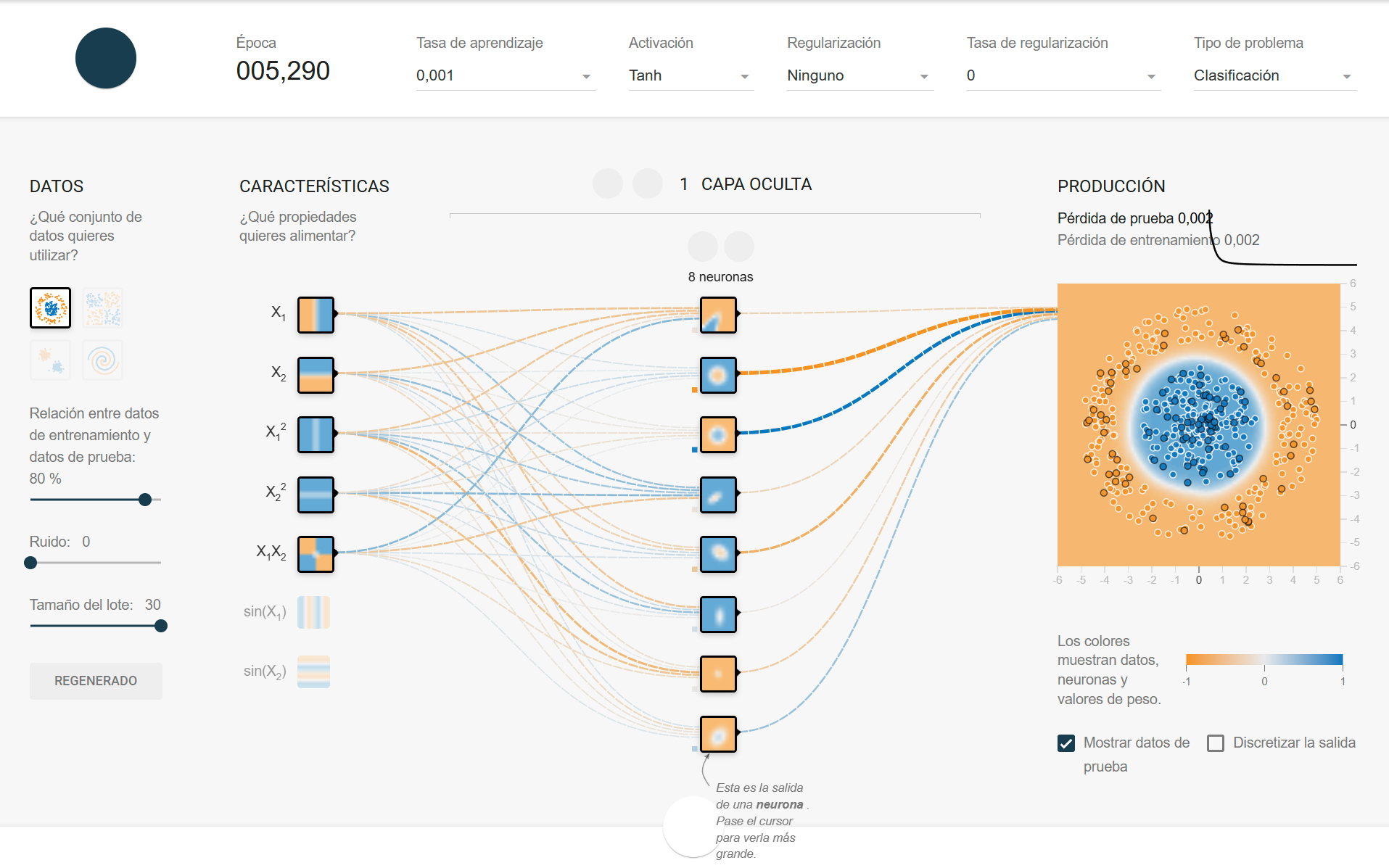
**Respuesta:** Para el proceso de prueba se hizo uso del conjunto de datos “Espiral”, teniendo en cuenta este tipo de distribución de los datos, se hacer uso de una función de activación no lineal, dentro de las pruebas se identificó que de las funciones no lineales “Tanh” al ser una función centrada en 0 y al oscilar entre -1 y 1 representa mejor el modelo, se emplean todas las características de entrada para procesar de forma más eficiente los datos, las cuadráticas para entregar información de la espiral, las trigonométricas para visualizar los patrones cíclicos y las de producto para identificar la interacción entre las dos variables.



Con la tasa de aprendizaje “0,1” se logra contar con un modelo con 0.006 de pérdida en prueba y 0.005 de pérdida en entrenamiento, obteniendo de esta forma un modelo óptimo y estable.

## Resultado de la segunda distribución

**Respuesta:** Para el proceso de prueba se hizo uso del conjunto de datos “Círculo”, teniendo en cuenta este tipo de distribución de los datos, se debe hacer uso de una función de activación no lineal, dentro de las pruebas se identificó que de las funciones no lineales “Tanh” al ser una función centrada en 0 y al oscilar entre -1 y 1, lo que los hace más eficiente para los modelos simétricos como el círculo, se emplean las primeras 5 características de entrada para representar de forma más eficiente las cuadráticas para representar mejor la simetría del círculo y las de producto para identificar la interacción entre las dos variables, especialmente para identificar la curvatura del círculo.



Se retira el ruido de este modelo y se reduce la tasa de aprendizaje “0,001” para llegar con mayor precisión al modelo óptimo, teniendo como resultado un modelo con 0.002 de pérdida en prueba y 0.002 de pérdida en entrenamiento, obteniendo de esta forma un modelo óptimo y estable.

# Conclusión

Como resultado de esta actividad se identifica la importancia de conocer en detalle la aplicabilidad de las funciones de activación de acuerdo al tipo de conjuntos de datos que se quieren modelar, al igual que la selección de las características de entrada, por otra parte al ajustar la tasa de aprendizaje, se debe buscar un equilibrio, ya que una tasa alta permite un aprendizaje rápido al principio, pero si es demasiado alta puede dificultar la convergencia óptima del modelo, por otra parte una tasa baja, permite realizar ajustes más precisos pero puede hacer que el entrenamiento tome demasiado tiempo.

Es así como el conocimiento de los datos permite al profesional en Ciencias de Datos hacer uso eficaz de los parámetros disponibles para la construcción de este tipo de procesos para llegar a un modelo óptimo y estable.