

Title	BERTを用いた歌詞表現の特徴分析
Sub Title	
Author	佐野, 貴紀(Sano, Takanori)
Publisher	慶應義塾大学AI・高度プログラミングコンソーシアム
Publication year	2023
Jtitle	AICカンファレンス予稿集 (2023.) ,p.41- 41
JaLC DOI	
Abstract	近年, 自然言語処理手法としてBERTモデルが注目を集めている。BERTは, Attention機構を用いることで, モデルが注意を当てている単語や文章の定量化が可能である。このモデルは翻訳や文章分類など様々な場面での社会実装がなされている強力な手法である。本研究では, BERTモデルとAttention機構を用いて歌詞表現の特徴抽出を試みた。その結果, どのような歌詞表現がアーティストの特色を示しているかの定量的な確認ができた。
Notes	会議名: AICカンファレンス2023 開催地: 慶應義塾大学日吉キャンパス 日時: 2023年3月4日 第2章ポスター発表要旨 ポスター要旨-13
Genre	Conference Paper
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=KO11003001-20230304-0041

慶應義塾大学学術情報リポジトリ(KOARA)に掲載されているコンテンツの著作権は、それぞれの著作者、学会または出版社/発行者に帰属し、その権利は著作権法によって保護されています。引用にあたっては、著作権法を遵守してご利用ください。

The copyrights of content available on the Keio Associated Repository of Academic resources (KOARA) belong to the respective authors, academic societies, or publishers/issuers, and these rights are protected by the Japanese Copyright Act. When quoting the content, please follow the Japanese copyright act.

BERTを用いた歌詞表現の特徴分析

佐野貴紀¹¹慶應義塾大学大学院社会学研究科心理学専攻

Abstract:

近年、自然言語処理手法としてBERTモデルが注目を集めている。BERTは、Attention機構を用いることで、モデルが注意を当てている単語や文章の定量化が可能である。このモデルは翻訳や文章分類など様々な場面での社会実装がなされている強力な手法である。本研究では、BERTモデルとAttention機構を用いて歌詞表現の特徴抽出を試みた。その結果、どのような歌詞表現がアーティストの特色を示しているかの定量的な確認ができた。

Keywords: BERT, Attention, Lyric

1. 研究背景・目的

近年、自然言語処理手法としてBERTモデル[1]が注目を集めている。BERTは、Attention機構[2]を用いることで、モデルが注意を当てている単語や文章の定量化が可能である。このモデルは翻訳や文章分類など様々な場面での社会実装がなされている強力な手法であり、近年では楽曲検索や楽曲推薦などへの応用に向けた研究も進んでいる[3, 4]。これらの検索や推薦アルゴリズムを検討する上で、楽曲の特徴や、様々なアーティスト間の関係性を定量的に把握する技術は応用の一助となる。そこで、本研究では、BERTモデルを用いて様々な楽曲の歌詞表現の特徴の抽出を行い、楽曲の特色の定量把握においてBERTモデルが有用であるかを確認することを目的とした。

2. 方法

はじめに、分析に使用する歌詞データを収集した。本研究では、音楽ストーリーミングサービスであるSpotifyで2022年に国内で再生数の高かった上位10のアーティスト(BTS, Official髭男dism, YOASOBI, Ado, 優里, Vaundy, back number, King Gnu, TWICE, Saucy Dog)の楽曲の歌詞を収集した。これらの一つの楽曲の歌詞情報を学習データ、アーティスト名をラベルデータとして、分類課題として分析を行った。モデルは、東北大学 乾・鈴木研究室が公開している事前学習済みの日本語BERTモデルを用いた。このモデルに、収集した歌詞データを用いて100epochのファインチューニングにより学習をさせ、ホールドアウト法による精度検証を行った。その後、正しく分類されていた楽曲に着目し、各アーティストに特徴的な歌詞表現について確認を行った。

3. 結果・考察

精度検証の結果、正答率は0.653であった。正解していたデータの中から、Attention機構によりモデルが注意を当てている単語を抽出し、可視化を行った。その結果の例をFigure 1, 2, 3に示す。例えば、King Gnuの歌詞の場合は「塗」「遊び」「行く」などが、Vaundyの場合は「僕達」「僕ら」「世界」「生きる」などが、Back Numberの歌詞の場合は「君」「僕」「毎日」などが重要な単語として抽出された。

```

塗["##れ"]["で"]["どこ"]["まで"]["でも"]["飛ん"]["で"]["いけ"]["そう"]["さ"]["UNK"]
["う"]["事"]["で"]["自分"]["自身"]["を"]["失っ"]["て"]["しまっ"]["て"]["も"]["夜"]["に"]["粉
れ"]["て"]["あなた"]["の"]["元"]["へ"]["遊び"]["に"]["行<"]["よ"]["眠"]["##れ"]["ない"]["こ
の"]["一"]["街"]["の"]["無意"]["##味"]["な"]["空"]["騒ぎ"]["に"]["は"]["うん"]["##ざ"]["##
り"]["さ"]["トー"]["##キ"]["##ヨー"]

```

Fig. 1 King Gnu 「Tokyo Rendez-Vous」の歌詞表現の特徴抽出結果の一部抜粋

```

["の"]["中"]["を"]["僕"]["##達"]["は"]["手"]["を"]["取り"]["##合っ"]["て"]["いる"]["な"]
["##あ"]["、"]["なんて"]["美しい"]["世界"]["だ"]["僕"]["##ら"]["何"]["度"]["裏切り"]["あ
っ"]["て"]["い"]["て"]["も"]["ま"]["##あ"]["、"]["なんとか"]["手"]["を"]["取り"]["##合う"]
["ん"]["だ"]["まるで"]["恋愛"]["映画"]["の"]["ラスト"]["##シーン"]["の"]["よう"]["な"]["「"]
["愛"]["で"][""]["UNK"]["世界"]["の"]["中"]["で"]["僕"]["##達"]["は"]["キス"]["を"]["し"]
["あっ"]["て"]["生き"]["て"]["いる"]["UNK"]["世界"]["の"]["中"]["を"]["僕"]["##達"]["は"]
["目"]["を"]["合わせ"]["あっ"]["て"]["生きる"]["w"]["##el"]["##co"]["##me"]["to"]

```

Fig. 2 Vaundy 「不可抗力」の歌詞表現の特徴抽出結果の一部抜粋

```

["な"]["##の"]["##に"]["君"]["は"]["何"]["度"]["も"]["何"]["度"]["も"]["僕"]["の"]["名前"]
["を"]["今"]["君"]["の"]["手"]["を"]["握っ"]["て"]["出来る"]["だけ"]["目"]["を"]["見"]["て"]["こ
んな"]["毎日"]["で"]["良かった"]["たら"]["残り"]["の"]["全部"]["まとめ"]["て"]["君"]["に"]["全
部"]["あげる"]["から"]["きっ"]["##と"]["僕"]["に"]["も"]["あっ"]["た"]["僕"]["##し"]["さ"]
["は"]["何"]["##より"]["こんな"]["気持ち"]["は"]["大"]["##袂"]["##袂"]["じゃ"]["なく"]
["て"]["君"]["と"]["出会っ"]["て"]["知っ"]["た"]["ん"]["だ"]["よ"]["だから"]["もう"]["僕"]
["は"]["君"]["の"]["これから"]["ずっと"]["僕"]["の"]["全"]["て"]["は"]["君"]["の"]["もの"]["だ"]

```

Fig. 3 Back Number 「僕の名前を」の歌詞表現の特徴抽出結果の一部抜粋

BERTモデルで分類に重要な単語を抽出した結果、アーティストごとに異なる単語が重要となる傾向が見られた。今後はより大規模なデータでの解析が必要となるが、BERTモデルはアーティストの歌詞表現の特色を定量的に把握するうえで有用であることが示唆された。

参考文献

- [1] J. Devlin *et al.*, arXiv, 1810.04805, (2018).
- [2] A. Vaswani *et al.*, Advances in neural information processing systems, **30**, (2017).
- [3] 仲村哲明, 内海彰, 坂本真樹., 人工知能学会論文誌, **27**, 163-175, (2012).
- [4] K. Tsukuda *et al.*, ISMIR, 544-551, (2017).