

日本語と英語の歌詞における性差のテキスト分類を用いた検討

木田菜月 久野雅樹

電気通信大学大学院情報理工学研究科

k2230049@edu.cc.uec.ac.jp, hisano@uec.ac.jp

概要

本研究では、歌手の性差による歌詞の違いに注目し、その違いを明らかにするためにニューラルネットワークモデルで日本語の歌詞を学習、歌詞から歌手の性別予測を行うテキスト分類を行い、その性能から性別による差異の実態を明らかにした。また、性別予測の際に算出される予測値を観察することで、性差による特徴や歌手ごとの違いが明らかになった。さらに、英語詞の楽曲でも同様に性差によるテキスト分類を行い、日本語歌詞での結果と比較すると違いが見られた。

1 はじめに

コーパスを対象とした特徴分析、テキストの分類や生成の研究は、言語資源の整備と自然言語処理技法の発展にともない盛んに行われている。本研究では音楽の歌詞を対象に研究を行う。音楽における歌詞には、言語圏ごとの言葉の使われ方、歌手やジャンルの違いによる構造の違いが反映されていると考えられるからである。

歌詞研究の代表的なものとして伊藤の研究が挙げられる。伊藤は松任谷由実をはじめとする J ポップ歌手の楽曲を対象に、対象語彙表や構造語彙表を用いた語彙分析やテーマ分析を行い、男女でのテーマや語彙の違いを明らかにしている[1]。また、社会心理学の分野における古典的な歌詞研究として見田の研究が挙げられる。見田はモチーフ分析や歌の主題の分析、用語分析などを行っており、モチーフ分析では『日本歌謡集』巻末の流行歌年表記載の楽曲を対象に、モチーフの比率の年代変化やその歌詞に注目しながら時代背景と照らし合わせて考察している[2]。このようにこれまでも歌詞研究は様々に行われているが、流行歌や数名の歌手の楽曲など特定の楽曲のみを対象としたものがほとんどであった。そこで、木田の研究ではより全般的な特徴を明らかにす

ることを目的として、大規模な歌詞コーパスを対象に、性差によるテキスト分類を行っている[3]。しかし、この研究のテキスト分類では、コーパスに性別情報がないために性別情報を付加したデータを確保することが難しく、学習データとテストデータに同じ歌手の楽曲を用いている。そのため、性差分類だけでなく歌手分類になっている部分があるという問題点があった。

そこで、本研究では、歌詞コンテンツデータ集[4]を用いた全般的な分析に加え、新たに取得した歌手の性別情報と組み合わせてデータセットを作成することで、学習データとテストデータで異なる歌手の楽曲を用いることができ、性別分類と歌手分類が混ざってしまう問題を解消した。また、歌詞コンテンツデータ集に含まれる英語の歌詞にも着目し、日本語の歌詞と同様に性差によるテキスト分類を行い、結果を比較した。

2 日本語歌詞のテキスト分類

2.1 使用したデータセット

分析対象としてシンクパワー社が提供する歌詞コンテンツデータ集[4]を利用した。この歌詞コンテンツデータ集には、約 43 万曲分の曲名、アーティスト名、アルバム名、作詞者、作曲者、発売日、歌詞などの情報が収録されている。この実験では日本語歌詞を対象とするため、このコーパスから歌詞に少なくとも平仮名またはカタカナが含まれる楽曲のみを抽出した。この抽出した日本語歌詞の楽曲の中から、人物辞典オンライン[5]の男性歌手、女性歌手それぞれの項目に載っている人物およびその人物が所属するグループの楽曲 63,722 曲を対象に分析を行った。これらの楽曲を男女である程度曲数を揃えたのち、それぞれで歌手の組数と曲数がある程度均等になるように 5 つに分割したデータセットを用いた。

表 1 データセットに利用した歌手の数と曲数

	男性	女性	合計
組数	576	876	1,452
曲数	31,846	31,873	63,722

データセットに用いた男性歌手女性歌手それぞれの数と曲数は表 1 の通り。

2.2 実験 1 テキスト分類の性能評価

2.2.1 実験手順

テキスト分類には GRU モデルを用いた。GRU は長期記憶と短期記憶のバランスも学習でき、LSTM よりもゲートが少なくセルも必要ないため、状態変数の数が同じであれば LSTM よりも少ない計算量・使用空間量で済ませることができるためである [6]。

まず、データセットに女性を 0、男性を 1 としてクラス情報を付加し、組数と曲数がある程度均等になるように 5 分割してデータセットを作成し、そのデータセットのうちの 1 つをテストデータ、その他 4 つを学習データとした。これらのデータセットからボキャブラリを構築、単語の ID 化とパディングを行い GRU に学習させ、モデルの性能評価を行った。この手順をテストデータとするデータセットを入れ替えながら計 5 回行った。データセットに利用した歌手の人数と曲数は表 1 の通り。

2.2.3 結果と考察

表 2 英語詞のテキスト分類の性能評価

	平均値 (5 回)
適合率	0.7857
再現率	0.7386
F1 値	0.7607

性能評価の平均値は表 2 の通り。歌手の性差によるテキスト分類の先行研究での性能評価は全ての指標において 0.90 以上であった[3]。これに対し、今回の実験ではすべての指標において 0.75 付近と少し低くなった。この結果から、テストデータと学習データで同じ歌手を利用していたテキスト分類では、データセットに使用していた歌手の数が少なかったこともあり、性差による分類だけでなく歌手の分類になっている部分もあったと考えられる。しかし、デ

ータセットに利用した歌手の数も曲数も大きく増えているにもかかわらず性能評価が依然高いことから、日本語の歌詞には性差による違いが反映されていると言える。

2.3 実験 2 テキスト分類の予測値分布

2.3.1 実験手順

テキスト分類にはこれまでと同様に GRU モデルを用いた。2.1 の計 63,722 曲のデータセットの中に含まれる曲数の多い歌手男女それぞれ 10 組と実験に用いた 80 年代女性アイドル 10 人は表 3 の通り。この歌手のうち 1 組の楽曲をテストデータ、その他の楽曲を学習データとして、ボキャブラリを構築、単語の ID 化とパディングを行い GRU に学習させた。学習したモデルを用いてテストデータの歌手の性別予測を行った。この手順を表 3 の歌手 30 組それぞれに対して行い、この予測で得られた予測値の分布を男性歌手 10 組、女性歌手 10 組、80 年代女性アイドル 10 人ごとに出した。

2.3.2 予測値

予測結果は、それぞれのクラスに対してすべてのクラスの予測値の合計が 1 となるように予測値を出したのち、予測値の値が一番高いクラスを予測結果としている。本実験は 2 値分類であるため、クラス 1 すなわち男性のクラスの予測値が 0 に近ければ近いほど女性的であり、1 に近ければ近いほど男性的な歌詞と言える。

2.3.3 結果と考察

コーパスに含まれる曲数の多い女性歌手 10 組と男性歌手 10 組それぞれの結果から、性別の特徴が強く出ている歌手と出ていない歌手など歌手ごとの差が顕著に出ていることが分かった。また、すべての女性歌手の中央値が 0.5 以下で 0 付近に寄っており、男性歌手の結果ではすべての歌手の中央値が 0.5 以上で 1 付近に寄っている。このことから、歌手ごとにその程度に差はあるが、歌詞には性別による特徴が反映されていると言える。その中で、特に特徴が顕著であった 80 年代女性アイドルに着目し、80 年代女性アイドル 10 人の予測値の分布を見ると、ほとんどの歌手の楽曲の予測値がより 0 付近に寄っており、この頃の女性アイドルの楽曲には女性歌手としての特徴が特に顕著に表れていたことが分かった。

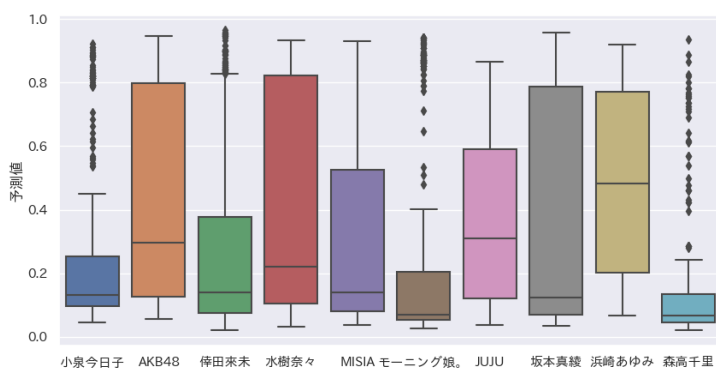


図1 女性歌手10組の予測値 (1に近いほど男性)

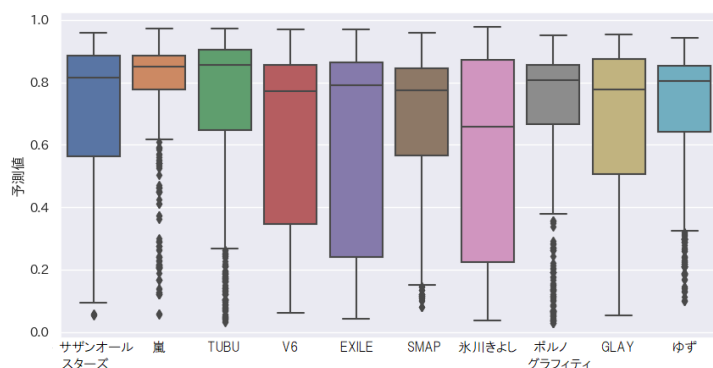


図2 男性歌手10組の予測値 (1に近いほど男性)

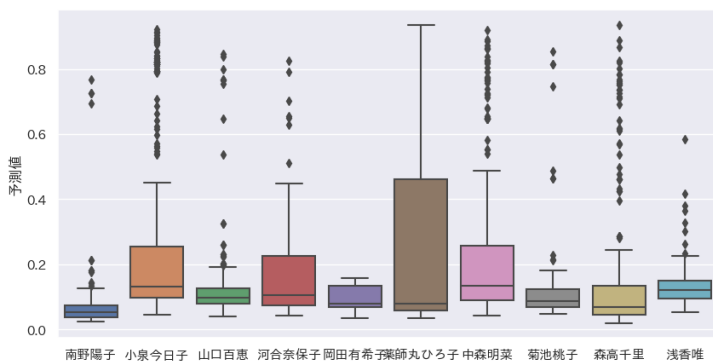


図3 80年代女性アイドル10人の予測値 (1に近いほど男性)

3 英語歌詞のテキスト分類

3.1 使用したデータセット

分析対象として日本語歌詞のテキスト分類と同様にシンクパワー社が提供する歌詞コンテンツデータ集[4]を利用した。この実験では英語の歌詞を対象とするため、言語判定ツールである pyclid2 を用いて言語判定を行った。この言語判定ツールで歌詞全体の英語の割合が 90%以上であると判定され、単語数が

50 語以上 600 語以下の楽曲 116,349 曲を英語歌詞の楽曲とみなして抽出した。このようにして作成した英語の歌詞コーパスの中から曲数の多い英語圏の歌手男女それぞれ 8 組の楽曲を用いてデータセットとした。このときデータセットに利用した歌手 16 組は表 3 の通りである。

表 3 データセットに利用した歌手

男性	女性
Elvis Presley	Britney Spears
Michael Jackson	Mariah Carey
Bon Jovi	Whitney Houston
Queen	Avril Lavigne
Maroon 5	Madonna
Backstreet Boys	Christina Aguilera
Eric Clapton	Alicia Keys
Aerosmith	Taylor Swift

3.2 実験3 テキスト分類の性能評価

3.2.1 実験手順

テキスト分類には日本語歌詞での実験と同様に GRU モデルを用いた。まず、データセットに女性を 0, 男性を 1 としてクラス情報を付加する。一番曲数の少なかった Taylor Swift の曲数 292 曲に合わせてすべての歌手の楽曲をサンプリングし、4,672 曲のデータセットとした。そのうち、男女それぞれ 1 組の楽曲をテストデータ、残りの歌手の楽曲を学習データとして学習と性能評価を行った。この手順をその都度サンプリングを行いながら全ての組み合わせについて計 64 回行い、性能評価の平均値を出した。

3.2.3 結果と考察

表 4 英語歌詞のテキスト分類の性能評価

	平均値 (64 回)
適合率	0.5356
再現率	0.5152
F1 値	0.5219

性能評価の平均値は表 4 の通り。性能評価の値は全ての指標で 0.50 付近の値となった。この結果から英語歌詞の性差による分類は決して簡単なものではないことが分かる。また、日本語歌詞のテキスト分

類の結果と比べてもかなり低くなっていることから、英語歌詞には日本語歌詞ほどはっきりとした性差が見られないと考えられる。

3.3 実験4 テキスト分類の予測値分布

3.3.1 実験手順

テキスト分類にはこれまでと同様に GRU モデルを用いた。実験1と同様に表4の歌手の楽曲をそれぞれ一番曲数の少なかった Taylor Swift の曲数に合わせてサンプリングした。このうち1組の楽曲をテストデータ、その他の歌手の楽曲を学習データとして、ボキャブラリを構築、単語の ID 化とパディングを行い GRU に学習させた。このようにして学習したモデルを用いてテストデータの歌手の性別予測を行った。この手順を表4の歌手16組それぞれに対して行い、この予測で得られた予測値の分布を男性歌手8組、女性歌手8組ごとに出した。

3.3.2 結果と考察

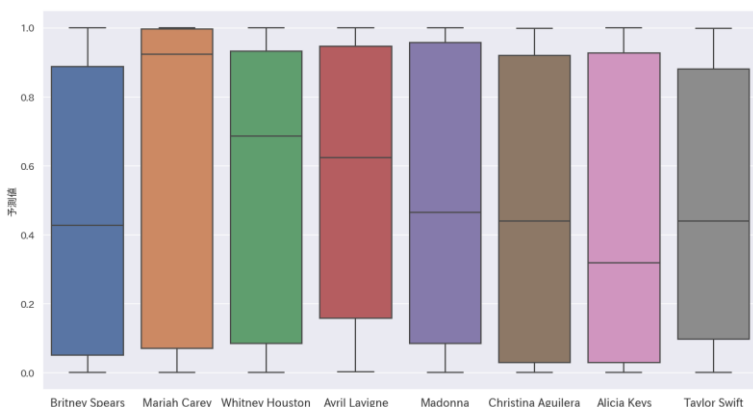


図4 女性歌手8人の予測値 (1に近いほど男性)

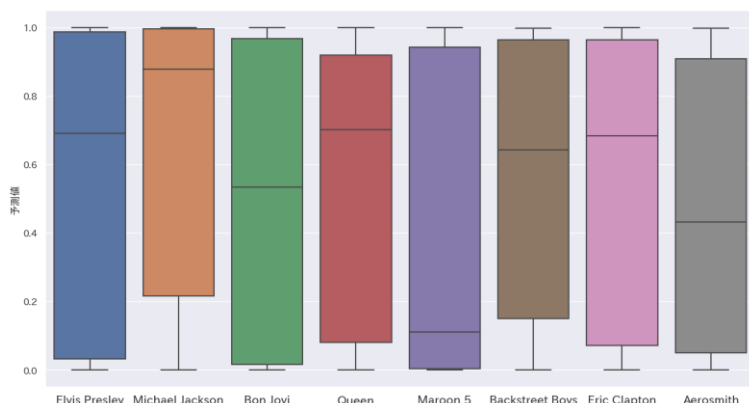


図5 男性歌手8組の予測値 (1に近いほど男性)

男女それぞれの予測値の分布は図4と図5の通り。

この結果をみると、ほとんどの歌手の分布が広がっており散らばりが大きく、日本語歌詞での結果と比べるとより顕著であることが分かる。また、日本語歌詞の結果のように男性歌手は中央値が1付近に寄り、女性歌手は中央値が0付近に寄るというような傾向も見られない。この結果からも、英語歌詞には日本語歌詞ほどはっきりとした性差が見られないと考えられる。

4 おわりに

日本語歌詞のテキスト分類の結果から日本語の歌詞には歌手の性差が反映されていることが示された。また、英語の歌詞においては日本語の歌詞のように男女での顕著な違いがあるわけではないことが示された。さらに、予測値の分布を見ることで歌手ごとに差があることや男性歌手の曲、女性歌手の曲という度合いの強さを観察することができた。

しかし、本研究で用いた英語歌詞のデータセットは、日本語歌詞のデータセットと比べ、歌手の数や曲数が圧倒的に少なく一部の曲や歌手の影響が大きく出ている可能性がある。そのため、コーパスの拡充などが今後の課題として挙げられる。

参考文献

- [1] 伊藤雅光. J ポップの日本語研究—創作型人工知能のために—. 朝倉書店, 2017.
- [2] 見田宗介. 近代日本の心情の歴史 流行歌の社会心理史. 講談社, 1978.
- [3] 木田菜月, 久野雅樹. 日本語歌謡曲の歌詞おける性差と時代変化のテキスト分類を用いた検討. 言語処理学会, 第29回年次大会発表論文集, 2023
- [4] 株式会社シンクパワー. 歌詞コンテンツデータ集, 2020.
- [5] 株式会社ジテンオン. 人物名鑑オンライン. (2023/06/21 閲覧). <https://jimbutsu.jitenon.jp/>.
- [6] K. Cho, B. Merriënboer, C. Gulcehere, F. Bougares, H. Schwenk, and Y. Bengio. Learning phrase representations using rnn encoder-decoder for statistical machine translation. **EMNLP**, pp.1724–1734, 2014.