

パターン認識 レポート課題

井上智裕

2021 年 1 月 16 日

1 選んだ論文

Alex Tamkin, Dan Jurafsky, and Noah Goodman. Language through a prism: A spectral approach for multiscale language representations, NeurIPS, 2020, <https://arxiv.org/abs/2011.04823>

2 導入

2.1 扱う問題

自然言語処理分野。特定の言語モデルに頼らない、自然言語におけるスケールの異なる構造（語彙、節、文書など）の発見・学習。

2.2 問題意識

単語の意味など語彙レベルの構造、節や文レベルでの構造、文書全体の主題構造や物語構造など、言語は異なるレベルでの構造を有している。従来の研究では、単語の分散ベクトル表現の学習 [1] や文の分散表現の学習 [2] といった個々のレベルでの構造を捉えるためのモデルの構築など、異なる階層の言語構造のそれぞれを明示的にモデル化する方法が示されてきた。この論文では、従来手法のように文や節といった特定の構造レベルの言語モデルに頼ることなく、あらゆるスケールの表現において、各レベルでの構造を発見・学習するための方法を提案している。

3 理論

言語における異なるレベルの構造の発見にあたり、信号処理分野などで広く使われているスペクトル分析の手法を導入する。入力系列に対し周波数領域で演算を行うためには、入力表現を周波数領域での表現に変換する必要がある。このスペクトル変換には、離散コサイン変換（DCT）を使用する。実数列 $\{x^{(0)}, \dots, x^{(N-1)}\}$ に対して、DCT（各周波数の重み）は次式で得られる。

$$f^{(k)} = \sum_{n=0}^{N-1} x^{(n)} \cos\left[\frac{\pi}{N}\left(n + \frac{1}{2}\right)k\right] \quad k = 0, \dots, N-1 \quad (3.1)$$

この DCT を言語表現に適用するにあたり、文脈に応じた単語表現を対象とする。文脈に応じた単語表現とは、トークン（単語やサブワード単位）の入力系列に対し、自然言語処理モデルが処理することで生成されるベクトル系列である。これは最近の自然言語処理モデルに共通する特徴量であり、前述の自然言語処理モデルには BERT[3] や GPT-2[4] のような Transformer[5] ベースのモデルや ELMo[6] のような LSTM[7] ベースのモデルを含む。文脈に応じた単語表現 v_0, \dots, v_{N-1} が与えられたとき、単語表現のあるニューロン（次元） i に沿った表現 $v_0[i], \dots, v_{N-1}[i]$ に DCT を適用し、得られる系列 $f_0[i], \dots, f_{N-1}[i]$ を第 i ニューロンのスペクトルとする。

この手法では、ニューロンのスペクトルに対して、特定の閾値を持つスペクトルフィルタを通すことによる特定成分の除去と逆離散コサイン変換 (IDCT) による元のドメインへの変換を行うことにより、系列から特定のスケールの構造を取り出すことができる。この研究では、入力サイズを 512 として、スケールに応じて表 3.1 に示す 5 種類のフィルタを用意した。

表 3.1 スペクトルフィルタ

フィルタ	対応する言語構造	周期 [tokens]	DCT index k
HIGH	単語	1 – 2	130 – 511
MID-HIGH	節	2 – 8	34 – 129
MID	文	8 – 32	9 – 33
MI-LOW	段落	32 – 256	2 – 8
LOW	文書	256 – ∞	0 – 1

4 実験

スペクトルフィルタにより、異なるスケールの言語構造が取り出せるか評価を行うための実験を行う。スペクトルフィルタによりフィルタリングされた表現を用いて異なるスケールのタスクを行った時、スペクトルフィルタの選択が分類器の能力にどのように影響するか比較する。以下の各データセットに対して、 $BERT_{BASE}$ モデル [3] を用いて 768 次元の単語表現を得る。次に、各次元に沿ってスペクトルフィルタを適用し、フィルタリングされた表現を用いて特定タスクを実行するソフトマックス分類器を学習させる。用いたデータセット、タスクは以下の通りである。

1. 文章タグ付け (単語レベル)。Penn Treebank データセットを使用する。タスクは、与えられたトークン表現から品詞 (例: 動詞過去形、wh 代名詞、数詞) を予測することである。
2. 対話の分類 (発話レベル)。Switchboard Dialog Speech Acts コーパスを使用する。このタスクは、与えられたトークン表現を含む発話の対話の分類 (例: 謝罪、言葉を濁す、感謝) を予測することである。
3. トピック分類 (文書レベル)。20 Newsgroups データセットを使用する。与えられたトークン表現を含む文書のトピック (ニュースグループ; 例: SCI.SPACE、COMP.GRAPHICS、REC.AUTOS) を予測する。
4. masked language modeling (MLM) タスク。単語の抜けがある文章の抜けを予測するタスク。BERT の事前学習で利用されるため、スペクトルフィルタ適用前の表現のターゲットタスクである。

各タスクの実行結果を図 4.1 に示す。図 4.1 より、単語レベルのタスクである文章タグ付けでは HIGH バンドを用いたときの精度が最も高い一方、フィルタリング前より性能が下がっている。これは、単語レベルの情報を主に使う一方で低周波数情報も必要としていることを示す。一方、文書レベルのタスクであるトピック分類タスクでは LOW バンドを用いたときの精度が最も高く、元の表現よりも高い性能を示した。これは、元の表現に存在する高い周波数変動がこのタスクにおいては、精度に悪影響を及ぼしている可能性があることを示している。また、発話を対象とした対話の分類タスクでは、MID フィルタを用いたときの精度がもっとも高い性能を示した。また、MLM タスクの結果は文章タグ付けの結果に最も類似しており、MLM タスクが局所的なタスクであることを確認できた。

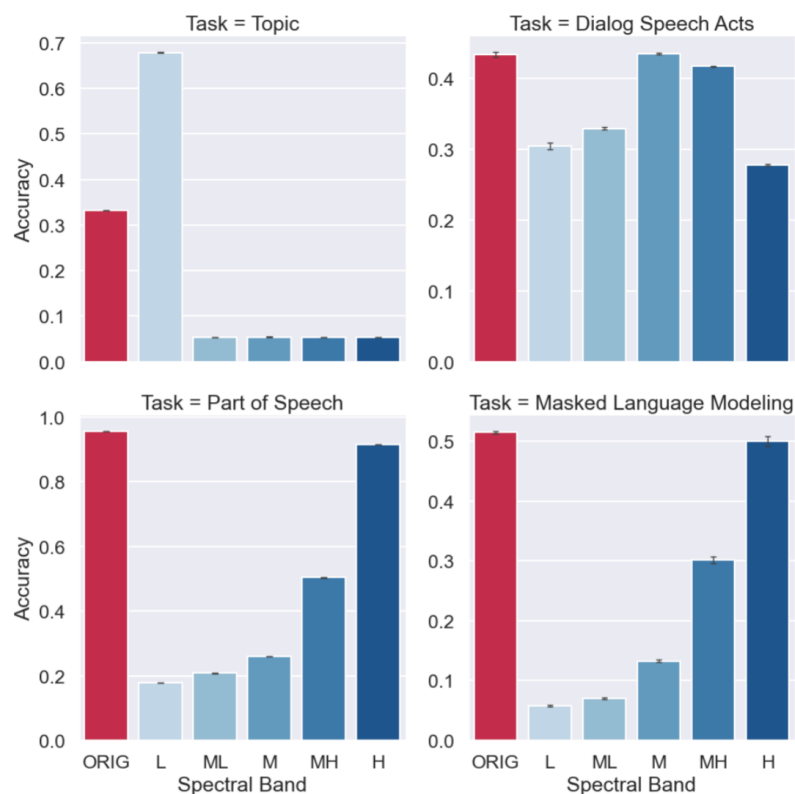


図 4.1 各タスクの実行結果

5 結論

6 論文に対する考察

ネットワーク全体ではなく、ニューロン単位で時系列の周波数情報を見ているのが面白い。

7 結論・展望

他に改善するとしたらどこか？

参考文献

- [1] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg Corrado, and Jeffrey Dean. Distributed representations of words and phrases and their compositionality, 2013.
- [2] Felix Hill, Kyunghyun Cho, and Anna Korhonen. Learning distributed representations of sentences from unlabelled data. In *Proceedings of the 2016 Conference of the North American Chapter of*

- the Association for Computational Linguistics: Human Language Technologies*, pp. 1367–1377, San Diego, California, June 2016. Association for Computational Linguistics.
- [3] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding, 2019.
 - [4] Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, and Ilya Sutskever. Language models are unsupervised multitask learners. 2018.
 - [5] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. Attention is all you need, 2017.
 - [6] Matthew Peters, Mark Neumann, Mohit Iyyer, Matt Gardner, Christopher Clark, Kenton Lee, and Luke Zettlemoyer. Deep contextualized word representations. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*, pp. 2227–2237, New Orleans, Louisiana, June 2018. Association for Computational Linguistics.
 - [7] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, Vol. 9, No. 8, pp. 1735–1780, 1997.