

## GPT-2 を用いた歌詞生成とその分析 Lyrics generation and analysis using GPT-2

笠松 雅史<sup>1)</sup> 青野 雅樹<sup>1)</sup>  
Masashi Kasamatsu Masaki Aono

### 1 はじめに

AI を用いた自動作詞の研究は、近年 Transformer[1] や GPT-2[2] をはじめに深層学習が自然言語処理を台頭してきたこともあり、興味深い分野である。自動作曲システム Orpheus[3] は作詞のみならず作曲まで自動で行うことができるシステムである。特に作詞に注目すると 4 つのジャンルと 1 行の長さやキーワードを指定することで自動作詞を行うことができる。この作詞は、主に bigram 確率とを用いた確率モデルと単語共起確率を応用して実現している。また、AI 作詞ツール Shikaki[4] は AI を用いた作詞が可能であり、A メロや B メロ等と 1 行の音数を指定することで自動作詞を行うことができる。この作詞は、LSTM と Transformer を組み合わせたモデルを用いて実現している。

上記に挙げた自動作詞システムは歌詞のジャンルや 1 行の音数をコントロールすることはできるが、作詞家の特徴を再現することはできていない。そこで我々は上記の作詞システムに加えて、作詞家の特徴を持った歌詞の自動生成を目的に研究を行っており、その目的を達成するために今まで作詞家の特徴の分析を行ってきた [5]。本稿では GPT-2 を用いて著名な作詞家 11 人の曲を用いて GPT-2 の学習を行う。そして学習したモデルを用いて学習した歌詞の冒頭を入力として歌詞の生成を行い、生成された歌詞と元の歌詞の BLEU スコアを算出し、生成された歌詞の特徴について分析を行う。

### 2 データセット

使用するデータセットは、Uta-Net[6] 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。

### 3 歌詞生成モデルの獲得

歌詞生成モデルには rinna 社が 2021 年 4 月に公開した学習済みの GPT-2[7] をファインチューニングすることで獲得する。ファインチューニングの際は、1 曲分の歌詞を 1 データとして扱い、エポック数は 100 としバッチサイズは 1 とする。ファインチューニングした結果の学習曲線を図 1 に示す。

図 1 の学習曲線からロスが収束していることがわかる。また次単語の予測精度は 0.99 であり、次単語の候補数を表す perplexity は 1.00 であった。

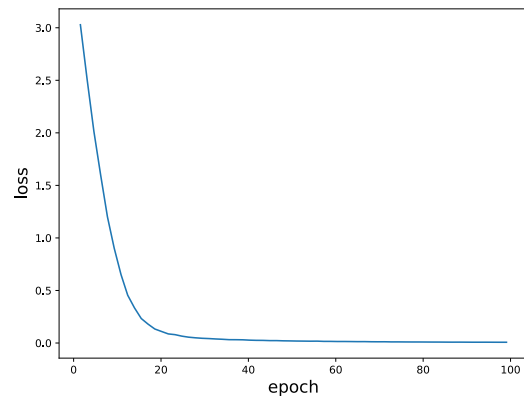


図 1 GPT-2 の学習曲線

### 4 実験

この実験では、ファインチューニングを行った GPT-2 を用いて、学習に使用した歌詞の冒頭を入力として歌詞を生成し、生成された歌詞と入力に用いた歌詞の BLEU を算出し、生成された歌詞の特徴について分析を行う。

#### 4.1 歌詞生成方法と評価方法

歌詞の生成方法は次の通りである。

- 生成に使用する入力文字列は歌詞の冒頭 10% 分の文字列とする。
- 生成する文字列長は入力に使用した歌詞の長さ+10 文字とする。
- 1 つの歌詞につき 3 回生成を行う。

次に生成された歌詞と元の歌詞の評価方法についてまとめる。

- 1 つの歌詞に対して生成された 3 つの歌詞との BLEU を算出する。
- 3 つ分の BLEU の平均を算出し、その平均 BLEU を評価指標とする。

#### 4.2 評価結果

図 2 に作詞家ごとの平均 BLEU のヒストグラムを示す。

図 2 を見ると、どの作詞家も平均 BLEU が高いところと低いところに集中している傾向があることがわかる。実際に平均 BLEU が低い歌詞と高い歌詞を確認してみると、平均 BLEU が低い歌詞には英語が含まれていたり、単語の羅列や繰り返しが多用されていた歌詞が確認できた。また、平均 BLEU が高い歌詞は、英語があまり含まれておらず、歌詞が文章の形態を取っているものが多く確認できた。

1) 豊橋技術科学大学 大学院工学研究科 情報・知能工学専攻

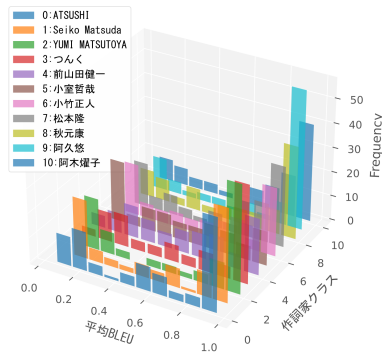


図 2 平均 BLEU のヒストグラム

## 5 分析

ここでは、平均 BLEU が高い歌詞と低い歌詞でそれぞれの特徴をより詳しく分析を行う。分析のために平均 BLEU が高い歌詞と低い歌詞の基準を次のように設ける。平均 BLEU が 0.8 以上を平均 BLEU が高い歌詞とし、その対象曲は 487 曲である。また平均 BLEU が 0.2 以下を平均 BLEU が低い歌詞とし、その対象曲は 198 曲である。これらの歌詞を対象に以降の方法で分析を行う。

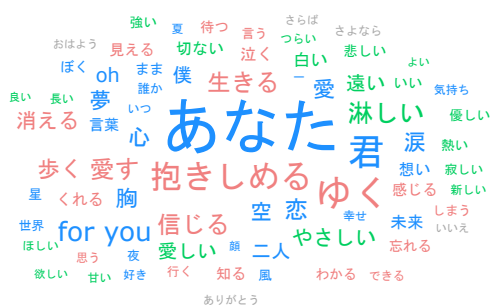
### 5.1 分析方法

分析には、ユーザーローカル提供のテキストマイニングツール [8] を用いる。このテキストマイニングツールは、文書のワードクラウドを作成したり、単語の出現頻度を算出したり、単語の出現パターンが似たものを算出し 2 次元空間にマッピングする共起キーワード分析などを行ったりすることができるツールである。今回は 2 つの文書を比較する機能を用いて BLEU が高い歌詞と低い歌詞の比較を行うことでそれぞれの歌詞の特徴を分析する。

### 5.2 分析結果

分析結果の一例であるワードクラウド分析の結果を図 3 に示す。

図 3 のワードクラウドにおいて各単語の色はその単語の品詞を表しており、赤字が動詞を、緑字が形容詞を、



(a) high

青字が名詞を、灰色が感動詞をそれぞれ表している。図 3 のワードクラウドを見ると、平均 BLEU が高い歌詞には日本語が多く名詞や動詞が多く出現していることがわかる。平均 BLEU が低い歌詞は英語が多く出現しており動詞や形容詞は余り出現していないことがわかる。

## 6 おわりに

本稿では、GPT-2 を作詞家 11 人の曲 1,100 曲分のデータでファインチューニングを行い、歌詞の冒頭 10% 分の文字を入力として歌詞の生成を行った。そして、生成された歌詞と元の歌詞の BLEU スコアを算出し、BLEU が高い歌詞と低い歌詞の特徴の分析を行った。分析の結果、BLEU スコアが高い歌詞は歌詞が長く文章の形態を取っており、また BLEU スコアが低い歌詞は歌詞が長く単語の羅列や繰り返しが多いといった知見が得られた。

今後の検討として、BLEU が高い低いの観点だけでなく、作詞家ごとにより詳細な分析を進めたいと考える。また、今回得られた知見をもとに、作詞家の特徴を持った歌詞の生成を行いたいと考えている。

### 謝辞

本研究は科研費基盤研究 (C) 22K12040 の支援を受けたものである。

### 参考文献

- [1] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. Attention is all you need. *CoRR*, Vol. abs/1706.03762, 2017.
- [2] Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, Ilya Sutskever, et al. Language models are unsupervised multitask learners. *OpenAI blog*, Vol. 1, No. 8, p. 9, 2019.
- [3] 嵯峨山茂樹. 自動作曲システム orpheus (小特集創造性・芸術性における ai の可能性). 電子情報通信学会誌 = The journal of the Institute of Electronics, Information and Communication Engineers, Vol. 102, No. 3, pp. 214–220, 3 2019.
- [4] tanikawa. Ai 作詞ツール shikaki (シカキ). <https://shikaki.diatonic.codes/>. (Accessed on 06/09/2022).
- [5] 笠松雅史, 青野雅樹. Bert とその拡張モデルを用いた作詞家推定手法と分析. 言語処理学会 第 28 回年次大会 発表論文集, pp. 1274–1278, 2022.
- [6] 歌詞検索サービス 歌ネット. <https://www.uta-net.com/>.
- [7] 趙天雨, 沢田慶. 日本語自然言語処理における事前学習モデルの公開. 人工知能学会研究会資料 言語・音声理解と対話処理研究会, Vol. 93, pp. 169–170, 2021.
- [8] Ai テキストマイニング by ユーザーローカル. <https://textmining.userlocal.jp/>.



(b) low

図 3 平均 BLEU が高い歌詞と低い歌詞のワードクラウド