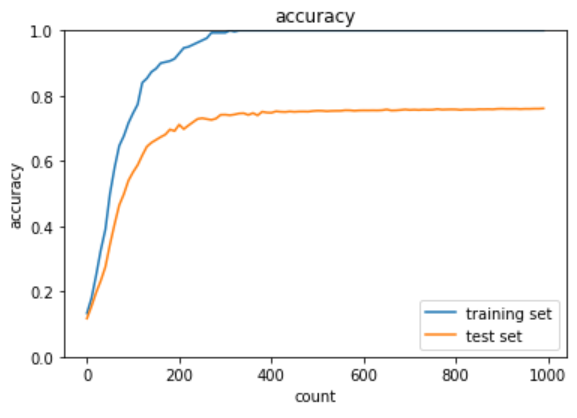
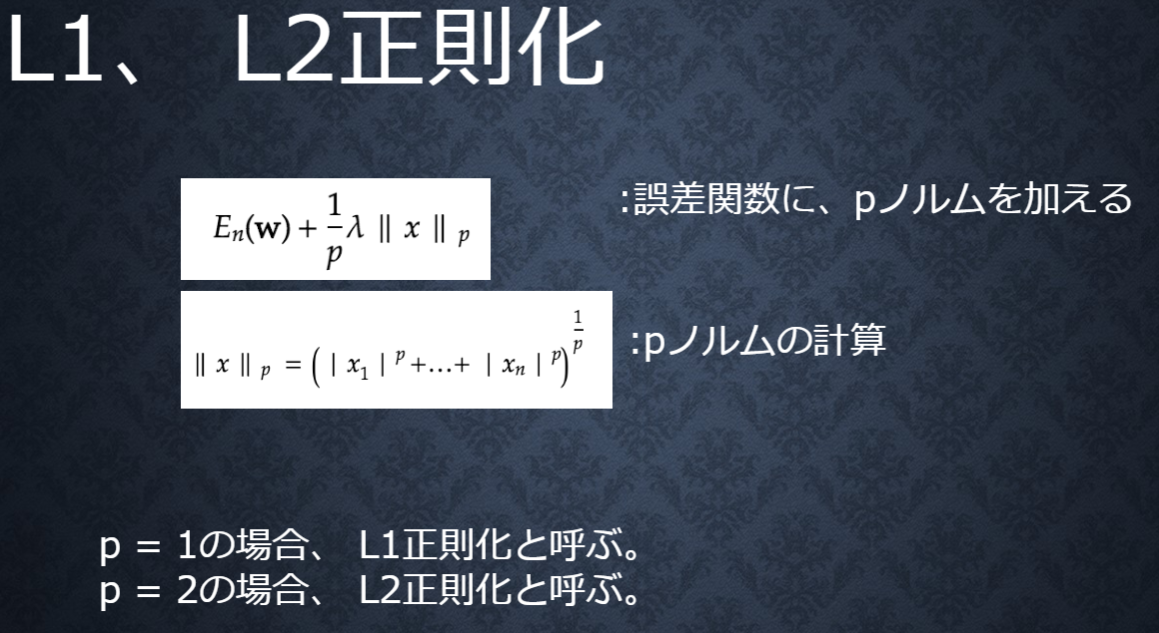
ラビットチャレンジ課題　作成者:内山 貴雄

●深層学習前編day2 過学習レポート

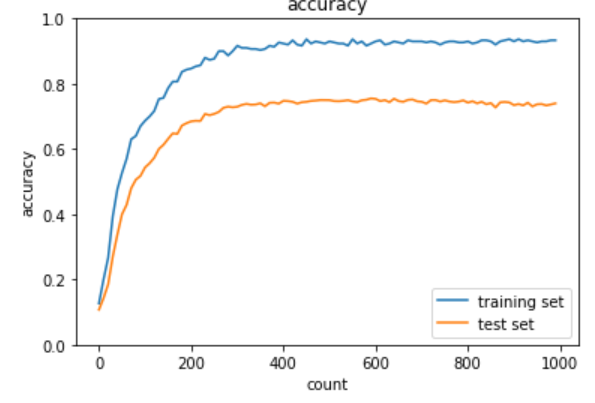
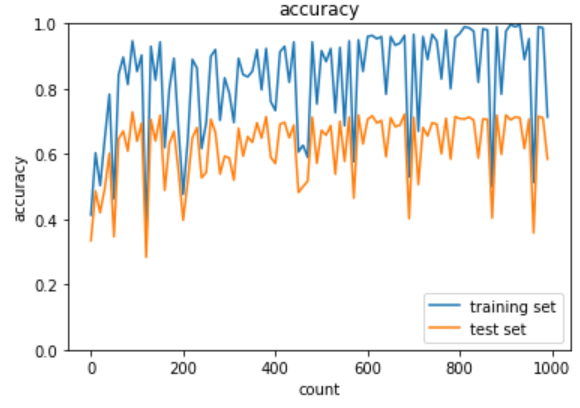
サンプルコード(2\_5\_overfiting.ipynb)において、Mnistデータを入力層(ノード数784)、中間層5 層(ノード数は全て100)、出力層(ノード数10)のニューラルネットワークにて1000回学習し、その際の学習数を横軸、誤差を縦軸としてプロットした。その結果が下図に示すものであるが、過学習が起きていることが分かる。



そのため、学習の際最小化する式に、各重みの値に依存する下記に示す正則化項を導入することにより、重みの大きさが過剰な値になり過学習となることを防止する。



　上記式に則り、L1,2正則化項を導入し、前述の計算を実施した結果をそれぞれ下記の左図(L1正則化) 、右図(L2正則化)に示す。



上記の結果を見ると、L1正則化の結果の方が精度にバラつきがあることが分かる。これは、一般的にL2正則化はモデルの過学習を防ぎ、精度を高めるために用いられ、L2正則化を用いた結果の方がより滑らかで表現力に優れているのに対し、L1正則化は学習において重みが0になりやすく、用途として不要な説明変数をそぎ落とす次元圧縮のために用いられることが多いためであることによる。

また、正則化項におけるコードの内の下記係数を0.2,0.15とし、前述の学習を実行した結果が下記のそれぞれ左図(0.2)と右図(0.15)である。左図においては、200回以降精度が低下し、右図においては訓練データとテストデータでの精度の差が下記係数が0.1のときよりも若干小さいが、テストデータでの精度がそれよりも低いため、この学習において、

下記係数は0.1付近が適当と考えられる。

# 正則化強度設定 ======================================

weight\_decay\_lambda = 0.15

# =================================================

