糖尿病データへの機械学習の適用　　　　　　　　　　 20c2007影山亮太

1.概要:

教科書で行った,ニューラルネットワークを使って,糖尿病のデータをテストデータと訓練データで分けて平均絶対誤差と1年後の糖尿病進行度の予測の正答率を出力すると

約:43.1,+-43.1の誤差と確認できた.

今回行うのは糖尿病データを線形回帰を使って,糖尿病データをテストデータと訓練データに分けて平均絶対誤差と1年後の糖尿病進行度の予測の正答率を出力する事が目的である.

2.解析結果:

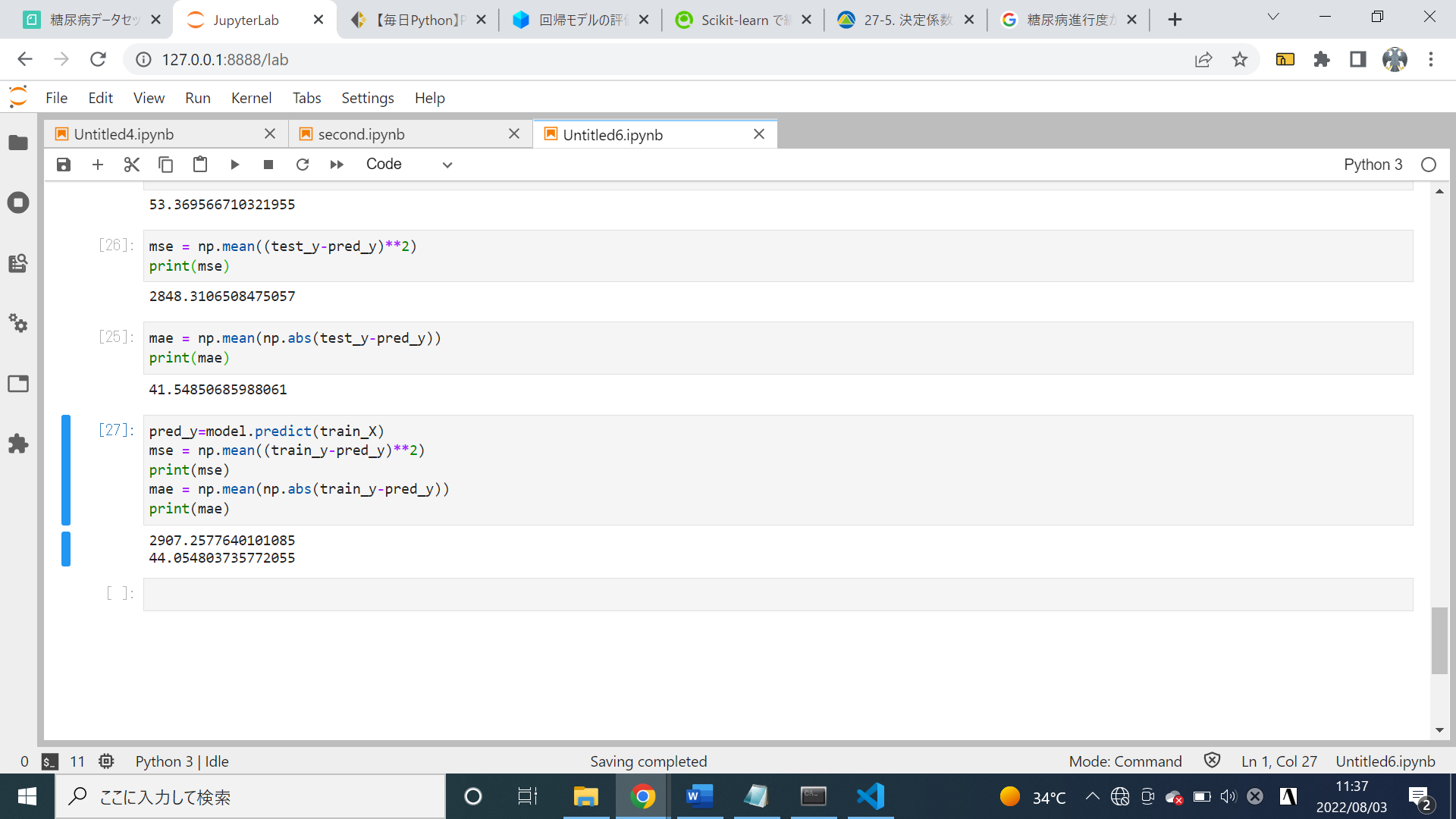


図 　線形回帰で解析した結果(テストデータ)

上の図は,糖尿病データを線形回帰で訓練データを解析した結果である.

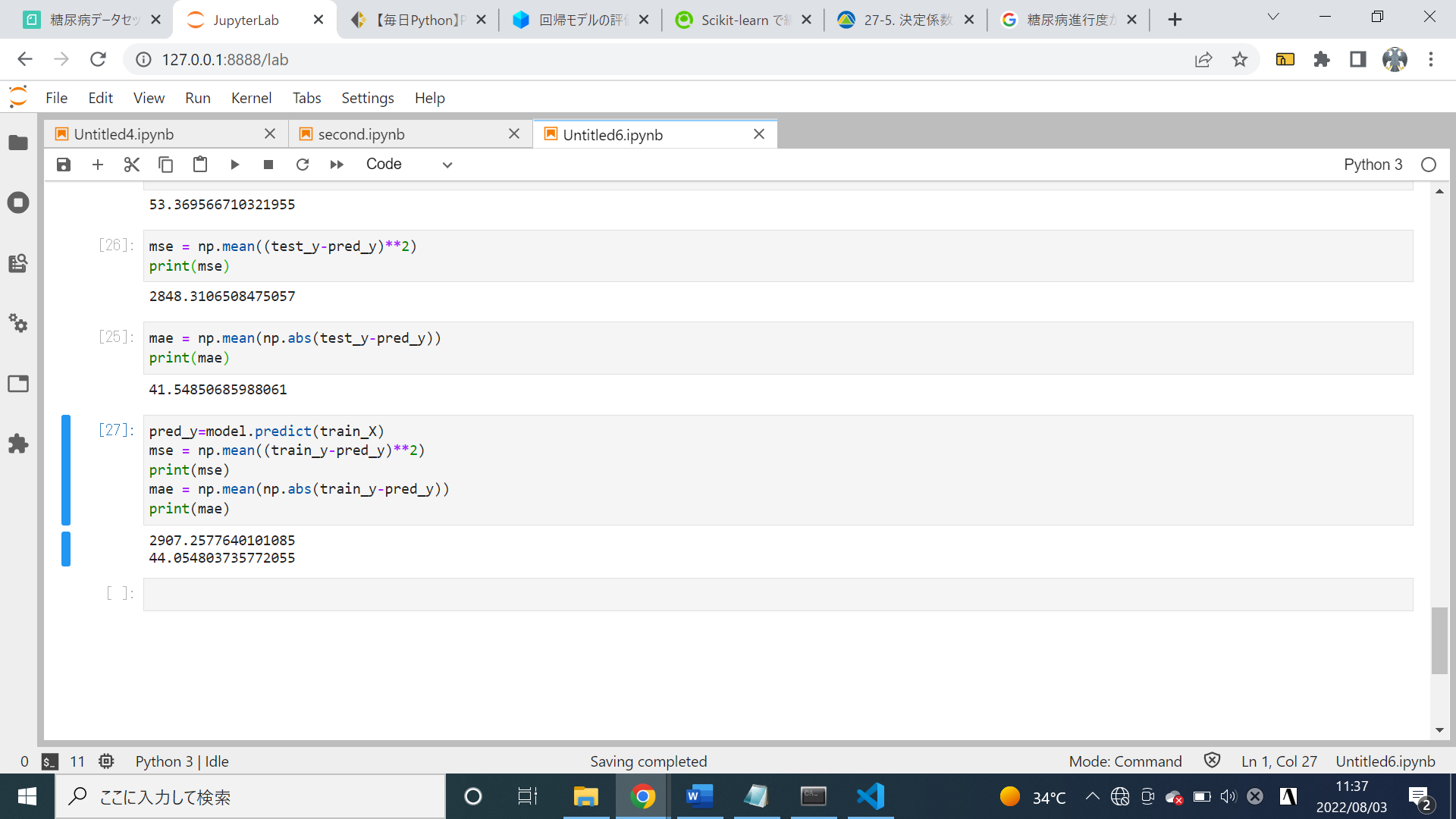


図 　線形回帰で解析した結果(訓練データ)

上の図は,糖尿病データを線形回帰で訓練データを解析した結果である.

さらに,ニューラルネットワークを使って,訓練データとテストデータを分割して解析した結果は以下の様になる.

グラフィカル ユーザー インターフェイス, テキスト

自動的に生成された説明

図 ニューラルネットワークで解析を行った結果

図1,図2と図3を見比べると同じデータで解析を行ったにもかかわらず,精度が異なったという事がいえる.

3.結果に対する考察:

解析結果で,同じ糖尿病データにもかかわらず線形回帰で解析した結果とニューラルネットワークで解析をした結果では,精度がわずかながら異なることがわかった.

ニューラルネットワークと線形回帰で行った時の精度が異なる理由としてニューラルネットワークでは,損失関数として,平均絶対誤差を用いているのに対して線形回帰では,平均二乗誤差を使ってテストクラスと訓練データを出力している.この損失関数の違いにより,ニューラルネットワークと線形関数で糖尿病データで分析した結果が異なることがいえる.