

進捗報告

1 今週やったこと

- hidden size が違う場合の実験を行いました

2 実験内容

Tarnsformer の入力を「サブ漢字, jsep_i, サブ漢字」のように分けて実験を行いました。そして hidden size が 256, 128, 64 の条件で実験を行いました。

3 実験結果

200 Epoch の実験を行った結果,hidden size が 256 の時 Train Loss は 1.020 に収束し,Valid Loss は 2.224 に収束し, 過学習になります。

そして中身を確認するため, 学習の同時に Train Data と Valid Data でサブ漢字を出力し, Precision, Recall, F1 を計算しました

Train Data について,Precision, Recall, F1 は各自 0.156, 0.183, 0.159 に収束しました。

Valid Data について,Precision, Recall, F1 は各自 0.148, 0.177, 0.147 に収束しました。

hidden size が 128 の時 Train Loss は 1.032 に収束し,Valid Loss は 1.278 に収束しました

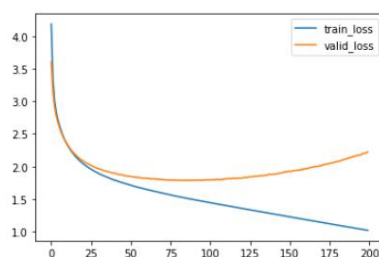
Train Data について,Precision, Recall, F1 は各自 0.405, 0.099, 0.141 に収束しました。

Valid Data について,Precision, Recall, F1 は各自 0.390, 0.091, 0.127 に収束しました。

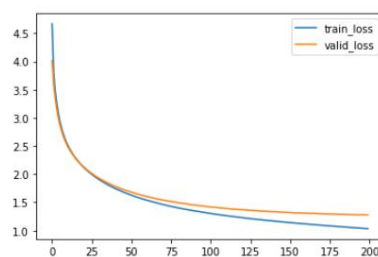
hidden size が 64 の時 Train Loss は 1.062 に収束し,Valid Loss は 1.278 に収束しました

Train Data について,Precision, Recall, F1 は各自 0.177, 0.258, 0.201 に収束しました。

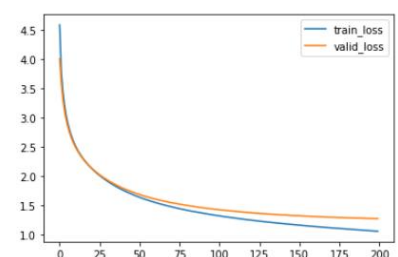
Valid Data について,Precision, Recall, F1 は各自 0.176, 0.255, 0.200 に収束しました。



(a) hid = 256

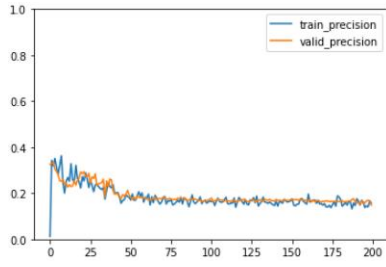


(b) hid = 128

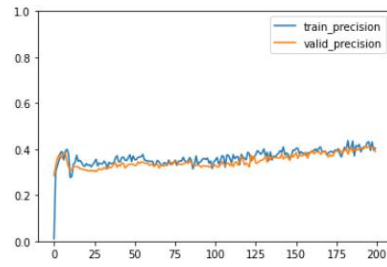


(c) hid = 64

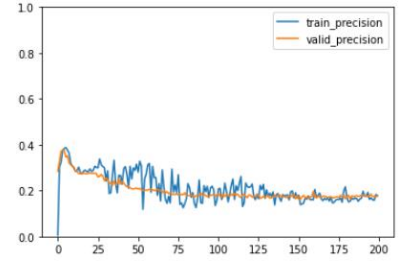
図 1: Loss



(a) hid = 256

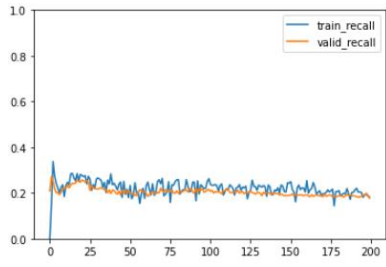


(b) hid = 128

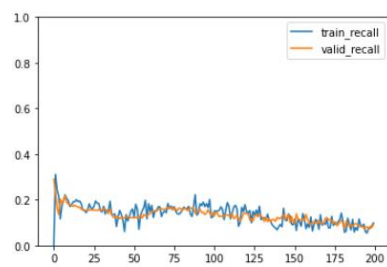


(c) hid = 64

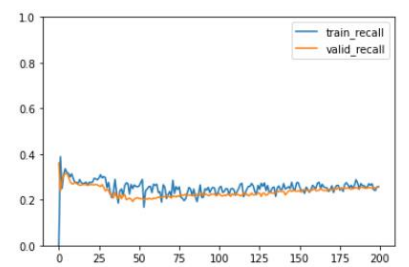
图 2: Precision



(a) hid = 256

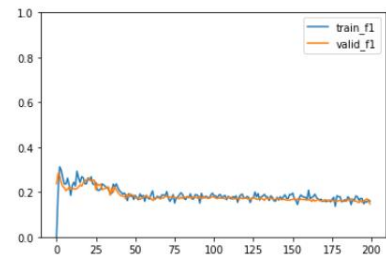


(b) hid = 128

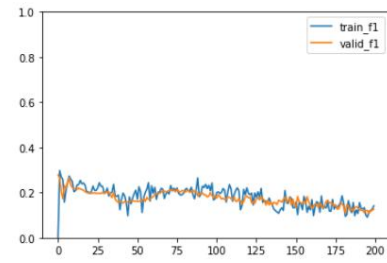


(c) hid = 64

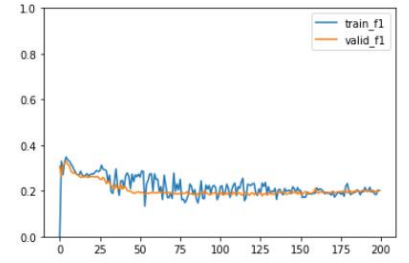
图 3: Recall



(a) hid = 256



(b) hid = 128



(c) hid = 64

图 4: F1

4 来週目標

- epoch 数を増やして実験します
- 他のパラメータを変わって実験します
- 出力が繰り返す原因を探します