
進捗報告

1 今週やったこと

- プログラムの調整

2 データ処理

データの処理は `torchtext` で行いますが、二つの問題が.

まずは問題から作られたボキャブラリサイズより, ラベルから作られたボキャブラリサイズの方が大きいです.

もう一つはバリデーションデータとテストデータに未知の漢字がある場合は訓練できません (key error が出ます).

解決法として, ボキャブラリの構成は訓練データだけでなく, バリデーションデータとテストデータも含まれます.

3 モデル

モデルは RNN です. これから色々なモデルで試したいと思います.

4 来週目標

- プログラムを完成すること