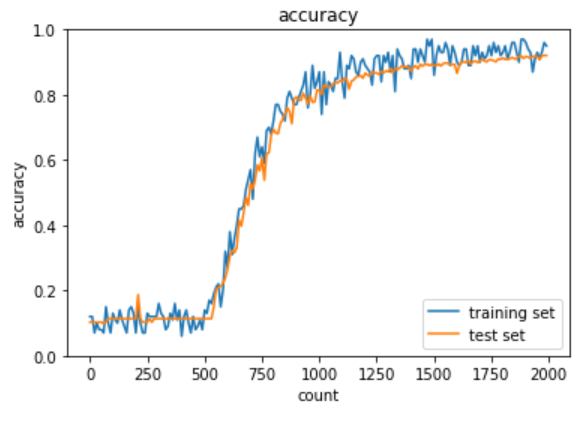
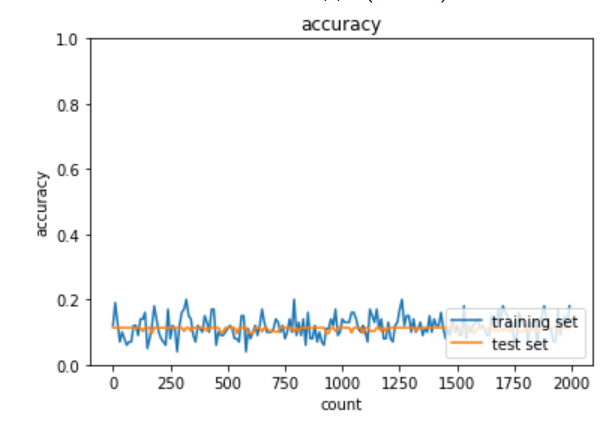
ラビットチャレンジ課題　作成者:内山 貴雄

●深層学習前編day2 勾配消失問題実装演習レポート

サンプルコード(2\_2\_2\_vanishing\_gradient\_modified.ipynb)において、

Mnistデータを入力層(ノード数784)、中間層2層(ノード数はそれぞれ40,20)、出力層(ノード数10)のニューラルネットワークにて2000回学習し、その際の学習数を横軸、誤差を縦軸としてプロットした。それらを下に置いたが、活性化関数をsigmoid関数としてプロットしたものが左図であり、ReLU関数としてプロッタしたものが右図である。

下図により、下の左図において勾配消失が起きていることが分かる。



　この勾配消失問題を解消する方法として、下記に示される重みの初期値を設定する際に、重み行列の各要素を前のノード数の平方根を用い調整する方法(XavierとHe)が挙げられる。

　サンプルコードでは、活性化関数がsigmoid関数の時、重みの初期値調整の方法としてXavierを用い、活性化関数がReLU関数の時、重みの初期値調整の方法としてHeを用いた際の計算を実施した。

重みの初期値調整の方法(Xavier):

重みの要素を、前の層のノード数の平方根で除算する。

重みの初期値調整の方法(He):

重みの要素を、前の層のノード数の平方根で除算した値に対し√２をかけ合わせる。

また、今回サンプルコード以外に、活性化関数がsigmoid関数の時、重みの初期値調整の方法としてHeを用い(下部左図)、活性化関数がReLU関数の時、重みの初期値調整の方法としてXavierを用いた際(下部右図)の計算を実施したした結果を下に示す。

