

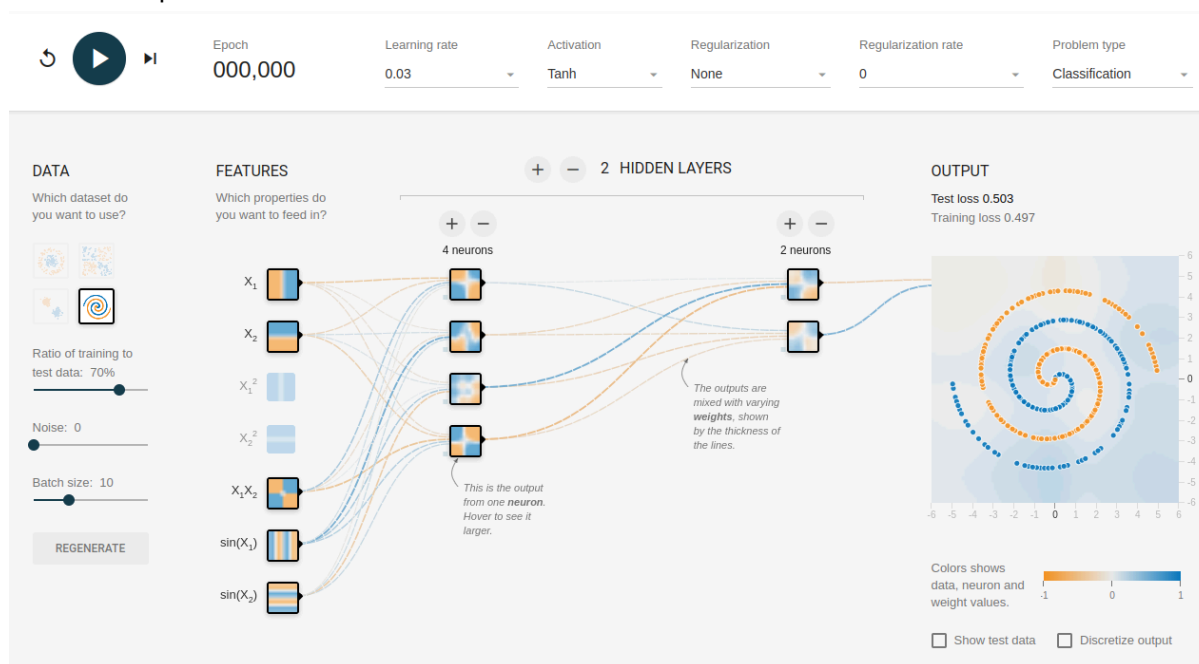
Evaluación del desempeño de una red Neuronal para clasificación, variando parámetros como: tasa de aprendizaje, función de activación y regularización.

José Manuel Ordóñez Silva.

Resumen: En este estudio se investigó el desempeño de una red neuronal multicapa para la clasificación de datos de prueba, dispuestos en la plataforma «Tensor Flow Playground [1]» se analizó el comportamiento del sistema para distintos valores de tasa de aprendizaje, seguido del comportamiento de la red al modificar la función de activación y finalmente se comparó el resultado al incorporar regularización. se encontró que los mejores parámetros fueron aumentar el número de capas de 2 a 3, y usando la función de activación Tanh.

Resultados y Análisis:

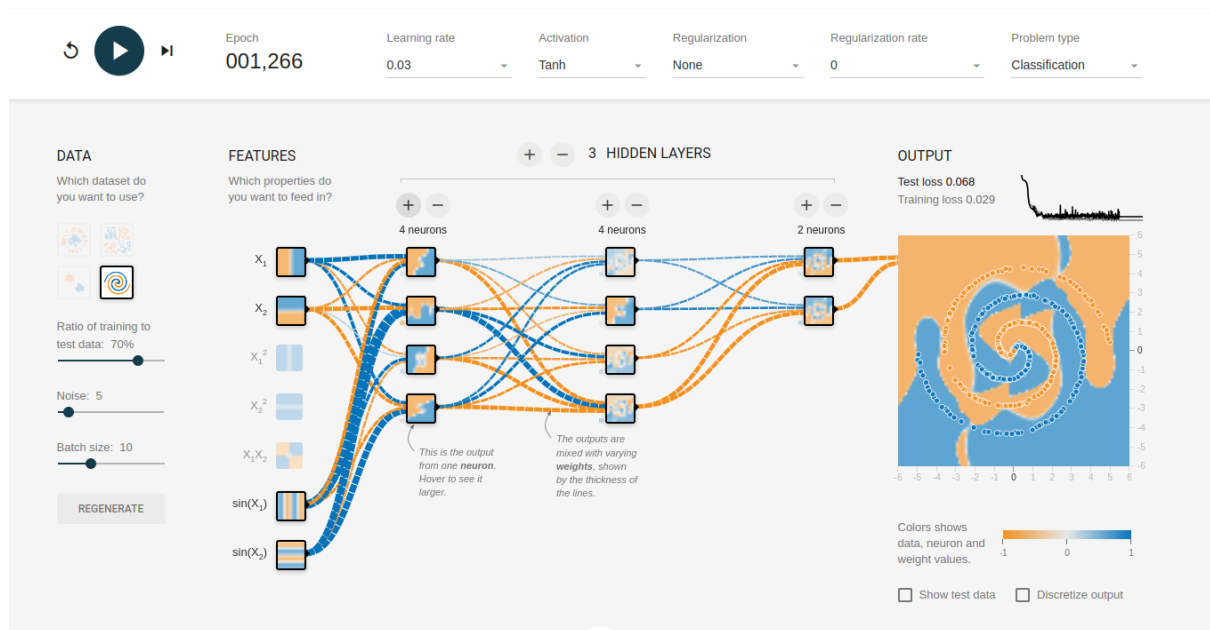
Modelo Base (Baseline): Se caracterizó el modelo utilizando la configuración por defecto obtenida al abrir por primera vez la interfaz, y se configuró para la clasificación de la muestra 4 (Panel “Data” Inferior Derecha) dividiendo los datos para un 70% para entrenamiento y 30% para validación, y se añadieron como características o descriptores a $\sin(X_1)$, $\sin(X_2)$, el resto de parámetros se mantuvieron por defecto.



Incremento de 2 capas , a 3 capas.

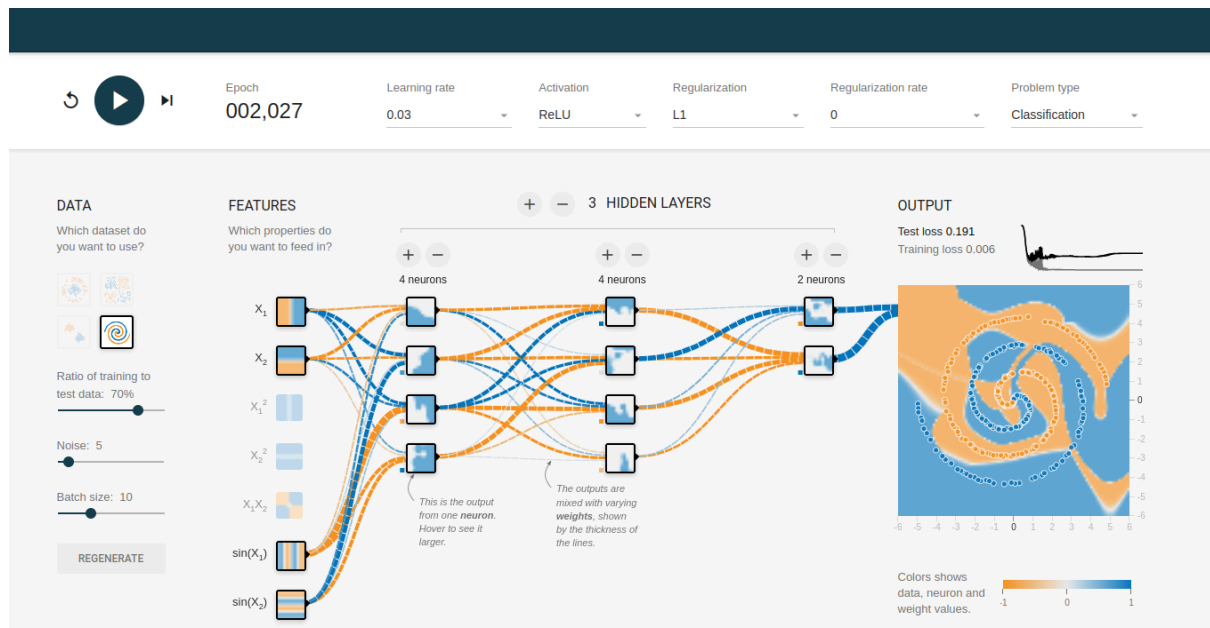
Función Tanh, sin regularización.

De esta forma la NN fue capaz de clasificar los datos al rededor de la epoca 1200, y se minimizaron las funciones de perdida de entrenamiento y validación en 0.068 y 0.029 respectivamente.



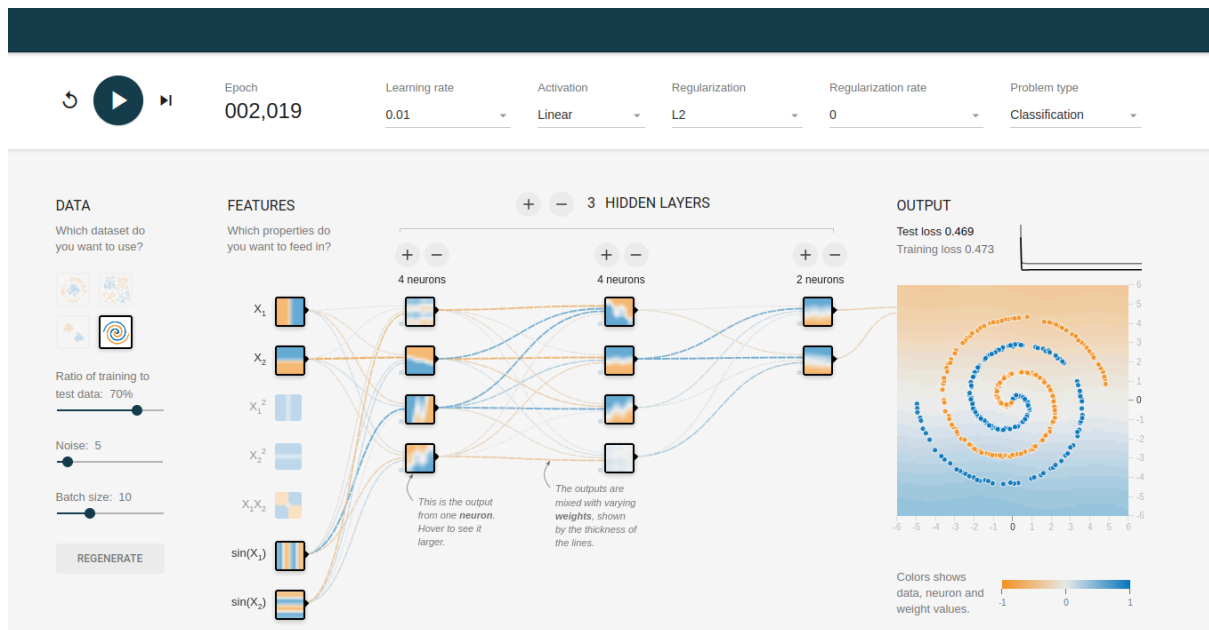
Función de activación Relu, y Regularización L1.

Se observa que en este caso, ocurre una diferencia entre la función de perdida para entrenamiento y validación lo que sugiere que el modelo tuvo un overfitting entre el conjunto de datos de entrenamiento y validación.



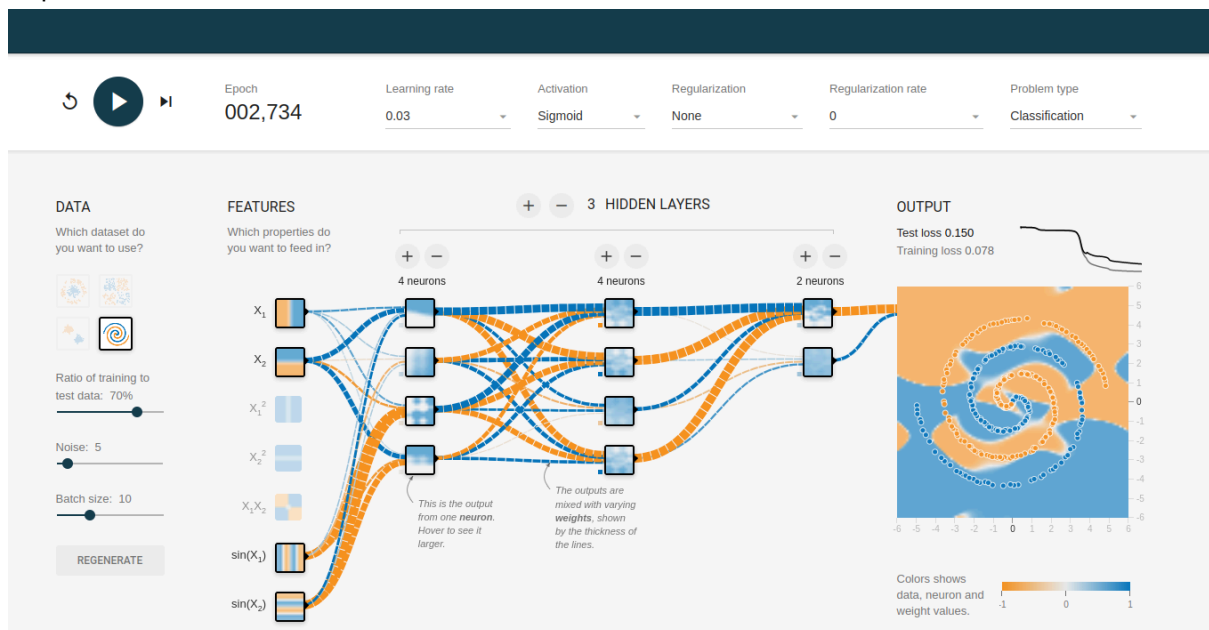
Función de activación Lineal y Regularización L2.

Despues de 2000 epocas la configuración no muestra ninguna convergencia en la clasificación de los datos, sin embargo hay que destacar que la complejidad de los datos no satisface una forma "lineal" por lo que la complejidad del problema debe estar acompañada de una función de activación que pueda permitir mas flexibilidad.



Función de Activación de la Sigmoid sin regularización.

Para la función de activación sigmoide los datos se clasificaron correctamente para la epoca 2500, y los valores de la función de perdida para entrenamiento y validación fueron 0.150, 0,028 respectivamente.



Conclusión:

Se observa que cada tipo de problema tiene una complejidad determinada y que la solución que se pretenda obtener a través de de una Red Neuronal puede tener distintas implementaciones, sin embargo se debe buscar los mejores parametros para satisfacer la complejidad de los datos. Los parametros que mejor funcionaron para la clasificación fue Función de activación Tanh y Sin regularizacion.