Topic 2

Q1: Using the default settings, what is the best activation function?

* 在 default dataset circle 中，只要是非線性的 activation function，都能完美的區分藍點和橘點，然後 tanh、relu 收斂的速度遠比 sigmoid 快很多。

Q2: Using the default settings, what happens if you increase # layer?

* 在 default dataset circle 中，當 hidden layer 數量為 2-4 層的時候收斂速度快且準確度高，從 hidden 大於等於 5 時，收斂速度大幅降低，且準確度明顯下降。

Q3 : Playing with the Spiral dataset (bottom right)

Q3-1: Changing the activation function

Q3-2: Changing the number of layers

Q3-3: Adding additional features

* 使用 tanh 的收斂速度和結果最佳
* hidden 兩層的收斂速度最快且準確度最好，且神經元的數量要用多一點。
* 額外加了 sin(X1)、sin(X2)，收斂速度提高非常多，且 test loss 大幅降低。
* 下圖提供了最後自己調整後的最佳參數

一張含有 文字, 螢幕擷取畫面, 地圖, 圖表 的圖片

自動產生的描述