

2023 年度 修士論文



川柳の理解と生成

Understanding and Generation of Senryu

指導教員 河原 大輔 教授
研究指導名 自然言語処理研究

早稲田大学 基幹理工学研究科 情報理工・情報通信専攻

学籍番号 5122F020

太田 聖三郎

2024 年 1 月 22 日

概要

川柳は日本の伝統文芸の一つである。詩や俳句に関する機械学習を用いた研究は散見されるが、川柳に関するものはない。本研究では、川柳を構成する要素を様々な指標に分解し、それらの予測を組み合わせることで、より高精度な川柳評価を予測する手法を提案する。クラウドソーシングを用いて川柳データセットを構築し、BERT-like モデルをファインチューニングすることより川柳の評価を予測する。それに加え、川柳をデータに事前学習モデルの追加学習を行い、川柳の特徴を理解した SenryuBERT を構築する。提案手法は、川柳の総合的な評価を直接予測するベースラインより、若干の精度向上を達成した。しかしながら、川柳の理解は難しいタスクであり、今後も研究の余地がある。また、川柳の理解に加え川柳生成モデルの構築を通じて、自然言語処理技術の創造性の向上に焦点を当てる。川柳生成において、音数やお題に沿った内容、おもしろさの学習を組み合わせ、end-to-end での生成を可能にする。学習には川柳を用いたファインチューニングの他に、川柳以外のデータを用いた学習や強化学習などを取り入れる。クラウドソーシングによる人手評価や自動評価により、提案手法はベースラインに比べて音数制御の精度、川柳のとしての適否、おもしろさの面で優れていることを示した。今後は、人間の川柳との相違点を埋めつつ、モデルの更なる洗練を目指す。

目次

1	はじめに	5
2	関連研究	6
2.1	文芸.....	6
2.2	言語モデル/LLM.....	7
3	川柳の理解	8
3.1	提案手法: 川柳理解モデルの構築.....	8
3.2	実験: 川柳のおもしろさ評価.....	11
4	川柳の生成	13
4.1	提案手法: 川柳生成モデルの構築.....	13
4.2	実験: 生成川柳の評価.....	16
5	おわりに	19
A	学習に使用したモデル	22
B	Wikipedia データセットの処理	22
C	川柳データの前処理	22
D	川柳嗜好データセットの詳細	23
E	ハイパーパラメータ	23

図目次

1	Structure-aware training.....	14
2	クラウドソーシングによる音数の評価.....	17
3	KWJA による音数の評価.....	18

表目次

1	追加学習における学習条件設定.....	9
2	[MASK] トークンの予測例 (入力: “歯ブラシと一緒に [MASK] あの”).....	9
3	[MASK] トークンの予測例 (入力: “[MASK] の財布の中身 500 円”).....	9
4	総合評価の予測結果.....	12
5	各指標の評価の予測結果.....	12
6	各指標における川柳評価の予測例と比較 (川柳: “札束を 数え快感 銀行員”, お題: “現実逃避”).....	13

7	総合評価の予測結果の比較	13
8	総合評価の予測結果の比較	14
9	クラウドソーシングによる川柳としての適否の評価	18
10	クラウドソーシングによるおもしろい川柳かの評価	19
11	報酬モデルによる川柳の評価	19
12	学習に使用した事前学習モデルの詳細	22
13	得票数の差による margin のアノテーション	23
14	RoBERTa の追加学習に使用したハイパーパラメータ	24
15	BERT-like モデルのファインチューニングに使用したハイパーパラメータ	24
16	LightGBM の学習に使用したハイパーパラメータ	24
17	CLM 学習に使用したハイパーパラメータ	24
18	PPO 学習に使用したハイパーパラメータ	24
19	報酬モデルの学習に使用したハイパーパラメータ	24

1 はじめに

人工知能による創作活動の発展はめざましい。自然言語を入力するだけで、その内容に沿った精巧な画像を生成するモデル [Rombach 21] も登場している。このように、創作モデルの精度が年々向上している中、モデルが人間に取って代わることを危惧する意見も存在する。しかし、これらのモデルは人間の代わりとしてだけではなく、人間の補助としても在り得る。つまり、これからは人間とモデルが共創していく時代である。

文芸に着目すると詩や俳句などの生成研究 [Bowman 15, Hitsuwari 23] がある。その一方で、それらの理解に関する研究はほとんどされていない。それらの理解には創作物の持つおもしろさや独創性などが関わってくる。おもしろさは人のコミュニケーションの中でも大きな役割を担っている。つまり、人と計算機のコミュニケーションを円滑にするためには、計算機によるおもしろさの理解が重要な要素となる。そして、人間とモデルがうまく協力することでより良い作品の創出が可能となる。

日本の伝統的な文芸の一つに川柳がある。川柳は俳句と同様に五・七・五の音数を持ち、元来は俳諧連歌の長句が独立したものである。短句に対しての前句付けが流行したことにより生まれた文芸であり、口語的で風刺的な内容が特徴である。しかし、川柳という文芸に関しては、理解・生成いずれにおいても機械学習による研究はされていない。川柳のおもしろさは人が理解するにおいても難しく、これを計算機に理解させるのはさらに難しいタスクである。実際、GPT-4などの基盤モデルも依然としておもしろい川柳の生成をすることはできない。本研究では川柳という文芸を対象とし、人が川柳のおもしろさをどのように理解しているか、また、計算機がそれを理解できるかを分析する。この問題に取り組むにあたって、川柳のおもしろさを構成する要素を細分化し、判断しやすい指標を設けることで、川柳の総合的なおもしろさの予測を目指す。さらには、得られた川柳理解の知見を応用することで、おもしろい川柳の生成を実現する。

川柳のおもしろさを構成する要素を以下のように細分化し、判断しやすい指標を設けることで、川柳の総合的なおもしろさの予測を目指す。

1. 句がお題に沿っている (along)
2. 場面をイメージしやすい・わかりやすい (imaginable)
3. 句に対比の構図がある (contrast)
4. あたりまえ・ありきたりなことは言っていない (usual)
5. 不適切な表現が含まれていない (appropriate)
6. 語のリズム感・軽快さがある (rhythmic)
7. 言い回しに独創性がある (unique)
8. 体験や体感への想いが感じ取れる・余韻がある (experience)

川柳のおもしろさを定量分析するために、クラウドソーシングによって各指標をアノテーションする。得られたアノテーションをもとに言語モデルのファインチューニングを行い、川柳

を入力し川柳の総合的な評価を予測するタスクを行う。川柳の評価を直接予測する場合と、細分化した指標の予測を組み合わせる場合の精度の差を比較する。

さらに、構築した川柳理解モデルを、強化学習時の報酬モデルとして応用することで、おもしろい川柳の生成を実現する。具体的には、以下の3つのステップにより達成する。

- 音数・トピックの学習
- 川柳の内容の学習
- 川柳のおもしろさの学習

本論文の主な貢献は以下である。まず、川柳理解に関して、川柳投稿サイト「まるせん」の投稿川柳から取得したデータを元に、お題や不適切な表現をフィルタリングし、クラウドソーシングで川柳の評価をアノテーションしたデータセットを構築した。川柳の特有の文脈やリズム感を理解するために、SenryuBERT と呼ぶ言語モデルを構築した。川柳のおもしろさを直接予測するだけでなく、複数の指標に細分化してから予測する手法を提案し、その有効性を検証した。各指標の評価をもとに総合評価を予測する実験を通じて、提案手法がベースラインに比べて僅かながらも高い精度で予測できることを示した。そして、BERT系のモデルでも Instruction Tuning が有効であることを示した。

次に、川柳生成に関して、言語モデルに end-to-end で川柳を生成させるための手法を3つのステップに分けて提案し、言語モデルが川柳の特徴を段階的に学習することを可能にした。クラウドソーシングと自動評価により、提案手法が全てのステップで、ベースラインの評価を上回り、その有効性を確認できた。特に、川柳としての適否は提案手法により、大幅な効果が見られた。

2 関連研究

2.1 文芸

2.1.1 詩・俳句生成

川柳について機械学習を用いた研究は無い。しかし、川柳に通ずるものとして、近年では詩や俳句の生成 [Bowman 15, Hitsuwari 23] が行われている。Hitsuwari ら [Hitsuwari 23] は人手評価において、人工知能が生成した俳句と人間が作った俳句との見分けがつかないほど完成度の高い川柳の生成を実現している。

2.1.2 大喜利データセット

中川ら [中川 19] は大喜利を複数の指標に分解し、クラウドソーシングを用いて各指標に対してアノテーションを行うことで、定量分析可能な大喜利データセットを構築している。

2.1.3 俳句の生成

川村ら [川村 21] は, LSTM を用いて俳句を生成する手法を提案している. しかし, この手法はモデルが end-to-end で俳句を生成することはできず, 俳句生成後にフィルタリングを行っている. また, Transformer [Vaswani 17] 登場以降に俳句や川柳の生成を行った研究は存在していない. 本研究では, end-to-end で川柳を生成できるモデルの構築を目標とする.

2.2 言語モデル/LLM

2.2.1 BERT/RobERTa

近年の自然言語処理研究の飛躍的な進歩には BERT [Devlin 19] の登場が大きく関わっている. BERT は, self-attention を導入することで自然言語処理にブレイクスルーをもたらした Transformer [Vaswani 17] の Encoder 部分のみを使用している. BERT の学習は事前学習とファインチューニングからなる. 事前学習は自己教師あり学習であり, 主な学習方法は Masked Language Model (MLM) である. MLM は入力文の一部のトークンをマスクし, マスクされたトークンを予測するタスクである. 予測されるトークンのうち 80% は [MASK] トークンに置き換えられ, 10% はランダムな他のトークンに置き換えられる. 残りのトークンは元のトークンのままで残される. これは, ダウンストリームタスクにおいて [MASK] トークンが出現しないことを考慮したためである. Liu ら [Liu 19] は BERT を改良したモデル RoBERTa を提案し, データセットの増大や動的なマスキングなどにより精度を向上させた.

2.2.2 GPT/GPT-2

GPT [Radford 18] は BERT と同様に Transformer を基にしたアーキテクチャであるが, BERT とは異なり, decoder 部分を使用している自己回帰モデルである. GPT の主な学習方法は Causal Language Model (CLM) である. CLM は前の入力から次の単語を予測していくタスクである. GPT-2 [Radford 19] は GPT の改良モデルであり, モデルのパラメータ数や学習データ量を増やすことで精度を向上させている. 最近では ChatGPT などの LLM (大規模言語モデル) が話題となっているが, これらは基本的に Transformer の decoder のアーキテクチャを採用しており, CLM で事前学習されている.

2.2.3 PoeLM

Ormazabal ら [Ormazabal 22] は教師無し学習による韻律や音韻を学習する手法, Structure-aware training を適用した形式詩生成モデル PoeLM を提案している. この手法では, 詩データを必要とせずに学習を行うことができる. 川柳は通常の学習コーパスと比べデータ (トークン) 数が圧倒的に少ないため, 安定した学習を行うことは難しい. しかし, この手法を適用することで川柳以外のデータも使用することができるため, 学習の安定化が期待できる.

2.2.4 LLM における強化学習

昨今の自然言語処理における学習の潮流は大規模言語モデル (LLM) に関する学習が主となっている。ChatGPT などに代表される LLM では強化学習が適用されている。RLHF (Reinforcement learning from human feedback) [Christiano 17, Stiennon 20] はその中でも特によく用いられる技術であり、強化学習の報酬モデリングに人間の嗜好を反映させた学習データを用いることで、モデルにより人間が好ましいと感じる応答を生成させることを可能にしている。Llama2 [Touvron 23] では報酬モデリング時のデータに“margin”というペア間の嗜好の差を導入することで報酬モデルの精度を向上させている。本研究では、川柳のおもしろさを人間に評価してもらうことで、モデルがよりおもしろい川柳を生成できるようになることを試みる。

3 川柳の理解

3.1 提案手法: 川柳理解モデルの構築

本節では、言語モデルの学習を通じて川柳の理解（評価）を行う。学習に使用する川柳は川柳投稿サイト「まるせん」¹より提供を受けたものを用いる。はじめに、川柳の文脈を理解する言語モデルを構築する。その後、クラウドソーシングにより集めた川柳評価データを用いてモデルをファインチューニングすることで、川柳を理解するモデルを構築する。学習のベースとなるモデルには日本語の RoBERTa-base²を使用する。

3.1.1 SenryuBERT の構築

既存の言語モデルは新聞やウェブテキストで学習しているため、川柳の風刺的な文脈や語順などのリズム感を理解できない可能性がある。この問題に取り組むために、川柳データを活用し、言語モデルの追加学習を行うことで SenryuBERT を構築する。「まるせん」よりいただいたデータは川柳、お題、作者を含む 581,336 句からなり、これらを学習データとして使用する。

ランダムトークン置換 (RTR) 2 節で述べた通り、BERT [Devlin 19] の MLM では、予測するトークンのうち、80%を [MASK] トークン、10%をランダムな他のトークン、10%をそのままのトークンにして学習を行う。川柳のような短い系列を入力とする場合、[MASK] トークンと異なり、ランダムなトークンに置き換えると、BERT の特徴量はランダムなトークンの影響を大きく受けるため、文脈が破綻する可能性がある。これを防ぐため、通常の MLM に加えて、ランダムなトークンの置き換えをしない学習も行い比較する。

追加学習 川柳データを用いた言語モデルの追加学習は、RoBERTa-base をベースとして行う。学習は表 1 に示す 4 種類の条件で行う。モデル 3 の場合は入力を“お題 [SEP] 川柳”とし、その他のモデルでは“川柳”とする。また、学習時にマスクされるトークンにお題は含まれないようにする。実験で使用するハイパーパラメータを付録 E に示す。

¹<https://marusenryu.com/>

²<https://huggingface.co/nlp-waseda/roberta-base-japanese-with-auto-jumanpp/>

表 1: 追加学習における学習条件設定

条件	[MASK]	同じ	ランダム	お題付き
1	80%	10%	10%	無し
2	90%	10%	0%	無し
2'	80%	20%	0%	無し
3	80%	10%	10%	有り

表 2: [MASK] トークンの予測例 (入力: “歯ブラシと一緒に [MASK] あの”)

条件	ランク 1		ランク 2		ランク 3	
	トークン	確率	トークン	確率	トークン	確率
BL	の	0.109	ない	0.058	いる	0.045
1	使う	0.191	使った	0.060	捨てる	0.038
2	消えた	0.441	消える	0.034	捨てた	0.027
3	シャンプー	0.116	消えた	0.049	帰った	0.028

表 3: [MASK] トークンの予測例 (入力: “[MASK] の財布の中身 500 円”)

条件	ランク 1		ランク 2		ランク 3	
	トークン	確率	トークン	確率	トークン	確率
BL	彼氏	0.039	付録	0.031	雑誌	0.020
1	ブランド	0.770	金持ち	0.013	母さん	0.007
2	休日	0.071	ブランド	0.060	コンビニ	0.031
3	コンビニ	0.061	プレゼント	0.043	誕生日	0.034

結果 各モデルの [MASK] トークンの予測例を表 2, 3 に示す。BL はベースライン (RoBERTa) を表す。例には条件 1, 2, 3 とベースラインの結果を示す。ベースラインと比べて、より川柳の文脈に沿ったトークンを予測できていることがわかった。定量的な評価は 3.2 節のダウンストリームタスクにて行う。

定性的評価 ベースラインの RoBERTa では助詞や関係のない名詞・助詞が予測され、正解と近い単語の出力は見られない。条件 3 のお題と川柳を入力として学習したモデルではベースラインほど遠くはないものの、正解に近い予測はほとんど見られなかった。条件 1 と 2 のモデルはどちらもある程度のニュアンスが捉えられており、正解に近い予測ができています。条件 2 のモデルでは生起確率が他と比べて高くなっており、これはランダムトークンによる文脈の揺れが抑えられたためと思われる。

3.1.2 川柳評価データセットの構築

川柳データの前処理・フィルタリング 本研究では、川柳の評価をクラウドソーシングを用いてアノテーションする。アノテーション対象の川柳を選別するために、川柳の前処理・フィルタリングを行う。投稿川柳にはしばしば絵文字を含んでいるものがある。これらは実験で用いる言語モデルの語彙に含まれていないため削除する。また、上記サイトではお題を作者が自ら設定し投稿するため、不適当なお題やイレギュラーなお題が散見される。そのため、

クラウドソーシングにかける川柳については、運営によりあらかじめお題が指定されている「コンテスト」の川柳を採用する。コンテストはサイト内で定期的に行われ、その度に10句ほどの入選作品が選出されるため、ある程度の川柳の質が担保される。データ収集時点までに開催されたコンテスト304回のうち、時事的なお題（例: VAR）や長いお題（例: 最近の若いもん）を除き名詞や形容詞を中心に250題について、フィルタリングを行った。各お題について入選作品+それ以外を合わせ20句ずつを選出し、合計5,000句の川柳を抽出した。

川柳のおもしろさの細分化と評価のアノテーション 前節で述べた5,000句の川柳に対し、クラウドソーシングを用いてアノテーションを実施する。プラットフォームはYahoo!クラウドソーシング³を用いる。川柳の解説書[新家11, 野林14]を参考に、川柳を8つの指標に細分化し、それぞれの指標について5段階評価を行ってもらう。細分化した指標を下記に示す。

1. 句がお題に沿っている (along)
2. 場面をイメージしやすい・わかりやすい (imaginable)
3. 句に対比の構図がある (contrast)
4. あたりまえ・ありきたりなことは言っていない (usual)
5. 不適切な表現が含まれていない (appropriate)
6. 語のリズム感・軽快さがある (rhythmic)
7. 言い回しに独創性がある (unique)
8. 体験や体感への想いが感じ取れる・余韻がある (experience)

これらの項目に、

- 総合的に良い川柳である (overall)

を加え、合計9個の指標について、1句あたり10人のクラウドワーカーを雇い、その平均値を正解ラベルとして付与する。クラウドワーカーにはお題、川柳が与えられる。

3.1.3 川柳のおもしろさの直接予測と間接予測

本研究では、川柳のおもしろさという曖昧な要素を直接予測する場合と、本節で示した8つのわかりやすい指標に分解してから間接的に予測する場合の精度を比較する。また、予測した各指標のスコアから総合評価を予測するタスクを行い、総合評価の予測精度の比較を行う。さらに、SenryuBERTの学習にInstruction Tuningを適用した際の精度予測の比較も行う。

実験では、事前学習済みBERT-likeモデルを回帰問題としてファインチューニングすることで、川柳から総合評価や各指標の評価を予測する。間接予測では、予測した各指標のスコアから決定木アルゴリズムを用いて総合評価を予測する。わかりやすい指標の予測から総合評価を予測することで、言語モデルの予測精度が向上することが考えられる。またWei[Wei 22]らは、Instruction Tuningを導入し、LLMが様々なタスクをインストラクションとともに大量に学習することで、未知のタスクでのパフォーマンスが向上することを示している。これをSenryuBERTの学習にも適用することでも精度の向上を図る。

³<https://crowdsourcing.yahoo.co.jp/>

3.2 実験: 川柳のおもしろさ評価

3.2.1 実験設定

本節では以下の4つの実験を行う。

1. 川柳から総合評価の予測
2. 川柳から各指標の評価の予測
3. 各指標の評価から総合評価の予測
4. Instruction Tuning [Wei 22] による総合評価の予測

実験1, 2には既存のモデルや3.1.1節で構築したモデルなどのBERT-likeモデルを使用し, 実験3には決定木アルゴリズムのLightGBM [Ke 17]を使用する。実験4ではSenryuBERTをRoBERTa-largeで構築したSenryuBERT-largeを使用する。

実験1, 2では3.1.1節で構築したSenryuBERTに加え, BERTやRoBERTaでの学習を行う。モデルの詳細と学習に用いたハイパーパラメータは付録A, Eに示す。学習に使用するデータセットは本節で構築した川柳評価データセットを用いる。データセットは訓練用4,000句, 検証用500句, テスト用500句に分割する。川柳はお題ありきであるため, モデルへの入力は“川柳 [SEP] お題”とする。各指標についてファインチューニングを10回行い, 精度を平均値と最大値で算出する。モデルの評価にはPearsonの相関係数を用いた。

実験3ではまず, 実験1, 2から最良のモデルを選択し, 訓練データの各指標を予測する(自動ラベルと呼ぶ)。各指標の自動ラベルからLightGBMを用いて総合評価を予測し, 実験1のスコアと比較する。LightGBMの学習に用いたハイパーパラメータを付録Eに示す。訓練データに自動ラベルを付与する手順は, まず, 訓練用データを10分割する。そのうち9個のデータ(3,500句)を用いて学習したモデルで, 残りの1個のデータに対して各指標のスコアを予測しラベルとする。これを10個の分割全てで行い, 4,000句に自動ラベルを付与する。

実験4では, これをSenryuBERTのファインチューニングに応用し, 分解したすべての指標と総合評価の予測を同時に学習することで, 総合評価の予測精度が向上することを検証する。モデルへの入力は“指標 [SEP] 川柳 [SEP] お題”とする。正解ラベルを入力する指標に対応するスコアに設定する。

3.2.2 川柳から総合評価の予測

川柳から直接総合評価を予測した際の評価結果を表4に示す。3.1.1節で提案したSenryuBERTはRoBERTaをベースとしたモデルであるが, RoBERTaベースのモデルに関してはすべてBERTよりも最大値で良い精度となったが, RoBERTaとSenryuBERTの間に大きな差はみられなかった。しかし, SenryuBERTにおいてRTRを行わなかった場合や, お題を入力に加えて学習した場合は精度の向上が見られた。最良のモデルはSenryuBERTにおいて, RTRを行わず, [MASK]トークンの割合を80%とした場合で, SenryuBERTとは最大で約0.07の差が見られた。ただし, お題を入力に加えて学習したSenryuBERT以外のモデルは精度が安定せず, データの順序などの影響を受けやすいことが分かった。表6に総合評価を川柳から直接予測する場合と提案手法との比較例を示す。提案手法が僅かに良い予測ができていることがわかる。表6に示す指標は3.1節で示した指標に対応している。

表 4: 総合評価の予測結果

モデル	Pearson (ave)	Pearson (max)
BERT _{BASE}	0.353	0.373
RoBERTa _{BASE}	0.314	0.412
SenryuBERT	0.246	0.419
– RTR (80%)	0.418	0.490
– RTR (90%)	0.350	0.462
+ お題	0.420	0.456

表 5: 各指標の評価の予測結果

モデル	average		max		average		max		average		max	
	along		imaginable		contrast		usual					
SenryuBERT	0.092	0.227	0.280	0.322	0.154	0.324	0.093	0.177				
– RTR (80%)	0.178	0.330	0.357	0.422	0.307	0.407	0.158	0.216				
– RTR (90%)	0.208	0.300	0.359	0.430	0.251	0.373	0.185	0.213				
	appropriate		rhythmic		unique		experience					
SenryuBERT	0.124	0.242	0.092	0.287	0.182	0.279	0.227	0.412				
– RTR (80%)	0.246	0.352	0.260	0.330	0.325	0.352	0.423	0.476				
– RTR (90%)	0.244	0.320	0.214	0.333	0.279	0.363	0.471	0.496				

3.2.3 川柳から各指標の評価の予測

川柳を入力として3.1節に示した各指標の評価を予測した結果を表5に示す。評価指標はPearsonの相関係数である。表5にはSenryuBERTのファインチューニングによる結果と、総合評価の予測にて高精度であったRTRを行わないモデルの結果を示す。いずれの指標においても、RTRを行わないモデルはSenryuBERTの精度を上回る結果となった。この内、総合評価の予測より高い精度となったのは、体験や体感への想いが感じ取れる・余韻がある (experience) のみであった。

3.2.4 各指標の評価から総合評価の予測

本実験では、川柳から各指標の評価の予測において精度の高かったSenryuBERT–RTR (90%)のモデルを用いて訓練データに自動ラベルを付与し、LightGBMの学習を行う。川柳から直接総合評価を予測した結果と、各指標のスコアの予測から総合評価を予測した結果を表7に示す。データセットの各指標の正解ラベルをもとに総合評価を予測する場合には、非常に高い精度で予測ができています。これは、細分化した8個の指標が、人が川柳のおもしろさをどう理解しているかをうまく表す基底になっていると理解できる。自動ラベルで学習した場合には、正解ラベルと比べて大幅に精度は下がるが、川柳から直接予測する場合と比べると、0.01ほど高い精度が得られ、提案手法がやや高い精度で予測できていることがわかる。自動ラベルの精度があまり高くない原因として、experienceしかoverallの精度に勝っていないことが考えられる。

表 6: 各指標における川柳評価の予測例と比較 (川柳: “札束を 数え快感 銀行員”, お題: “現実逃避”)

手法	along	imaginable	contrast	usual	appropriate	rhythmic	unique	experience	overall
正解ラベル	2.9	3.9	2.7	2.8	3.5	3.3	3.0	3.4	2.9
直接予測	-	-	-	-	-	-	-	-	3.264
提案手法	3.366	4.038	3.585	3.495	3.434	3.197	3.074	3.714	3.212

表 7: 総合評価の予測結果の比較

モデル	Pearson
LightGBM (正解ラベル)	0.8912
LightGBM (自動ラベル)	0.4760
SenryuBERT (直接予測)	0.4618

3.2.5 Instruction Tuning による総合評価の予測

SenryuBERT-large の学習に Instruction Tuning を適用し、総合評価の予測精度向上を図る。総合評価を予測した結果を表 8 に示す。個別予測は SenryuBERT-large を 1 つの指標のみでファインチューニングした際のスコアである。評価指標は Pearson の相関係数である。Instruction Tuning を行うことにより、imaginable 以外の指標で精度が向上したことが確認できた。GPT 系の LLM で Instruction Tuning が多く適用される中、モデルサイズの大きくない BERT 系のモデルでの有効性が示された。

4 川柳の生成

4.1 提案手法: 川柳生成モデルの構築

前節までの実験により、川柳を理解するモデルを構築した。このモデルを活用し、おもしろい川柳を生成するモデルの構築を目指す。本章では、本研究において言語モデルに良い川柳を end-to-end で生成させるための学習手法を述べる。本研究では 3 つのステップに分けて川柳の特徴を学習する。学習のベースとなるモデルには日本語 GPT-2⁴を使用する。

4.1.1 音数・トピックの学習

通常、言語モデルは音数を明示的に学習することはないが、正確に 5-7-5 の音数で川柳を生成させるためには音数の情報を学習させる必要がある。また、川柳はあるお題に対して作られるものであり、そのお題に沿った内容である必要がある。このお題に対して生成するという構図を明示的に学習することで、川柳を一貫して生成するモデルを構築する。

学習に使用可能な川柳のトークン数は、一般的な言語モデルの学習に使用されるコーパスのトークン数に比べて圧倒的に少ない。そこで 2 節で述べた structure-aware training を用いて、川柳データを使わずに川柳の形式を学習する。PoeLM では形式詩の形式を韻律や音韻

⁴rinna/japanese-gpt2-medium

表 8: 総合評価の予測結果の比較

指標	個別予測	Instruction Tuning
overall	0.459	0.568
along	0.590	0.653
imaginable	0.529	0.523
contrast	0.573	0.604
rhythmic	0.298	0.543
inappropriate	0.545	0.576
usual	0.544	0.564
unique	0.348	0.395
experience	0.321	0.405

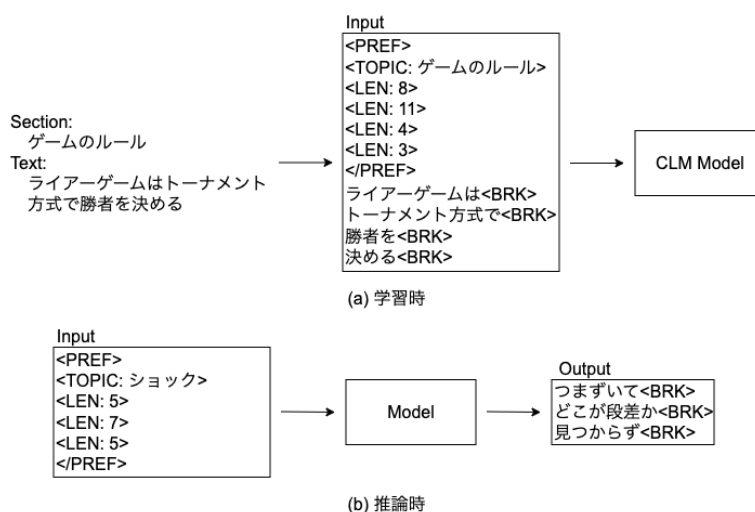


図 1: Structure-aware training

数の情報を付加することで学習させている。本研究では、川柳の特徴的な要素のうち「5-7-5の音数を守った生成」、「お題に沿った生成」を目的に学習を進める。図1は本研究におけるstructure-aware trainingの概要図である。

データセット 学習には日本語のWikipediaデータセット⁵を使用する。このデータセットには記事のタイトル、セクション名、テキストが段落ごとに格納されている。本研究では、セクション名を川柳のお題、テキストを句に見立て、お題や音数の制約に則った生成を行うための学習用に整形を行う。川柳における5-7-5のようにテキストをある長さで区切り、区切りごとの音数を付与する。また、セクション名を対応するテキストのトピックとする。本研究では区切りの目安を文節とし、文節の取得には構文解析器KNPを、音数の取得には形態素解析器Juman++から得られる単語の読みを用いる。最終的なデータセットはあるテキストに対するトピック、文節のリスト、文節ごとの音数のリストから成る。学習をスムーズに行うためにデータセットの前処理・フィルタリングを行う。まず、テキストは段落の一文目を使用し、丸括弧で囲まれている部分を削除する。また、Juman++では英数字の音数が正確に取

⁵singletongue/wikipedia-utils

得できないため英数字を含む文を除外する。そして、長（短）すぎる文は除外し、最終的なデータ数は2,282,186となる。

モデルの入力は<PREF><TOPIC: >(<LEN: {}>)*n</PREF>({}<BRK>)*nとし、<PREF>、</PREF>、<BRK>、TOPIC、LENは特殊トークンとしてトークナイザの語彙に追加する。TOPICの後には入力文のセクション名、LENの後には各文節の音数を入力する。nは入力文の文節数である。学習のタスクはCLM (Causal Language Model)で、次のトークンを予測する学習手法である。推論時には<PREF> </PREF>を入力し、川柳の部分を生成する。学習したモデルを“Poelm-ja”と呼ぶ。学習に使用したハイパーパラメータを付録Eに示す。

4.1.2 川柳の内容の学習

前節で学習したモデルを川柳生成用に調整する。前節では川柳の形式を学習させたに過ぎないため、川柳を学習データに用いて川柳の内容を学習する。前節では文節を分割の区切りとしていたが、ここでは川柳の5-7-5を区切りとすることで、川柳の音数に合わせたファインチューニングを行う。具体的には、4.1.1節で適用したstructure-aware trainingを川柳を用いて行う。

川柳分割モデルの学習 川柳投稿サイト「まるせん」より提供された利用可能な川柳データ約58万句の内、約65%は5-7-5に分かたれていない川柳となっている。これらの川柳はstructure-aware trainingでは適用することができないため、ファインチューニングに使用することができない。この問題を解消するため、あらかじめ5-7-5に分かたれている川柳を用いて、川柳5-7-5に分割するモデルを構築する。5-7-5の間の空白を削除した川柳をモデルの入力とし、5-7-5の間に空白を含んだ川柳を正解として学習を行う。モデルはGPT-2とT5、GPT-NeoXを使用する。モデルの詳細は付録Cに示す。

ファインチューニング 構築した川柳分割モデルの推論により、5-7-5に分割されていない川柳を分割する。各モデルに対し分割された川柳を生成し、正確な分割がなされているかフィルタリングを行う。川柳が正確に5-7-5に分割されているかを判定するのは困難なため、ここでは単純に空白を2つ未満、2つより多く含む川柳を除外する。また、モデル間で異なる生成をした場合は多数決をとり、同様の生成が多いパターンを採用する。この操作により分割された川柳と、あらかじめ5-7-5に分割されている川柳を合わせてファインチューニング用のデータセットとする。このデータセットを用いて前節で学習したモデルをファインチューニングする。入力の形式は前節の学習と同様であり、“TOPIC”の後には川柳のお題、“LEN”の後には5-7-5を順番に入力する。学習のタスクはCLMである。学習したモデルを“SenryuGPT”と呼ぶ。

4.1.3 川柳のおもしろさの学習

前節までのステップにより川柳の音数やお題に沿った生成を学習し、さらに、生成する内容を川柳に整えた。本節では、人間の嗜好をとらえた報酬モデルを使用した強化学習を適用し、よりおもしろい川柳を生成するモデルを構築する。

川柳嗜好（プリファレンス）データセット 強化学習時の報酬モデリング用に川柳の良し悪しを比較した川柳嗜好データセットを構築する。Yahoo!クラウドソーシング⁶を用いて川柳のペアに対してアノテーションを行う。あるお題に対する2つの川柳をクラウドワークに提示し、良い（おもしろい）と感じる川柳を選択してもらう。川柳は300個のお題に対し5句ずつ用意、すなわち、1つのお題に対し10ペアを作成する。お題は川柳投稿サイト「まるせん」で開催されるコンテストのお題とChatGPTにより生成したものを使用する。お題ごとに用意する5句の川柳の内、「まるせん」のコンテストの入選作があるお題は1句を入選作とし、4句をSenryuGPTで生成したものとする。入選作がない場合は5句すべてをSenryuGPTで生成する。各ペアに対して10人で評価し、得票数の多かった川柳をお題に対する“chosen”，少なかった川柳を“rejected”と呼ぶ。得票数が等しい場合は、クラウドワークに先に提示された句をchosenとする。また得票数の差に応じて“margin”を設定する。marginの設定の詳細は付録Dに示す。

報酬モデルの構築 川柳嗜好データセットを使用して、お題に対する川柳の良し悪しを回帰により評価する報酬モデルを構築する。ベースモデルには3.1.3節で構築したInstruction Tuningを適用したSenryuBERTを採用する。報酬モデルは、好まれる川柳と好まれない川柳のギャップを最大化するように学習する。損失関数は次式で定式化する。

$$L = -\log(\sigma(R_{chosen} - R_{rejected}) - M) \quad (1)$$

ここで、 σ はシグモイド関数、 R_{chosen} と $R_{rejected}$ はそれぞれ“chosen”，“rejected”の報酬を示す。 M は“margin”である。

強化学習 強化学習にはProximal Policy Optimization (PPO)を使用する。構築した報酬モデルを使用し、SenryuGPTに強化学習を適用する。学習時は、SenryuGPTの入力形式に則って、お題と音数を入力し、川柳部分をモデルに生成させる。それらを元に報酬を最大化するように学習する。学習に使用のお題は「まるせん」のコンテストのお題や、ChatGPTの生成により1,000個収集した。学習に使用したハイパーパラメータを付録Eに示す。学習したモデルを“SenryuGPT-ppo”と呼ぶ。

4.2 実験: 生成川柳の評価

4.2.1 実験設定

4.1節で構築したモデルや既存のモデルによる川柳生成の質を評価する。本実験では3つの要素についての評価を行う。

- Structure-aware training による音数制御の精度
- 川柳としての適否
- おもしろさの評価

これらについて、クラウドソーシングによる人手評価とテキスト解析ライブラリKWJA[植田 22]や強化学習時に用いた報酬モデルでの自動評価を実施する。

⁶<https://crowdsourcing.yahoo.co.jp/>

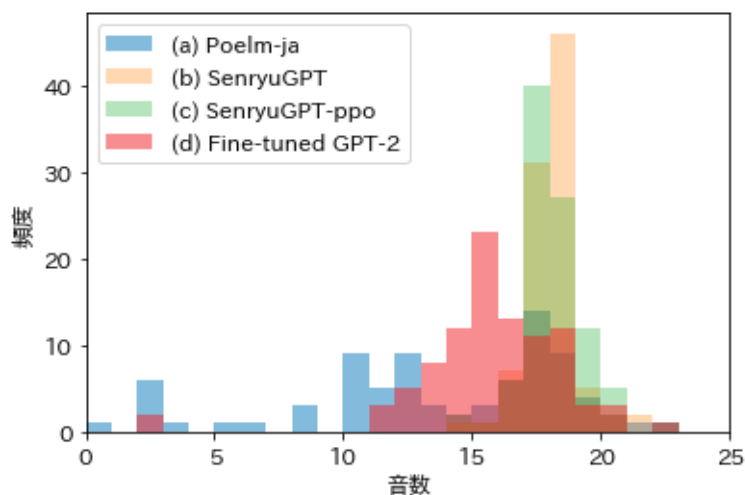


図 2: クラウドソーシングによる音数の評価

ベースライン CLMにより川柳でファインチューニングしたGPT-2 (“Fine-tuned GPT-2”と呼ぶ)をベースラインとする。学習時の入力は“次のお題に対する川柳を生成してください。お題: {}, 川柳: {}”とする。{}内にはお題や川柳が入る。

モデル 4.1節で述べた手法の、ステップごとに構築したモデルについて評価を行う。具体的には次のモデルについて実験を行う。

- (a) Poelm-ja (Structure-aware training)
- (b) SenryuGPT (ファインチューニング)
- (c) SenryuGPT-ppo (強化学習)
- (d) Fine-tuned GPT-2 (ベースライン)

以降、モデルは(a), (b), (c), (d)で表すこととする。

実験の詳細 4.1.2節と4.1.3節の学習時に使用したお題 (seen) のうち50個と、使用していないお題 (unseen) 50個の合計100個のお題について10句ずつ生成し、それぞれのモデルに対して1,000句を自動評価の対象とする。(a), (b), (c)の生成時は入力を“<PREF><TOPIC: {お題}><LEN: 5><LEN: 7><LEN: 5></PREF>”とする。サンプリングにより生成する際のtop_pの値は0.9とする。人手評価では、このうちseenとunseenから25個ずつお題を選びその中から2句ずつ選択して合計100句を評価対象とする。

音数の自動評価では各テキストをKWJAによる形態素解析を行った後に、各形態素の読みから音数を取得する。おもしろさの自動評価では4.1.3節で使用した報酬モデルを評価モデルとして使用する。川柳としての適否の自動評価は行わない。

4.2.2 音数の評価

図2にクラウドソーシングによる音数の人手評価結果を示す。図3にKWJAを用いた音数の自動評価結果を示す。(a)は音数の制御を学習しているが、サンプリングにより音数の乱れ

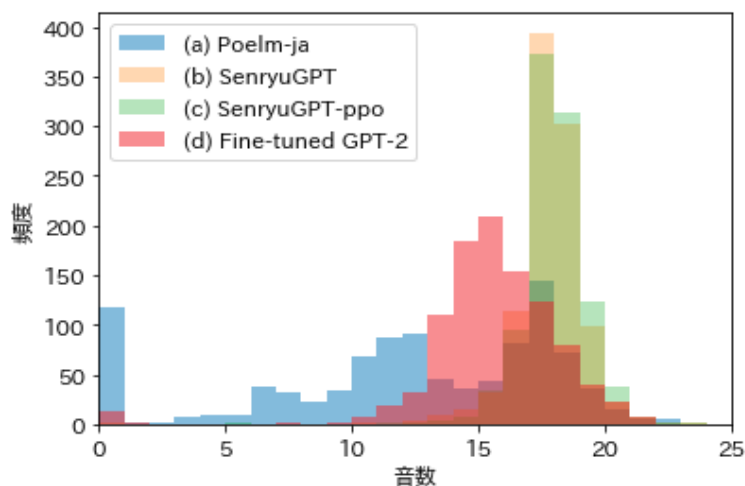


図 3: KWJA による音数の評価

表 9: クラウドソーシングによる川柳としての適否の評価

モデル	カウント
(a) Poelm-ja	4
(b) SenryuGPT	70
(c) SenryuGPT-ppo	71
(d) Fine-tuned GPT-2	18

た生成が多くあることがわかった。一方、川柳を使用して学習した後のモデルはある程度の制御ができており、(d)では17音となっている生成が最も多くなった。ベースラインは15音あたりの生成が多くなり、音数の学習はできていない。また、強化学習が音数の制御に及ぼす影響も見られなかった。自動評価においても、人手評価と類似した結果が得られた。

4.2.3 川柳としての適否の評価

表9には生成の川柳としての適否をクラウドソーシングにより評価した結果を示す。5人の評価者のうち3人以上が「川柳である」と回答したものを川柳とした。(a)は川柳の体裁のみを学習しているため、川柳であるとされた生成は少ないが、川柳で学習したモデルがベースラインと比べ圧倒的に川柳と認められることがわかった。

4.2.4 おもしろさの評価

表10に川柳のおもしろさをクラウドソーシングで評価した結果を示す。10人の評価者の最頻値をその句のスコアとした結果を示している。表11には4.1.3節で使用した報酬モデルによる自動評価結果を示す。強化学習ではこの報酬モデルの報酬を大きくすることを目的として学習しているため(c)の値が突出しているが、強化学習前のモデルがベースラインよりおもしろい川柳を生成できていることがわかる。また、人手評価の結果から、強化学習により生成川柳がおもしろくなっていることが確認できた。

表 10: クラウドソーシングによるおもしろい川柳かの評価

モデル	スコア (最頻値)
(b) SenryuGPT	2.202
(c) SenryuGPT-ppo	2.390
(d) Fine-tuned GPT-2	1.701

表 11: 報酬モデルによる川柳の評価

モデル	スコア (seen)	スコア (unseen)
(a) Poelm-ja	0.0434	0.0370
(b) SenryuGPT	0.0867	0.0694
(c) SenryuGPT-ppo	0.1550	0.1227
(d) Fine-tuned GPT-2	0.0641	0.0493

4.2.5 生成例

以下に各モデルの生成例を示し、講評を行う。“【】”内はその文のトピックやお題を表す。

1. 【イベント】その他には ミニマラソンが開かれる (a)
2. 【わくわく】うなぎの塔私も挑戦したが (d)
3. 【スポーツ】故障したオレはアイススケートで良い (b)
4. 【スポーツ】義母からの口撃はいつも ブレーキ無し (c)
5. 【マイペース】スマホゲームに没頭し過ぎて飛び起きた (c)

1, 2 は音数としては 17 音を守っているが川柳と認められなかった例である。(d) は内容が支離滅裂になってしまう場合が多く見られた。3 と 4 の比較では強化学習によるおもしろさ向上が確認できた。ただ、5 のように最終的なモデルでも川柳と呼べない生成があることも判明した。

5 おわりに

本研究では、機械学習を用いて川柳という文芸の理解と生成に取り組んだ。まず川柳の理解では、ベースラインの川柳の直接的な総合評価を予測するアプローチに対し、提案手法として、川柳を相対的にわかりやすい指標に分割し、それぞれの項目の予測を組み合わせて総合的な評価を予測する手法を提案した。提案手法は、ベースラインに比べて若干の精度向上が見られたが、今後の研究では川柳の面白さの理解に向けてさらなる精度向上が必要である。また、構築した川柳理解モデルを強化学習に応用することで、end-to-end でおもしろい川柳を生成するモデルを構築した。学習を 3 つのステップに分けて段階的に川柳の特徴を学習した。実験の結果、提案手法は音数、川柳としての適否、おもしろさの 3 つの項目すべてにおいてベースラインを上回る結果となったが、人間の作品とは依然として大きな乖離があり、さらなる研究の余地がある。

謝辞

川柳データを提供いただいた川柳投稿サイト「まるせん」に感謝する。また、本研究を行う上で、いくつものアドバイスをいただいた河原教授をはじめ、研究のアイデアや結果の分析において貢献いただいた京都大学の野村理朗准教授に感謝申し上げます。

参考文献

- [Bowman 15] Bowman, S. R., Vilnis, L., Vinyals, O., Dai, A. M., Jozefowicz, R., and Bengio, S., “Generating Sentences from a Continuous Space” (2015)
- [Christiano 17] Christiano, P. F., Leike, J., Brown, T., Martic, M., Legg, S., and Amodei, D., “Deep Reinforcement Learning from Human Preferences”, in Guyon, I., Luxburg, U. V., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., and Garnett, R. eds., *Advances in Neural Information Processing Systems*, Vol. 30, Curran Associates, Inc. (2017)
- [Devlin 19] Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K., “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding”, in *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pp. 4171–4186, Minneapolis, Minnesota (2019), Association for Computational Linguistics
- [Hitsuwari 23] Hitsuwari, J., Ueda, Y., Yun, W., and Nomura, M., “Does human–AI collaboration lead to more creative art? Aesthetic evaluation of human-made and AI-generated haiku poetry”, *Computers in Human Behavior*, Vol. 139, p. 107502 (2023)
- [Ke 17] Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., and Liu, T.-Y., “Lightgbm: A highly efficient gradient boosting decision tree”, *Advances in neural information processing systems*, Vol. 30, pp. 3146–3154 (2017)
- [Liu 19] Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., and Stoyanov, V., “RoBERTa: A Robustly Optimized BERT Pretraining Approach” (2019)
- [Ormazabal 22] Ormazabal, A., Artetxe, M., Agirrezabal, M., Soroa, A., and Agirre, E., “PoeLM: A Meter- and Rhyme-Controllable Language Model for Unsupervised Poetry Generation”, in Goldberg, Y., Kozareva, Z., and Zhang, Y. eds., *Findings of the Association for Computational Linguistics: EMNLP 2022*, pp. 3655–3670, Abu Dhabi, United Arab Emirates (2022), Association for Computational Linguistics
- [Radford 18] Radford, A., Narasimhan, K., Salimans, T., Sutskever, I., et al., “Improving language understanding by generative pre-training” (2018)
- [Radford 19] Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., and Sutskever, I., “Language Models are Unsupervised Multitask Learners” (2019)

- [Rombach 21] Rombach, R., Blattmann, A., Lorenz, D., Esser, P., and Ommer, B., “High-Resolution Image Synthesis with Latent Diffusion Models” (2021)
- [Stiennon 20] Stiennon, N., Ouyang, L., Wu, J., Ziegler, D., Lowe, R., Voss, C., Radford, A., Amodei, D., and Christiano, P. F., “Learning to summarize with human feedback”, in Larochelle, H., Ranzato, M., Hadsell, R., Balcan, M., and Lin, H. eds., *Advances in Neural Information Processing Systems*, Vol. 33, pp. 3008–3021, Curran Associates, Inc. (2020)
- [Touvron 23] Touvron, H., Martin, L., Stone, K., Albert, P., Almahairi, A., Babaei, Y., Bashlykov, N., Batra, S., Bhargava, P., Bhosale, S., Bikel, D., Blecher, L., Ferrer, C. C., Chen, M., Cucurull, G., Esiobu, D., Fernandes, J., Fu, J., Fu, W., Fuller, B., Gao, C., Goswami, V., Goyal, N., Hartshorn, A., Hosseini, S., Hou, R., Inan, H., Kardas, M., Kerkez, V., Khabsa, M., Kloumann, I., Korenev, A., Koura, P. S., Lachaux, M.-A., Lavril, T., Lee, J., Liskovich, D., Lu, Y., Mao, Y., Martinet, X., Mihaylov, T., Mishra, P., Molybog, I., Nie, Y., Poulton, A., Reizenstein, J., Rungta, R., Saladi, K., Schelten, A., Silva, R., Smith, E. M., Subramanian, R., Tan, X. E., Tang, B., Taylor, R., Williams, A., Kuan, J. X., Xu, P., Yan, Z., Zarov, I., Zhang, Y., Fan, A., Kambadur, M., Narang, S., Rodriguez, A., Stojnic, R., Edunov, S., and Scialom, T., “Llama 2: Open Foundation and Fine-Tuned Chat Models” (2023)
- [Vaswani 17] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L. u., and Polosukhin, I., “Attention is All you Need”, in Guyon, I., Luxburg, U. V., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., and Garnett, R. eds., *Advances in Neural Information Processing Systems*, Vol. 30, Curran Associates, Inc. (2017)
- [Wei 22] Wei, J., Bosma, M., Zhao, V., Guu, K., Yu, A. W., Lester, B., Du, N., Dai, A. M., and Le, Q. V., “Finetuned Language Models are Zero-Shot Learners”, in *International Conference on Learning Representations* (2022)
- [植田 22] 植田 暢大, 大村 和正, 児玉 貴志, 清丸 寛一, 村脇 有吾, 河原 大輔, 黒橋 禎夫, “KWJA : 汎用言語モデルに基づく日本語解析器”, 第 253 回自然言語処理研究会, 京都 (2022)
- [新家 11] 新家完司 : 川柳の理論と実践, 新葉館出版 (2011)
- [川村 21] 川村 秀憲, 山下 倫央, 横山 想一郎 : 人工知能が俳句を詠む AI 一茶くんの挑戦, オーム社 (2021)
- [中川 19] 中川裕貴, 村脇有吾, 河原大輔, 黒橋禎夫, “クラウドソーシングによる大喜利の面白さの構成要素の分析”, 言語処理学会年次大会発表論文集, Vol. 25, pp. ROMBUNNO.B3–2 (WEB ONLY) (2019)
- [野林 14] 野林正路 : 詩・川柳・俳句のテキスト分析, 和泉書院 (2014)

Appendix

A 学習に使用したモデル

3.1 節の実験 1,2 で使用した BERT-like モデルの詳細を表 12 に示す.

表 12: 学習に使用した事前学習モデルの詳細

model	corpus	#params	#dimension	vocabulary s
BERT _{BASE} (cl-tohoku/bert-base-japanese-whole-word-masking)	Wikipedia	110M	768	32,000
RoBERTa _{BASE} (nlp-waseda/roberta-base-japanese-with-auto-jumanpp)	Wikipedia & CC-100	125M	768	32,000

B Wikipedia データセットの処理

Structure-aware training 用データセットの構築は Wikipedia データセットを基に行う. テキストは段落の一文目を使用し, 丸括弧で囲まれている部分を削除する. また, Juman++ では英数字の音数が正確に取得できないため英数字を含む文を除外する. そして, 長 (短) すぎる文は除外する.

C 川柳データの前処理

4.1.2 節で使用する川柳データの前処理について述べる. 「まるせん」より提供された利用可能な川柳データ約 58 万句の内, 約 65% は 5-7-5 に分かたれていない. これらの川柳は structure-aware training では適用することができないため, ファインチューニングに使用することができない. この問題を解消するため, あらかじめ 5-7-5 に分かたれている川柳を用いて, 川柳 5-7-5 に分割するモデルを構築する. 5-7-5 の間の空白を削除した川柳をモデルの入力とし, 5-7-5 の間に空白を含んだ川柳を正解として学習を行う. 学習に使用したモデルを次に示す.

- rinna/japanese-GPT-2-medium
- rinna/japanese-gpt-neox-3.6b
- megagonlabs/t5-base-japanese-web

GPT 系のモデルでは学習時の入力を “分かたれていない川柳<\s>分かたれた川柳” とし, T5 では分かたれていない川柳の入力に対し, 分かたれた川柳をラベルとして学習を行った. 3 つのモデルの推論結果から多数決を取り, 多数派の分割を 5-7-5 の正解と仮定して学習データとする.

D 川柳嗜好データセットの詳細

4.1.3 節のデータセット構築時に設定した“margin”の設定の詳細を表 13 に示す。得票数の差に応じて設定している。

表 13: 得票数の差による margin のアノテーション

比率	margin
10:0	3
9:1	2
8:2	2
7:3	1
6:4	0
5:5	0

E ハイパーパラメータ

表 14 に、3.1.1 節の実験で使用したハイパーパラメータを示す。また、表 15 には、3.1 節の実験 1, 2 で使用したハイパーパラメータを示す。表 16 に示すのは、3.1 節の実験 3 で使用したハイパーパラメータである。表 17, 18, 19 に、4.1 節の学習で使用したハイパーパラメータを示す。表内の英字は次のモデルに対応している。

- (a) Poelm-ja
- (b) SenryuGPT
- (c) SenryuGPT-ppo
- (d) Fine-tuned GPT-2
- (e) 報酬モデル（強化学習時）

表 14: RoBERTa の追加学習に使用したハイパーパラメータ

hyperparameter	
learning rate	1e-5
batch size	16
#epoch	50
#warm up steps	50000

表 15: BERT-like モデルのファインチューニングに使用したハイパーパラメータ

hyperparameter	
learning rate	5e-5
batch size	8
#epoch	15
#warm up steps	500

表 16: LightGBM の学習に使用したハイパーパラメータ

hyperparameter	
metric	l2 (mse)
#leaves	31
learning rate	0.01
feature fraction	0.7
bagging fraction	0.8
bagging frequency	5
min child samples	20
#epoch	15

表 17: CLM 学習に使用したハイパーパラメータ

ハイパーパラメータ	(a)	(b)	(d)
learning rate	6e-5	5e-5	5e-5
batch size	128	256	16
#epoch	5	5	2
weight decay	0.2	0	0

表 18: PPO 学習に使用したハイパーパラメータ

ハイパーパラメータ	(c)
learning rate	5e-5
batch size	32
#epoch	5
#ppo epoch	3
clip range	0.1
clipping value	0.2

表 19: 報酬モデルの学習に使用したハイパーパラメータ

ハイパーパラメータ	(e)
learning rate	6e-5
batch size	16
gradient accumulation steps	3
#step	5
weight decay	0.5