

数理・情報のフロンティア
2020 年度採択研究者

2021 年度 年次報告書

高瀬 翔

東京工業大学 情報理工学院
助教

ランダムベクトルを用いた軽量な埋め込み表現の構築

§ 1. 研究成果の概要

画像情報処理や音声情報処理、自然言語処理においてニューラルネットワークを用いたモデルによる性能向上が目覚ましい。これらニューラルモデルは学習コーパスの量やモデルのパラメータ数、学習に費やした計算時間と性能とが対数比例すると報告されており、モデルのパラメータ数や学習時間が著しく増加している。このような状況を受け、パラメータ数や学習時間のような計算コストに対し、効率の良い手法を考えることが本研究の目的である。本年度は翻訳や要約のような系列変換タスクにおいて学習時間に効率良く性能を向上させられる手法を探求した¹⁾。系列変換タスクでは訓練時に、モデルの出力を確率的に入力として使用することや敵対的な摂動を入力に適用することで性能が向上することが知られている。しかしながら、これらの手法は多量の計算時間を必要としてしまう。本年度の研究において、入力の単語をランダムに別の単語に置換するような単純な手法は上記のような手法と比べると極めて高速にも関わらず同程度の性能を達成でき、計算時間に対して効率良く性能を向上させられることを示した。また、自然言語処理分野で広く用いられているニューラルモデルである Transformer について、高い性能を達成可能な性質を維持しつつ多層にした際の学習を安定化させる手法を研究した。従来手法では学習を安定化させようとすると性能が損なわれており、多層にする利点が少なくなっていたが、これを解消し、パラメータ数に対して効率的な手法を考案した²⁾。

【代表的な原著論文情報】

- 1) Sho Takase, Shun Kiyono. Rethinking Perturbations in Encoder-Decoders for Fast Training. In Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pp. 5767-5780, 2021.
- 2) 高瀬翔, 清野舜, 小林颯介, 鈴木潤. Transformer を多層にする際の勾配消失問題と解決法について. 言語処理学会第 28 回年次大会, 2022.