

PFN2019 インターン課題

氏名：関口舜一

1 課題 3

1.1 結果

図 1 と図 2 に SGD を用いた場合, 図 3 と図 4 に Momentum SGD を用いた場合の, トレーニングデータへの平均損失と平均精度, バリデーションデータへの平均損失と平均精度を示す. 移動平均は連続する 75 個のデータの平均で描画している. また, iterations は, 学習を行った回数を意味する. ハイパーパラメータについては, 課題の説明に記された例と同じものを使用した. Batch size は 10, epoch は 50 とした.

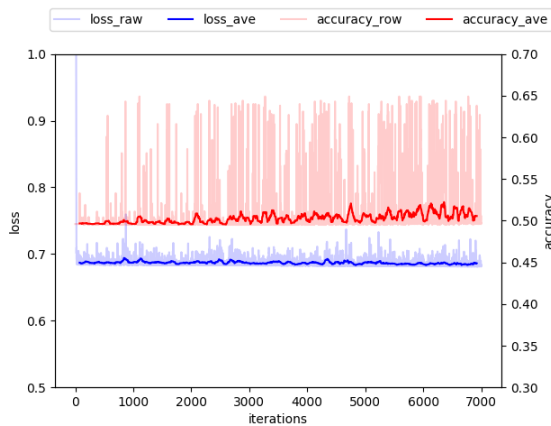


図 1 トレーニングデータでの
平均損失と平均精度 (SGD)

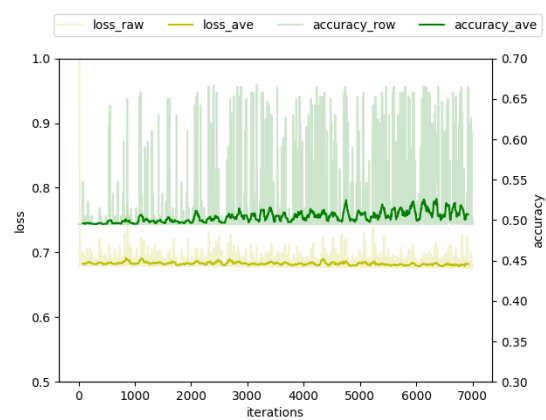


図 2 バリデーションデータでの
平均損失と平均精度 (SGD)

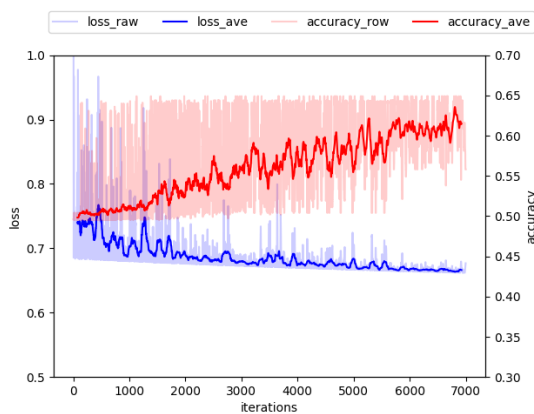


図 3 トレーニングデータでの
平均損失と平均精度 (MomentumSGD)

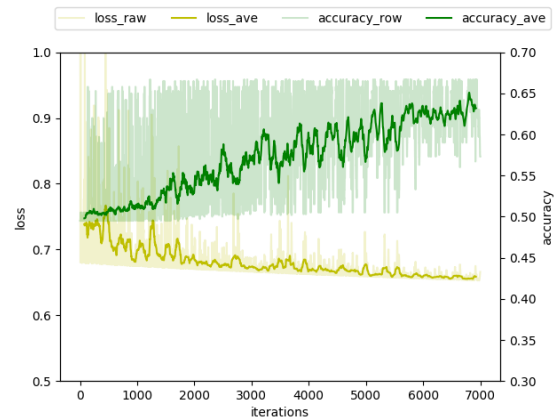


図 4 バリデーションデータでの
平均損失と平均精度 (MomentumSGD)

1.2 考察

図 1, 2 と図 3, 4 を比較すると, SGD よりも, Momentum SGD において精度が高くなり, 損失も小さくなっていることが分かる. これは, Momentum SGD が 1 つ前の勾配を使うため, 局所解から抜け出しやすい性質をもっているためである.

ると考えられる。

また、図 1 と図 2, 図 3 と図 4 を比較すると、トレーニングデータとバリデーションデータでの挙動がほぼ一致していることが分かる。これは、トレーニングデータとバリデーションデータでの入力データに違いが少ない（分布が似ている）ことを意味している。（データを分割する際に、全体的なラベルの割合を保ったまま分割するように工夫した。）なお、分割の際に選ぶ **index** を乱数を用いてシャフルし学習させているのにもかかわらず、毎度、同じようにトレーニングデータとバリデーションデータの挙動が似るため、今回の用意されたデータは比較的偏りが少ないものではないかと考えられる。

2 課題 4

2.1 取り組んだ内容

課題として、集約ステップで用いる際の線形関数を多層パーセプトロン（2 層）に置き換えることを行った。活性化関数には **Relu** 関数を用いた。

2.2 結果

図 5 と図 6 にトレーニングデータとバリデーションデータに対する平均損失と平均精度の結果を示す。ハイパーパラメータについては、課題の説明に記された例と同じものを使用した。Batch size は 10, epoch は 20 とした。また、予測結果を **prediction.txt** に記した。

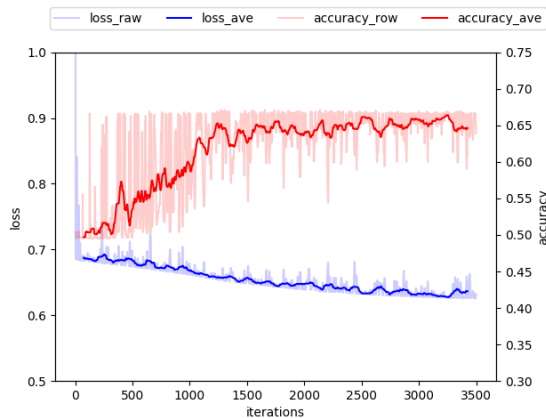


図 5 トレーニングデータでの平均損失と平均精度（MLP）

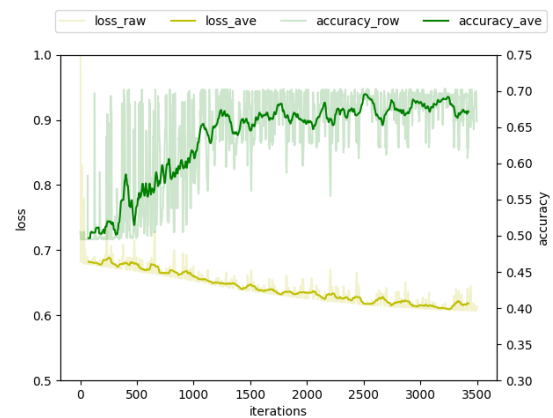


図 6 バリデーションデータでの平均損失と平均精度（MLP）

2.3 考察

図 5 から、課題 3 よりも平均損失、平均精度ともに向上していることが分かる。これは、グラフの特徴が線形関数よりも多層パーセプトロンを用いた方がより抽出しやすいからであると考えられる。