

BERT とその拡張モデルを用いた作詞家推定手法と分析

笠松雅史¹ 青野雅樹¹¹ 豊橋技術科学大学 情報・知能工学専攻

kasamatsu.masashi.wx@tut.jp masaki.aono.ss@tut.jp

概要

深層学習を用いた自然言語処理の中ではまだあまり着目されていない歌詞を対象に、従来手法やBERTとその拡張モデルを用いて歌詞の作詞家推定を行った。また、各推定モデルを用いて、推定モデルが誤分類した歌詞に着目した、間違えやすい歌詞とクラスの基準を提案した。そして、提案した基準のもと歌詞データの分析を行い、作詞家の特徴の推察を行った。

1 はじめに

自然言語処理においても深層学習が浸透してきている。その中でGoogleが2018年にBERT (Bidirectional Encoder Representations from Transformers) [1] という自然言語処理モデルを発表した。BERTは、大量のテキストデータから学習を行い、その学習で得られた特徴量を様々なタスク (文章理解や感情分析など) に応用することができるという汎用性の高さが特徴として挙げられる。ここで、BERTの汎用性の高さが歌詞にも適用されるかの興味がある。

また、テキストデータの著者推定タスク [2] は、近年のソーシャルメディアの発達に伴うテキストデータの増加により、注目を集めているタスクである。しかし、題材とされているテキストデータには小説やニュース記事が多く、歌詞は比較的少ない。

そこで本研究では、歌詞の著者、すなわち作詞家の推定を行う。本研究の目的は歌詞データに対するBERTの汎用性の解明と分析である。目的を達成する手段として、まずBERTを用いた作詞家の推定モデルの獲得、また従来手法の機械学習や深層学習でも推定モデルを獲得、そして得られたモデルで推定精度の比較・分析を行う必要がある。本稿では、これらについて報告を行う。

2 データセット

使用するデータセットは、Uta-Net[3]より取得した作詞家と歌詞がセットになった独自のデータである。データセットでは多くのデータがあることが望ましいため、作詞家には著名で多くの歌詞を作詞した方を対象にしている。したがって、作詞家は11クラスで作詞家1クラスそれぞれ100曲分のデータであり、全部で1,100件のデータセットである。表1に作詞家とクラスの対応表を示す。

表1 作詞家とクラスの対応表

作詞家	対応クラス
ATSUSHI	クラス0
Seiko Matsuda	クラス1
YUMI MATSUTOYA	クラス2
つんく	クラス3
前山田健一	クラス4
小室哲哉	クラス5
小竹正人	クラス6
松本隆	クラス7
秋元康	クラス8
阿久悠	クラス9
阿木耀子	クラス10

3 作詞家推定モデル

ここでは、11クラスの作詞家の分類を行う推定モデルについて述べる。

3.1 従来手法による作詞家推定モデル

自然言語処理において、BERTが登場する以前よりよく用いられる従来手法として、次の3種類を用いる。

- SVM (Support Vector Machine)
- MLP (Multi Layer Perceptron)
- LSTM (Long Short Term Memory)

まず SVM では、入力特徴量に TF-IDF (Term Frequency-Inverse Document Frequency) で重み付けされた BoW (Bag of Words) を用いる。入力特徴量の次元数は 9,179 次元である。

次に MLP では、入力特徴量に SVM と同様の TF-IDF で重み付けされた BoW を用いる。MLP の中間層は 2 層で、中間層のユニット数は入力から順番に 1024, 256 である。また活性化関数は全て ReLU 関数である。

そして LSTM では、1 単語 200 次元のランダムベクトルを 1 曲分、時系列データとして入力する。LSTM の各内部表現の次元数は 128 で、最終層の内部表現から 11 クラス分類を行う。

3.2 BERT を用いた作詞家推定モデル

使用する BERT は、東北大学の乾研究室が公開している訓練済みの BERT-base である [4]。BERT を用いた作詞家推定のモデル図を図 1 に示す。[CLS] トークンに対応する最終章の埋め込みベクトルである [CLS] ベクトルを全結合層を通して作詞家推定を行う。

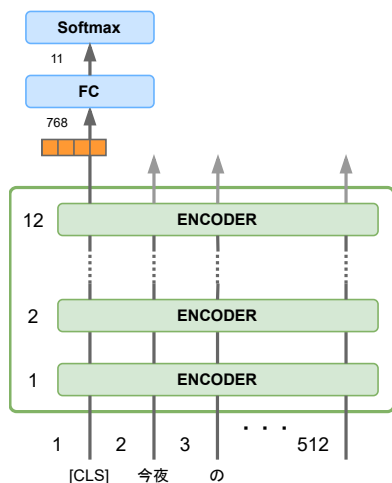


図 1 BERT を用いたモデル図

3.3 BERT を拡張した作詞家推定モデル

図 1 に示すような作詞家推定モデルでは、BERT の ENCODER のうち最終層のみの [CLS] ベクトルから作詞家推定を行う。しかし BERT を拡張した作詞家推定モデルでは、中間層の ENCODER より得られる [CLS] ベクトルも利用することを考える。この考えの概要図を図 2 に示す。図 2 のように最大で 12 層分の [CLS] ベクトルを利用する構造 (図 2 中の Classifier) を考える。この考えのもと、表 2 に示す

ように 6 種類の Classifier の構造を持つ作詞家推定モデルを提案する。

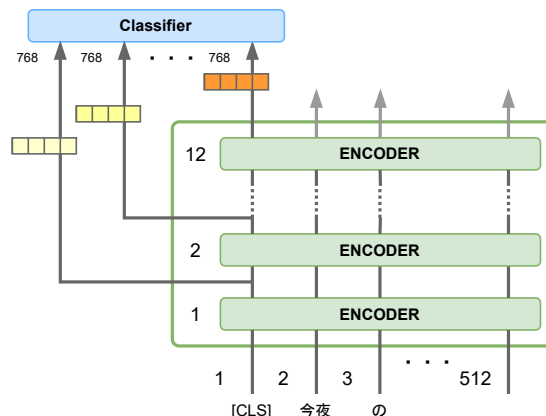


図 2 BERT を用いた拡張モデルの概要図

表 2 拡張モデル 6 種類の Classifier の構造

拡張モデル	Classifier の構造
拡張モデル 1	1 層 MLP
拡張モデル 2	2 層 MLP
拡張モデル 3	3 層 MLP
拡張モデル 4	LSTM
拡張モデル 5	階段構造な MLP
拡張モデル 6	木構造な MLP

4 実験

この実験では、各作詞家推定モデルの精度調査を行う。

4.1 実験方法

各作詞家推定モデルの精度調査を K-fold 交差検証で行う。交差検証の回数は 10 回として、テスト結果 10 回の平均 Accuracy を記録する。交差検証を行う際のデータ分割は、学習データとテストデータで登場するクラス数が均等となるように分割を行う。すなわち、交差検証ごとに学習データ 990 件 (90 件 × 11 クラス)、テストデータ 110 件 (10 件 × 11 クラス) となるように分割を行う。

4.2 学習の設定

まず SVM の設定は、カーネルにガウスクーネルを用いる。次に MLP の設定は、オプティマイザーに Adam を、学習率は 0.001、バッチサイズは 64、エポック数は 100 である。また LSTM の設定は、MLP の場合と同様である。そして BERT を用いた作詞家推定モデルは、オプティマイザーは Adam であり、

BERT 部分の学習率は 0.00005 でそれ以外の学習率は 0.0001, バッチサイズは 16, エポック数は 30 である。最後に BERT を拡張した作詞家推定モデル 6 種類は, BERT を用いた作詞家推定モデルと同様である。

4.3 実験結果

表 3 に従来手法による作詞家推定方法 3 種類と BERT とその拡張モデルによる作詞家推定方法 7 種類の平均 Accuracy と標準偏差を示す。

モデル	平均 Accuracy [%]	標準偏差
SVM	56.36	0.039
MLP	63.27	0.047
LSTM	33.82	0.036
BERT	78.72	0.039
拡張モデル 1	79.18	0.034
拡張モデル 2	78.55	0.046
拡張モデル 3	77.64	0.042
拡張モデル 4	77.45	0.030
拡張モデル 5	78.45	0.038
拡張モデル 6	79.64	0.040

表 3 より, 従来手法である SVM や MLP, LSTM と比べて BERT の平均 Accuracy が高いことがわかる。このことから歌詞データに対しても BERT は有効であると考えられる。

また BERT と拡張モデルでは, 拡張モデルのうち 2 種類が BERT よりも精度が高いことが確認された。そこで, BERT より精度が向上した拡張モデル 1 と 6 を図 3a と図 3c に, 精度が低下した拡張モデル 4 を図 3b に示す。図 3a, 図 3c, 図 3b を見ると, 精度が向上した拡張モデル 1 と 6 では, Classifier の構造が拡張モデル 4 と比べて単純であることから, 複数の [CLS] ベクトルを扱う際には単純な構造が有効であると考えられる。

5 間違えやすさの分析

ここでは, 実験で得られた 10 種類の作詞家推定モデルを用いてどのクラスの歌詞がどのクラスに間違えられやすいかの分析を行う。

5.1 間違えやすい歌詞とクラスの基準

はじめに, 間違えやすい歌詞の基準を次のように定義する。ある 1 つの歌詞に対して, 10 種類の推定

モデルが半数以上あるクラスに誤分類すると, その歌詞は間違えやすい歌詞とする。そして, あるクラスの歌詞がまたあるクラスへ間違えやすい歌詞が 3 曲以上あれば, そのクラスへ間違えやすいクラスとする。

5.2 分析方法

上記で定義した間違えやすい歌詞とクラスの基準をもとにして, 作詞家 11 クラス全てに対して間違えやすい歌詞とクラスを算出する。次に, 間違えやすいクラスの関係図を辺の重みを間違えやすい歌詞の数として作成する。

また, 作成した関係図から PageRank を算出する。PageRank とは, 1998 年に S.Brin と L.Page によって発表され, 参照リンクが多い Web ページは重要度が高くなるという考えのもと, Web ページの重要度を測るアルゴリズムである [5]。ここでは, 間違えやすい歌詞が多いクラスはそれらの歌詞に重要な特徴があると考えのもと, PageRank による分析を行う。

そして PageRank の高いクラスに間違われている歌詞に対して, ユーザーローカル提供のテキストマイニングツール [6] を用いてワードクラウドを作成し, 間違われやすい歌詞にどのような特徴があるか分析を行う。

5.3 分析結果

図 5 に作成した間違えやすい歌詞とクラスの関係図と PageRank を示す。図 5 において, 丸の中の数字がクラスを, 矢印に重なっている数字が間違えやすい歌詞の数を, 赤字がそのクラスの PageRank をそれぞれ現している。図 5 から, PageRank が高いクラスはクラス 2, クラス 4, クラス 5, クラス 10 などであることがわかる。

次に, PageRank の高いクラスに間違われている歌詞に対するワードクラウドを図 4a, 図 4b, 図 4c にそれぞれ示す。図 4a は 17 曲分の, 図 4b は 12 曲分の, 図 4c は 13 曲分の歌詞のワードクラウドである。また, 赤字が動詞を, 緑字が形容詞を, 青字が名詞をそれぞれ表している。

図 4a から, クラス 2 に間違われやすい歌詞には英単語やネガティブな形容詞がよく使用されると推測できる。次に図 4b からは, クラス 5 に間違われやすい歌詞にはクラス 2 と同様英単語が多く, またネガティブな形容詞が多いが, 動詞にはポジティブなものが多い特徴があると推測できる。そ

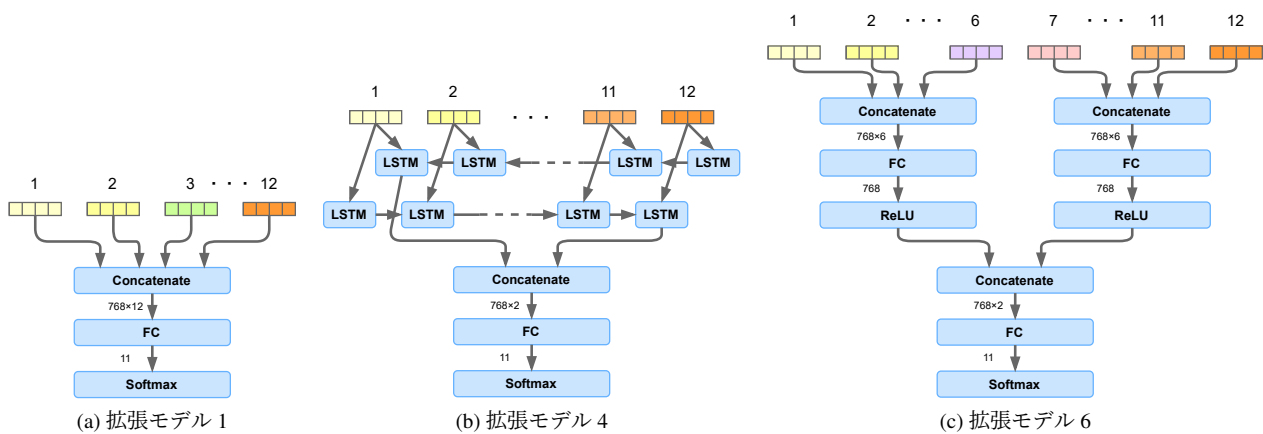


図3 精度が向上した拡張モデルと低下した拡張モデル



図4 間違われた歌詞のワードクラウド

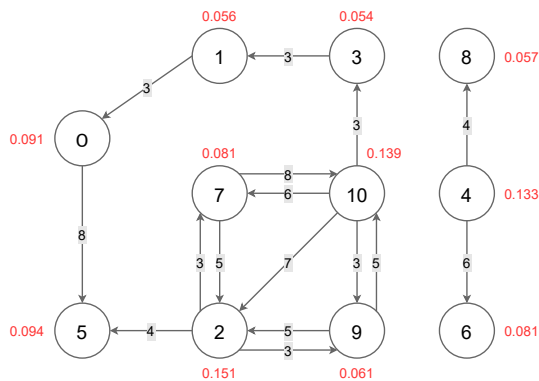


図5 間違えやすい歌詞のクラスの関係図と PageRank

して図 4c からは、クラス 10 に間違われやすい歌詞には、ネガティブな動詞や形容詞がよく使用されていると推測できる。

6 おわりに

本稿では、歌詞データに対しても BERT が有効であるかを K-fold 交差検証を用いて作詞家推定を行い、従来手法と比較を行った。比較の結果、BERT はどの従来手法よりも優れた結果であった。

また、BERT の拡張モデルを 6 種類考案し、BERT と比較を行った。比較の結果、拡張する際は単純な構造が有効であると考えられる。

そして、従来手法と BERT と BERT の拡張モデルの計 10 種類を用いて、間違えやすい歌詞とクラスの基準の提案を行った。提案した基準のもと簡単な分析を行い、間違えやすい歌詞の特徴の推察を行った。このような推定モデルの誤分類を着目した方法は、データ分析の新たなアプローチの 1 つになると考えられる。

今後の検討として、間違えやすい歌詞とクラスの詳細な分析を進める必要があると考えられる。また、間違えやすい情報を利用して、例えば GAN (Generative Adversarial Network) にその情報を付加してテキスト生成を行いたいと考えている。

参考文献

- [1] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. BERT: pre-training of deep bidirectional transformers for language understanding. **CoRR**, Vol. abs/1810.04805, , 2018.
- [2] Janek Bevendorff, Bilal Ghanem, Anastasia Giachanou, Mike Kestemont, Enrique Manjavacas, Iliia Markov, Maximilian Mayerl, Martin Potthast, Francisco Rangel Pardo, Paolo Rosso, Guenther Specht, Efstathios Stamatatos, Benno Stein, Matti Wiegmann, and Eva Zangerle. **Overview of PAN 2020: Authorship Verification, Celebrity Profiling, Profiling Fake News Spreaders on Twitter, and Style Change Detection**, pp. 372–383. 09 2020.
- [3] 歌詞検索サービス 歌ネット. <https://www.uta-net.com/>.
- [4] Github - cl-tohoku/bert-japanese: Bert models for japanese text. <https://github.com/cl-tohoku/bert-japanese>.
- [5] S. Brin and L. Page. The anatomy of a large-scale hypertextual web search engine. In **Seventh International World-Wide Web Conference (WWW 1998)**, 1998.
- [6] Ai テキストマイニング by ユーザーローカル. <https://textmining.userlocal.jp/>.