

表 1 実験結果

	二文節単位	文単位
[BERT]	92.01%	71.53%
[BERT ⁻]	89.48%	65.36%
[RNNLM+SVM] ([5])	85.75%	59.19%
[RNNLM ⁻ +SVM] ([3])	85.84%	58.59%

Linear 層と Sigmoid を介し、入力が読みやすい語順である確率を出力する。なお、入力系列長が BERT の入力上限数を超える場合は、2 つの係り句の系列長の比を維持しつつ、1 文節分の系列長より短くならないように調節した。

4 評価実験

提案手法の有効性を示すために、新聞記事文を用いた語順整序実験を実施した。なお、本研究では新聞記事文は読みやすい語順であるとみなす。

4.1 実験概要

実験データは我々の先行研究 [5] と同一であり、京大コーパス Ver.4.0 のうち、1 月 1 日から 8 日までと 1 月 10 日から 6 月 9 日までの 25,388 文を学習データとし、1 月 9 日と 6 月 10 日から 6 月 30 日までの 2,368 文から、構文情報のみから 1 文全体の語順が確定する文を取り除いたもののうち、1,050 文を開発データ、1,164 文をテストデータとした。モデルは PyTorch^{*1} を用いて実装した。BERT の事前学習モデルには、京都大学が公開しているモデル (BASE WWM 版)^{*2} を用いた。Linear 層 2 層の次元数はそれぞれ 768 と 200 とし、それぞれの入力を 0.1 の確率でドロップアウトさせた。学習アルゴリズムは AdamW を用い、パラメータの更新にはミニバッチ学習 (学習率 1e-6, バッチサイズ 16) により行った。損失関数には BCELoss を使用した。エポック数は、同一パラメータで 5 つのモデルを作成し、開発データを用いてエポックごとに文単位一致率を計測した結果、最良だったエポック 5 とした。

評価指標は二文節単位一致率 (2 つずつ係り句を取り上げ、それらの根の文節間の順序関係が元の文と一致しているものの割合) [2] と文単位一致率 (元の文の語順と完全に一致している文の割合) [5] を採用し、5 つのモデルの平均をそれぞれ測定した。比較のために、提案手法 [BERT] において、BERT への入力として、ボトムアップに語順整序した結果を用いない (具体的には、 d' から取り出した 2 つの係り句の代わりに、 c' から取り出した 2 つの文節を用いる) 手法 [BERT⁻] を用意した。また、我々の先行研究 [RNNLM+SVM] ([5]) と、同モデルでボトムアップに語順整序した結果を用いない高須らの手法 [RNNLM⁻+SVM] ([3]) の実験結果を示す。

4.2 実験結果

表 1 に実験結果を示す。提案手法 [BERT] は、二文節単位と文単位の両一致率において、他の全ての手法を上回り、語順整序において BERT を用いてボトムアップに語順整序することの有効性を確認した。また、[BERT⁻] は、両一致率において提案手法を除く他の手法を上回り、語順整序において BERT を用いることの有効性を確認した。

4.3 提案手法による影響の分析

まず BERT の成功例を分析する。文単位において、提案手法 [BERT] (5 つのうち、最も高い一致率となったモ

*1 <https://pytorch.org/>

*2 https://nlp.ist.i.kyoto-u.ac.jp/?ku_bert_japanese/

提案手法の出力 (正解)

経営悪化に/陥り、/自助努力に/限界が/ある/金融機関には、/国民の/税金への/シロ寄せを/最小に/すべく、/早期の/対策が/欠かせない。

[RNNLM+SVM] の出力 (不正解)

経営悪化に/陥り、/限界が/自助努力に/ある/金融機関には、/最小に/国民の/税金への/シロ寄せを/すべく、/早期の/対策が/欠かせない。

図 2 提案手法の成功例

提案手法の出力 (不正解)

サミット終了後、/村山富市首相は/仏に/立ち寄り、/パリで/シラク大統領就任後/初の/日仏首脳会談に/臨む。

[RNNLM+SVM] の出力 (正解)

村山富市首相は/サミット終了後、/仏に/立ち寄り、/パリで/シラク大統領就任後/初の/日仏首脳会談に/臨む。

図 3 提案手法の失敗例

デル) が正解し、[RNNLM+SVM] では不正解となった文は 231 文存在した。その典型例を図 2 に示す。この例の「限界がある」及び「最小にする」のように、2 文節がひとまとまりで表現される傾向が強い表現に対して、その内部に他の語を割り込ませることなく、提案手法は正しく連続して並べることができている。先行研究では、[RNNLM+SVM] も [SVM 単独手法] と比べて、同様の傾向があると報告していたが [5]、提案手法は [RNNLM+SVM] 以上に、2 文節間の結びつきを考慮できるものと考えられる。次に、BERT の失敗例を分析する。文単位において、[RNNLM+SVM] が正解し、提案手法 [BERT] が不正解となった文は 78 文存在した。その典型例を図 3 に示す。この例では、「サミット終了後、」を「村山富市首相は」よりも前に配置することで [BERT] が誤った。しかし、[BERT] による出力文も正解文と同程度に読みやすいと感じられる。不正解であっても [BERT] による出力文が読みやすいと感じられる文が数多く存在した。

5 おわりに

本稿では BERT を用いて、ボトムアップに日本語文を語順整序する手法を提案した。語順整序実験の結果、BERT モデルの有効性及びボトムアップに語順整序を行うことの有効性を確認した。今後は、人手による主観評価や、サブワードの長さごとに一致率を比較するなどして、評価をより充実させたい。

謝辞 本研究は、一部、科学研究費補助金基盤研究 (C) No. 19K12127 により実施した。

参考文献

- [1] 日本語記述文法研究会, “現代日本語文法 7,” くろしお出版, 2009.
- [2] 内元ら, “コーパスからの語順の学習,” 自然言語処理, Vol.7, No.4, pp.163 – 180, 2000.
- [3] 高須ら, “RNNLM と SVM を用いた日本語文の語順整序,” JPSJ 第 82 回全国大会講演論文集, pp.453 – 454, 2020.
- [4] T. Kuribayashi et al., “Language Models as an Alternative Evaluator of Word Order Hypotheses: A Case Study in Japanese,” Proc. ACL2020, pp.488 – 504, 2020.
- [5] 山添ら, “言語モデルと構文情報を用いた日本語文のボトムアップ語順整序,” FIT 講演論文集, Vol.20, No2, pp.295 – 296, 2021.
- [6] J. Devlin et al., “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding,” Proc. ACL2018, pp.4171 – 4186, 2019.