

# Sentence-BERT を用いたネット小説の文体評価

## Evaluation of Web Novel Style using Sentence-BERT

2022Y22 瀬尾 慎介 (Shinsuke Seo)

担当教員 出口 利憲 (Toshinori Deguchi) ・ 田島 孝治 (Koji Tajima)

### 1. 序論

ネット小説とはネット上に公開されている小説のことである。多くは小説投稿サイトで公開されており、人気を集めて書籍・コミック化する作品も多い。しかし、ネット小説は膨大な作品が存在するため、ランキング上位以外の作品が注目を集めることは難しい。そのため、本文の特徴から検索する機能があれば、埋もれた良作が注目される可能性が高まると思われる。小説の面白さにつながる要素は無数にあるが、本研究では読みやすさなどに繋がる文体要素に注目した。文体を評価することで面白いと感じる作品の文体を知ることができ、文体が優れている作品が注目されると考える。

本研究では、ネット小説を文体で評価して好みの作品を推薦することを目的に、ネット小説の本文から文体を評価するモデルを生成する。このモデルを用いて文体が近い小説を検索し、その結果が妥当かを評価する。

### 2. BERT

BERT[1]とは自然言語処理の学習モデルの一つで、Transformer ベースの機械学習手法である。ネットワークに双方向 Transformer encoder を、事前学習タスクに Masked Language Model (MLM) と Next Sentence Prediction (NSP) を採用しており、前後の単語の関係性を学習してコンテキストを考慮したベクトル表現を生成する。また、大量の生の文章データで事前学習を行ったモデルに対して比較的少ない学習データでファインチューニングすることで、各種言語タスクに特化したモデルを作成することができる。

### 3. Sentence-BERT

Sentence-BERT[2]とは、BERT をより精度の高いモデルにファインチューニングする手法の一つである。BERT から生成される二つの文章の埋め込み表現を比較し、類似度の高い文章同士を近づける

ように学習する (Fig. 1)。BERT より優れた埋め込み表現が出力でき、埋め込み表現の類似度によってデータの関係性を計算することができる。そのため、大量のデータを分類する際の計算時間を大幅に減らせることから多クラス分類問題などに用いられる。本研究で作成するモデルも作品同士の本文の比較に使用するため Sentence-BERT による学習手法を採用した。

### 4. 提案手法

本研究では、「同作品内の文章は文体が類似する」と仮定し、同小説内の文章かどうかを判別するタスクで Sentence-BERT による学習を行うことで、文体の特徴を持つ埋め込み表現を出力する手法を提案する。また、埋め込み表現が小説推薦に活用可能かどうかの検証を行う。

### 5. 実験方法

学習データには小説投稿サイト「小説家になろう」の作品の本文を使用した。「恋愛」「ファンタジー」「文芸」「SF」の4ジャンルからそれぞれ総合ポイントが高い250作品を選び、50話以上の長編作品のみを抽出する。そして各作品の1~50話の書き出しをウェブスクレイピングで取得してモデルへの入力データとする。

学習手法には Triplet Loss を使用する。基準とするサンプル (anchor) と、それに類似したサンプル (positive)、類似していないサンプル (negative) を定義し、それぞれの関係性を学習する。

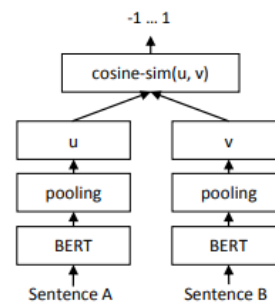


Fig.1 SBERT architecture at inference.

実験では anchor、positive を同作品の異なる話の書き出し、negative をランダムな異作品の書き出しとしたデータを使用する。同作品の1~50話の書き出しの全組み合わせである 2450 通りのデータに negative を追加して、合計 1080450 個のデータを作成した。このデータを学習用、評価用、テスト用で 8 : 1 : 1 の割合に分割し、864360 個のデータを学習に使用した。

学習に使用する BERT モデルは東北大学研究チームが公開している日本語事前学習モデルを使用する[3]。このモデルに Mean Pooling を結合したモデルをファインチューニングする。

モデルの評価は他手法との比較で行う。提案手法との比較対象としてファインチューニング前の BERT モデルによる手法、Word2vec で単語ベクトルの平均をとる手法を用いて、テスト用データでの正解率を求めて比較する。また、positive 類似度と negative 類似度の差分の平均を求め、分類の明確さを評価する。

## 6. 実験結果

1000 個のテスト用データによる評価を Table. 1 に示す。提案手法の正解率が 99.4% と、他手法と比較して高い精度であることが分かる。また、positive 類似度と negative 類似度の差分が大きく取れている。このことから、提案手法は作品内の特徴を他手法より明確に捉えていることが分かる。

しかし、この実験データでは捉えている特徴が文体かどうか判断できない。そのため、正しい文体の評価かどうかを評価するために、同作者の作品は文体が類似すると仮定して、同作者の作品かどうかを判断可能か検証する。検証では、ランダムに選出した 5 人の作者のそれぞれ 5 つの作品から一話書き出しをウェブスクレイピングで取得し、同作者の 5 作品と、各作者から 1 作品ずつ選択した異作者の 5 作品を評価する。5 作品の全組み合わせである 20 組の類似度の平均を 5 作品の類似度として、同作者と異作者の類似度を比較する。今回は 5 人の作者で検証したため、同作者と異作者のデータを①~⑤の 5 組ずつ作成し、それぞれで類似度の計算を行った。

結果を Table. 2 に示す。異作者のデータの類似度が約 0.4~0.5 に対して同作者は約 0.5~0.7 と比較的高いことが分かる。このことから、提案手法のモデルは作者の違いもある程度判別すること

Table.1 Evaluation on the test set.

モデル	正解率	差分(平均)
提案手法	0.994	0.361
BERT(東北大)	0.781	0.024
Word2vec(平均)	0.746	0.013

Table.2 Evaluation on the set by the same authors and different authors.

データ	類似度(平均)	データ	類似度(平均)
同作者①	0.686	異作者①	0.480
同作者②	0.696	異作者②	0.477
同作者③	0.674	異作者③	0.522
同作者④	0.558	異作者④	0.502
同作者⑤	0.603	異作者⑤	0.434

ができることが分かる。しかし、5 人の作者のみでの結果であるため、信憑性を高めるために追加の検証が必要だと考えられる。

## 7. まとめ

今回の実験では提案手法のモデルを作成し、テストデータで高い精度が確認できた。また、同作者と異作者の作品を比較すると、同作者の作品のほうが類似度が高くなることが確認できた。

課題点として、このモデルの文体評価の精度、有用性が確認できていない点がある。作品や小説が同一かどうかを判別するだけでは文体で判別しているかどうかは定かではない他、他作品推薦での有用性も証明できない。そのため、読者のブックマークを用いた検証やランキング上位の作品の特徴解析等を行うことで、モデルの文体評価の妥当性や有用性を検証したい。また、学習データを同作者の作品や同ブックマーク内などに広げること、更に精度の良いモデルの作成も検証したいと考える。

## 参考文献

- [1] Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. arXiv:1810.04805. 2018
- [2] Nils Reimers, Iryna Gurevych. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. arXiv:1908.10084. 2019
- [3] Tohoku NLP Group. “cl-tohoku/bert-base-japanese-whole-word-masking”. Hugging Face. 2021. <https://huggingface.co/cl-tohoku/bert-base-japanese-whole-word-masking>