
進捗報告

1 今週までにやったこと

- ライブドアコーパスから Word2Vec の学習モデルの獲得
- 論文のソースコードから Word2Box の学習モデルの獲得

2 今週やったこと

2.1 Word2Box の利用

学習した Word2Box のモデルから単語の「箱」表現の取得と単語間の類似度の取得する方法を確認することで学習モデルと各種クラス・関数の挙動の理解を深めた。Word2Box の類似度を表す数値の意味については理解が及ばないところがあることがわかった。

2.2 Word2Box の類似度の書き出し

2 つの単語の意味の合成から得られる単語を獲得する準備として学習した Word2Box のモデルでのすべての組み合わせにおいて単語間の類似度を求め、ファイルへと書き出しを行った。

3 次やること

- 学習した Word2Box と条件を揃えた Word2Vec の学習モデルの獲得
- simirality_dataset に得られた類似度が準じているかの確認
- 2 つの単語の意味の合成から得られる単語を獲得
- データセットを日本語に置き換えたモデル獲得の検討
- Word2Box での意味の集合論的關係性確認方法の検討 (論文を読む)