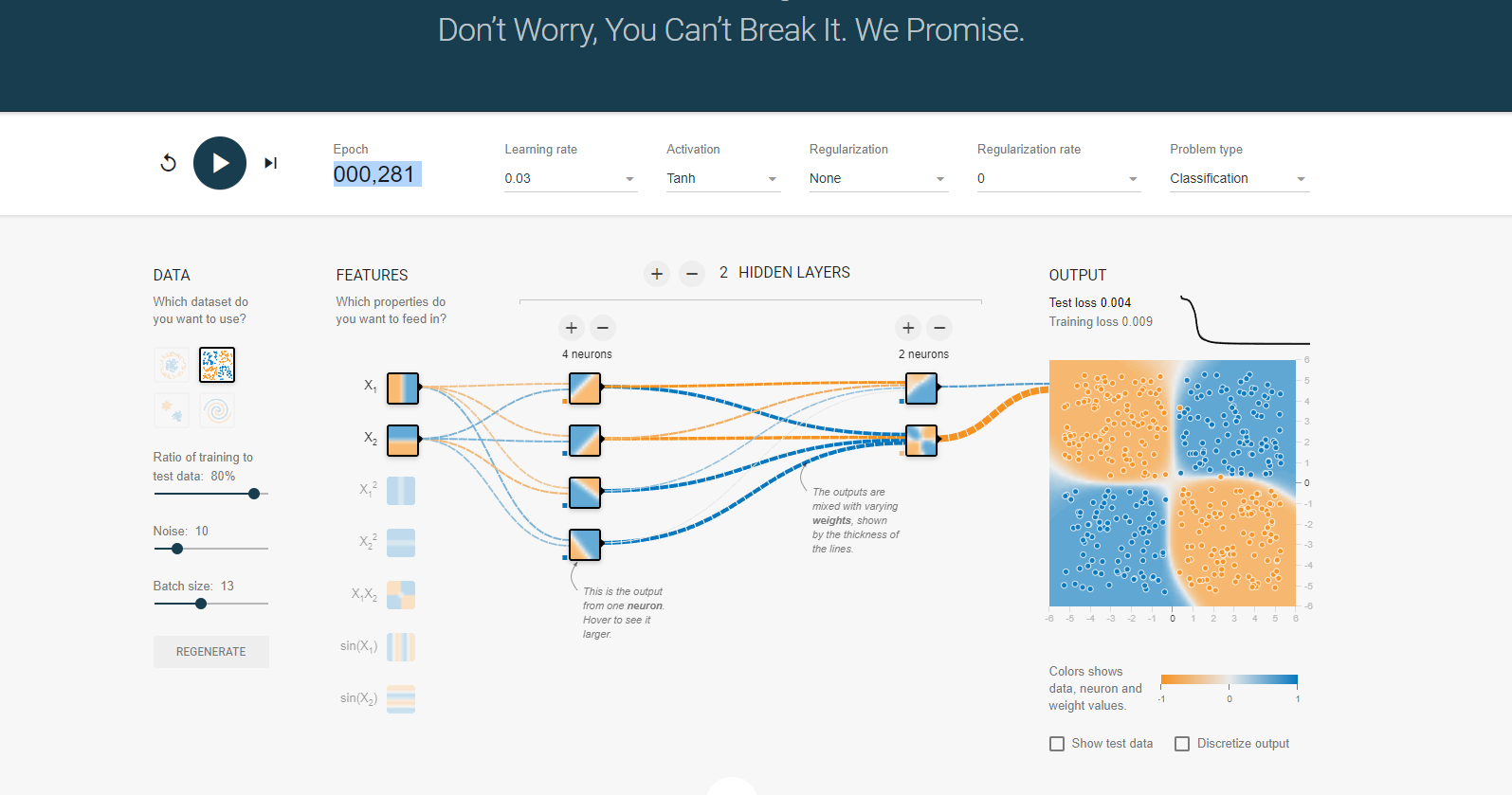
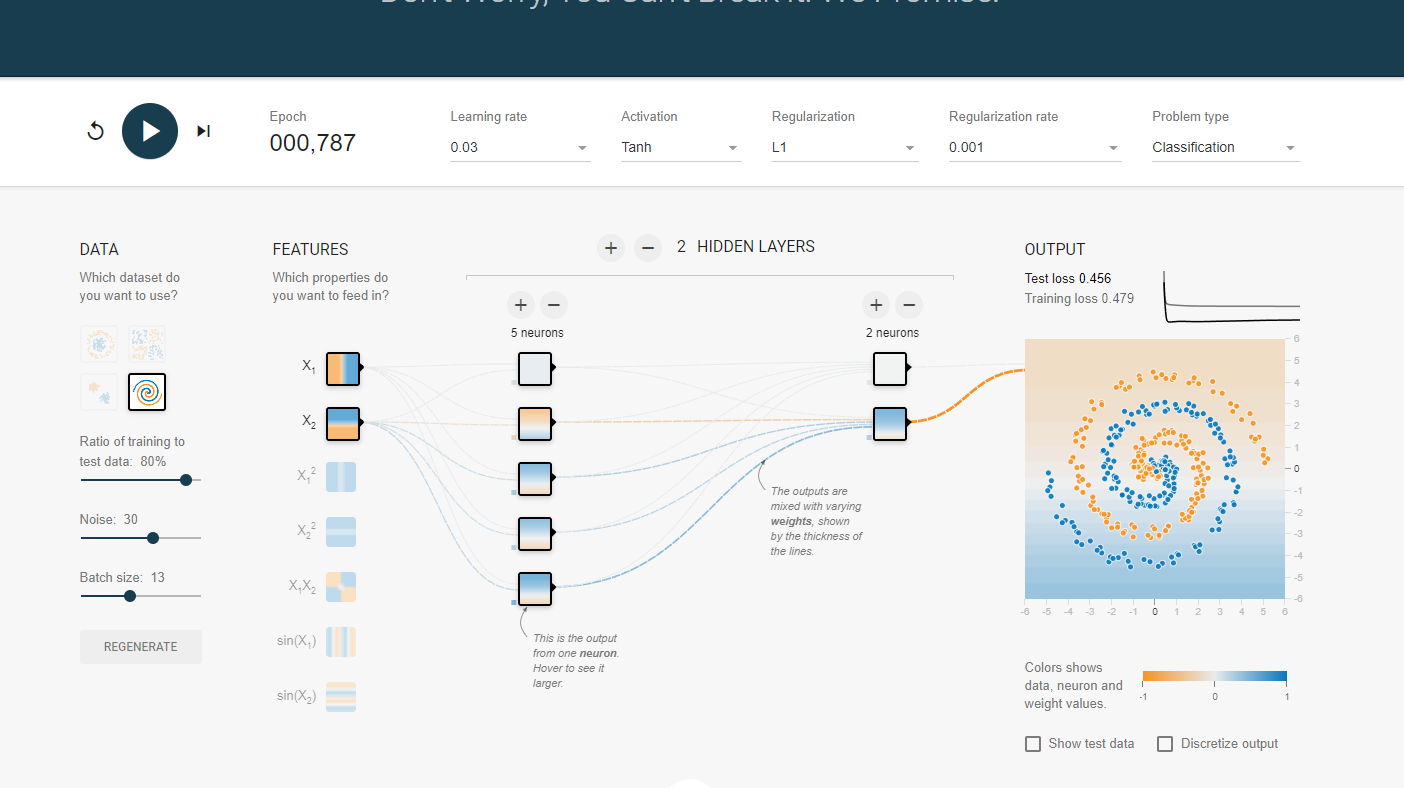
1. Se cambia la cantidad de datos usados para el conjunto de entrenamiento y test dejando 80% y 20% respectivamente. Adicional, se agrega ruido a los datos de 10 para tener datos más realistas. Se dejan los demás hiper parámetros por default.



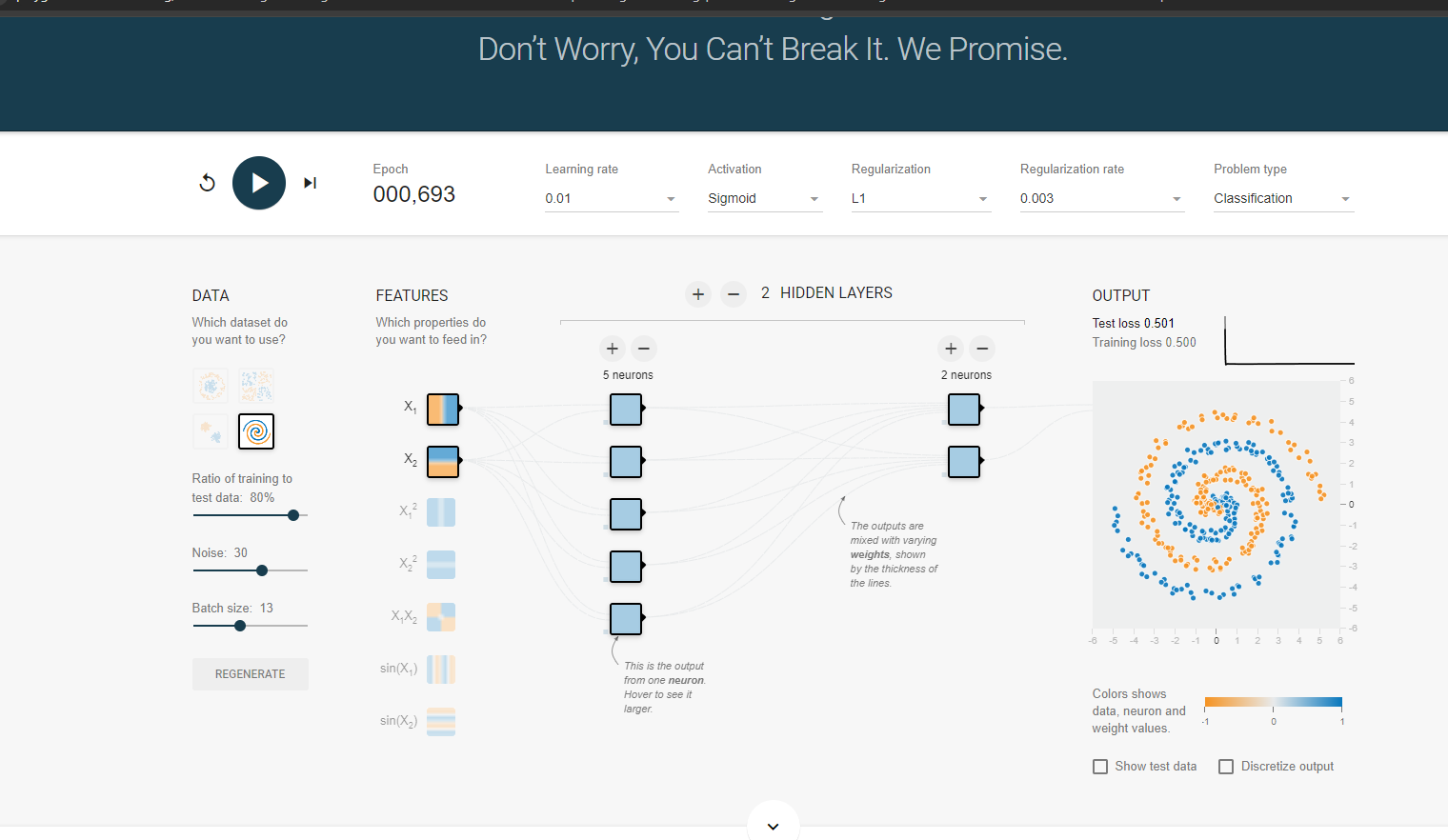
De acuerdo con los resultados obtenidos se puede ver un entrenamiento correcto en 281 épocas, donde es capaz de clasificar la mayoría de los datos correctamente.

1. En este ejemplo se cambió la distribución de los datos a la forma spiral, se aumentó el ruido y el batch size. Adicional, se le agrega regularización de 0.001.



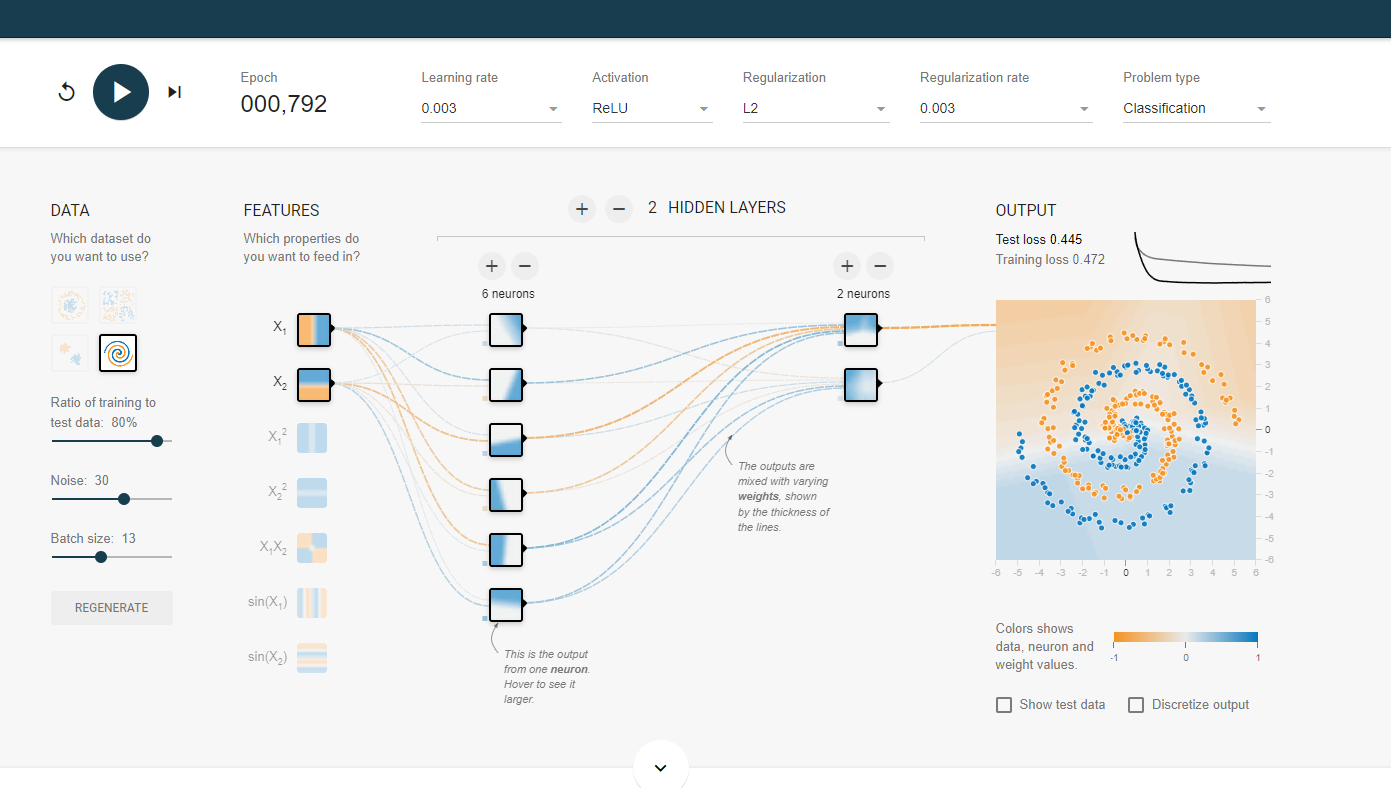
Se evidencia que después de 787 épocas el modelo no logra ajustarse bien a los datos, para poder clasificar de forma correcta. Incluso agregando una neurona adicional en la capa oculta.

1. Se continua con el ejemplo anterior, se cambia la función de activación a Sigmoid y se aumenta el learning rate.



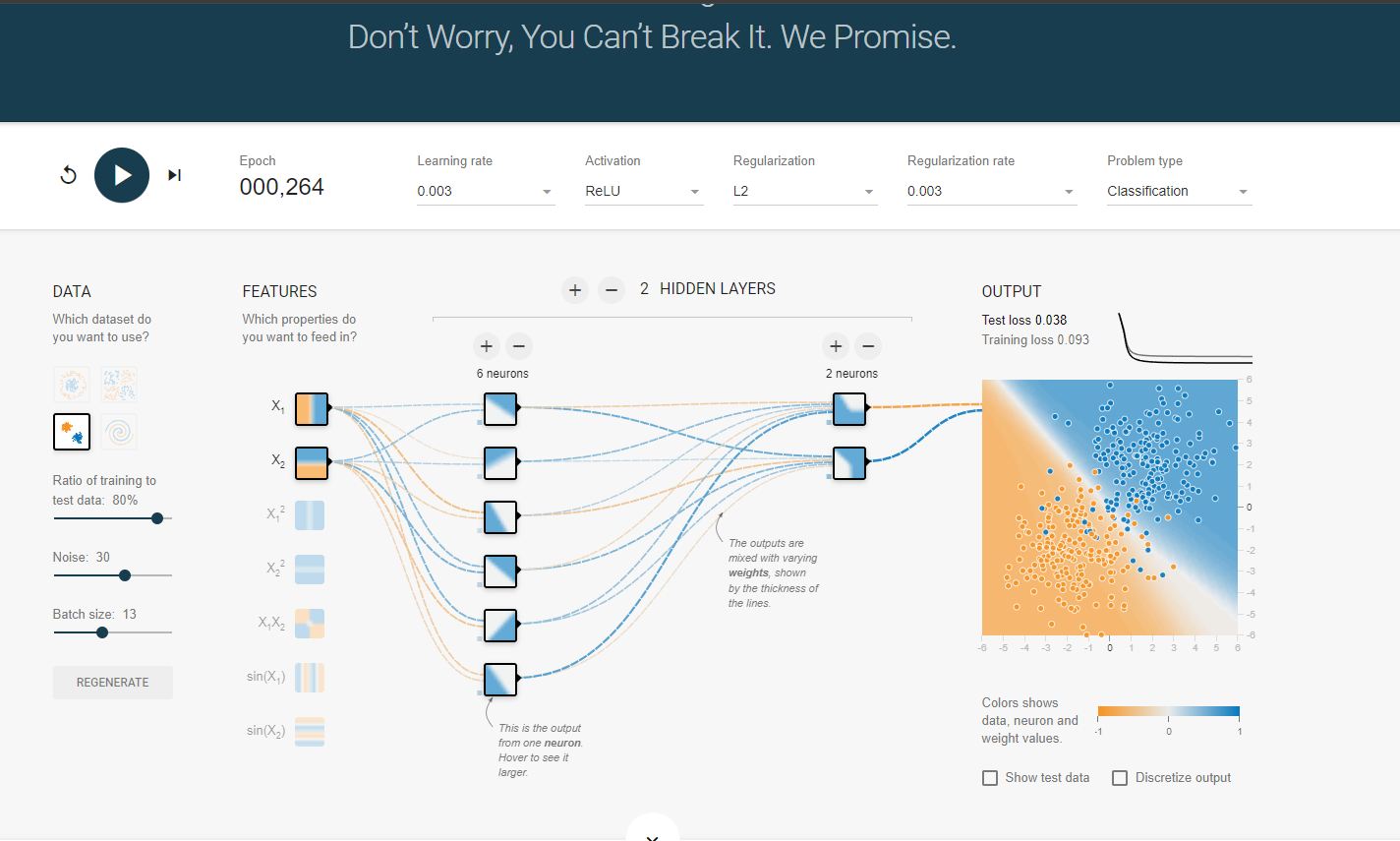
El modelo después de 693 épocas no logra ajustarse y en cambio se muestra con un resultado peor que el obtenido con los hiper parámetros del punto anterior.

1. Se cambia el learning rate, la función de activación y la regularización. Y se agrega una neurona en la capa profunda.



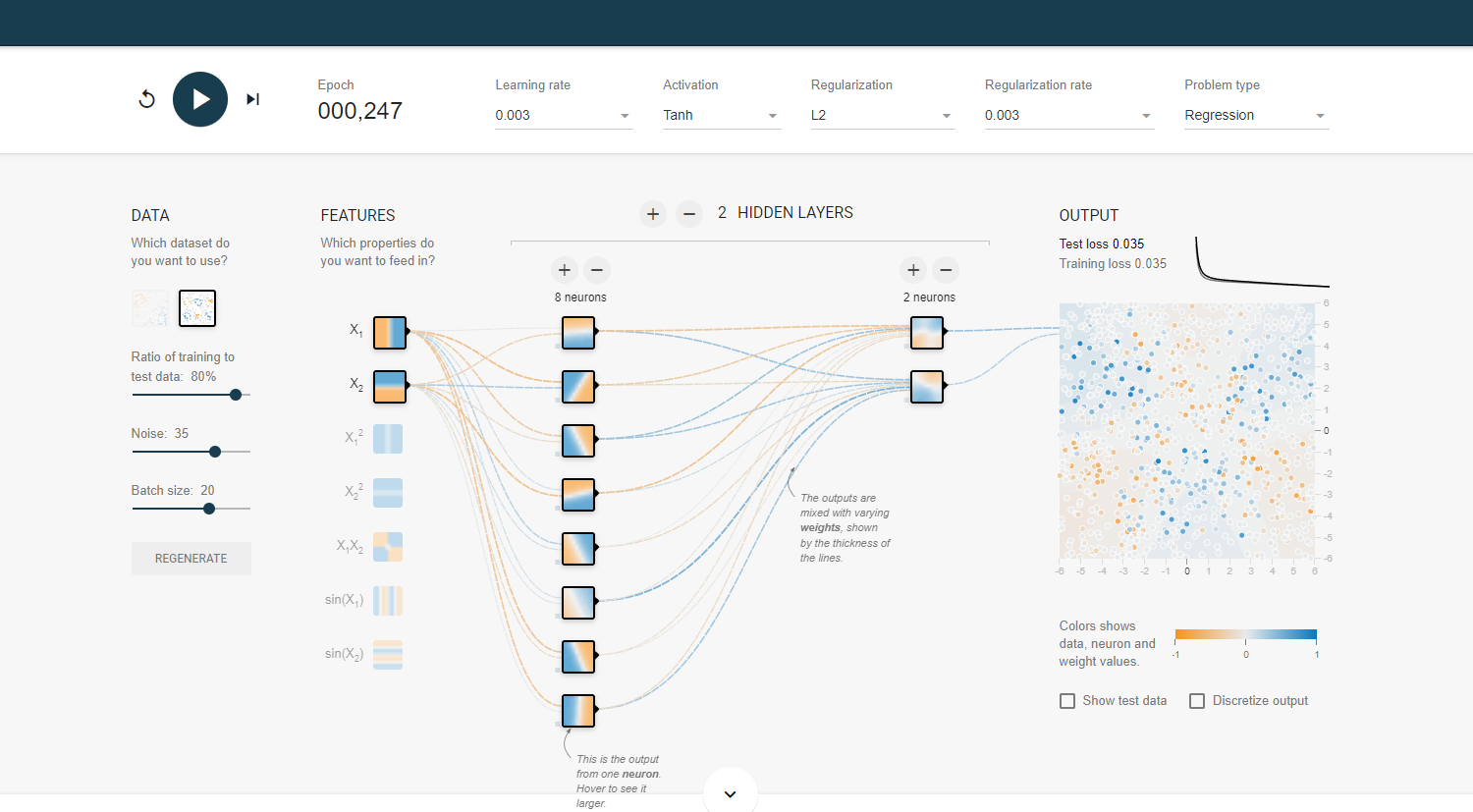
Se obtuvieron resultados similares al punto 2 después de 792 épocas.

1. Se mantienen los hiper parámetros del punto 4, pero se cambia la distribución de los datos a Gaussian.



Se evidencia un comportamiento bueno, alcanzando métricas de clasificación deseadas.

1. Se realiza un ejemplo para regresión aumentando las neuronas de la capa profunda a 8, cambiando la función de activación a Tanh, los demás parámetros se conservaron igual al punto 4. La distribución de datos usada fue MultiGaussian.



Los resultados arrojados son buenos.