
進捗報告

1 今週やったこと

- biRNN に関する取り組む

2 biRNN

機械翻訳の用いる seq2seq モデルは, エンコーダーが入力のインフォメーションを抽出して, 固定的なベクトルを出力するので, 長い入力は扱いにくい, データセットより長いデータは特にそうである.

よって, エンコーダーに biRNN を導入する.

3 biRNN の実装

biRNN は pytorch の bidirectional パラメータで調整することで導入できるが, 入力 tensor の形ははっきりしてないので, プログラムはまだ出来ていない.

4 他にある問題

日本語データを処理する janome ライブラリは, 時間かかるので, 次は MeCab で実験する.

5 来週目標

- biRNN を実装すること