

学習データの珍しさを考慮した多様な広告文生成

黒木 開¹ 石塚 湖太¹ 川上 孝介¹¹negocia 株式会社

{kai_kurogi, kota_ishizuka, kosuke_kawakami}@negocia.jp

概要

リスティング広告は、様々なユーザーにリーチするために、多様な表現や訴求を用いた広告文を作成することが重要である。しかし従来の広告文生成モデルは、トレードオフの関係にある多様性と妥当性を両立することが難しい課題があった。そこで我々は、学習データに含まれる各広告文の表現の珍しさ(レア度)を計算し、レア度に応じた重み付けを行うことで、妥当性を維持しながら多様な広告文を生成する手法を提案した。結果として、既存手法以上の妥当性を獲得しながら、多様性を29%向上させることができた。

1 はじめに

リスティング広告とは、ユーザーの検索キーワードに関連した広告がリスト形式で表示される広告である。リスティング広告の入稿形式の1つであるレスポンス検索広告(RSA)は、1つの広告に複数の広告見出し、説明文を入稿し、配信媒体側でその組み合わせを自動で調整することで、ユーザーごとに最適な広告文を表示できる広告である。例えばGoogleやYahoo!のRSAでは、最大15個の広告見出し、及び最大4個の説明文を一つのRSAに入れることができる。RSAの運用では、様々なユーザーにリーチするためにその入稿枠を全て埋めることが推奨されている。しかし、広告ごとに大量のアセットを考える業務負荷は非常に大きい。そのため、制作現場では多様な広告文を生成出来る技術が求められている。

この課題に対して川本ら[1]は、変分AutoEncoderにより広告文の潜在空間を獲得し、品詞系列で条件づけることで、多様な広告文を生成する手法を提案している。しかし上記手法では、多様性が向上する代わりに、妥当性や流暢性が低下するトレードオフの問題が確認された。また、近年ではTransformer[2]に代表される事前学習済み汎用言語モデルを活用

することで、高い流暢性を持った広告文自動生成技術が数多く提案されている。脇本ら[3]や黒木ら[4]は大規模言語モデルの入力にキーワードを利用することで、大規模言語モデルが元々備える流暢性と、キーワードによる方向付による妥当性の両立に成功した。しかし上記手法は、多様性向上を目的としていないため、何度生成しても同じような広告文ばかりが生成されてしまうという問題があった。そこで本研究では、高い妥当性を維持しながら、既存の手法よりも多様性を高くするモデルの学習方法を提案する。具体的には学習データに含まれる各広告文の表現の珍しさ(レア度)を事前に計算し、そのレア度に応じて学習データの重み付けを行った。レア度で重み付けを行うことで、モデルは珍しい表現が使われた広告文を重視して学習するため、生成時に多様性が向上することが期待できる。レア度の計算にはPerplexityを用いた。また、モデルの評価にはPairwise-BLEU[5], local-distinct_n(n=1, 2)[1], recall-BLEU[6]を用いた。

2 提案手法

本研究では、事前にて学習データの各広告文のレア度をPerplexityによって計算し、モデル学習時にその値に応じた重み付けを行うことで、様々な表現が出やすい広告文生成モデルを作成した。

2.1 Perplexity

Perplexityは言語モデルの複雑さを示す指標であり、以下の式で計算される。

$$\text{Perplexity} = \prod_{c=1}^N P(c)^{-\frac{1}{N}} \quad (1)$$

ただし $P(c)$ はトークン c の発生確率である。これはモデル分布が正解トークンを選択する難しさを示しており、Perplexityが小さいモデルは、より正解トークンを選ぶ確率が高い、すなわちより流暢なモデルであると考えられている。ただし広告文生成タスクにおいては、一つの入力に対して、複数の出

力が正解となることが一般的である。このような場合、Perplexity は「レア度」と解釈することができる。すなわち、ある広告文の Perplexity を計算した際、その値が小さい場合は、対象の広告文にはよく見る表現が多く含まれていることを示している。逆に Perplexity が大きい場合、その広告文には珍しい表現が多く使われていると考えることができる。本研究では、事前に準備した非常に軽量の学習済み言語モデルを用いて、データセットの個々の広告文に対して Perplexity を計算した。

2.2 Perplexity による重み付け学習

データセット全体の広告文を $T_m = \{T_1, T_2, \dots, T_m\}$ とする。各広告文がトークンの長さ = N_i のトークン $T_i = (t_{i1}, t_{i2}, \dots, t_{iN})$ で表される時、 T_i の Perplexity = $weight_i$ を用いて、以下の値を最大化するようモデルのパラメータを学習させた。

$$\sum_{T_i \in T_m} weight_i * \sum_{1 \leq n \leq N} (\log p(t_{in} | input, t_{i < n})) \quad (2)$$

ただし $t_{i < n} := (t_{i1}, t_{i2}, \dots, t_{i(n-1)})$ である。すなわちモデルは、より Perplexity が高い広告文を重視するように学習した。

3 実験

本実験では、Perplexity の重みづけによる多様性の変化を観察するため、以下で指定されたモデルに対し、Perplexity による重みづけのあり (reference)、なし (baseline) の 2 パターンでモデルを学習し、それぞれの生成文の品質を比較した。

3.1 モデル

黒木ら [4] の手法を参考に、言語モデルとして T5[7] を採用した。また、モデルの入力は、キーワード、及び広告の遷移先 URL であるランディングページ (LP) の HTML から取得したタグ情報を連結した文字列を用いた。埋め込み層には、HTML のタグ情報による Tag Embeddings を追加した。つまり、本モデルは一つのキーワードと LP のペアから広告文 (広告見出し、及び説明文) を生成するモデルとなる。

3.2 データセット

学習に用いるデータセットは独自にスクレイピングすることで得られたデータを用いた。各データは検索キーワード、キーワードによって表示された広

表 1 評価結果 (広告見出し)

	Pairwise-BLEU	local-distinct _n		recall-BLEU
		n=1	n=2	
reference	0.127	0.232	0.377	0.401
baseline	0.164	0.177	0.291	0.363

表 2 評価結果 (説明文)

	Pairwise-BLEU	local-distinct _n		recall-BLEU
		n=1	n=2	
reference	0.216	0.210	0.373	0.446
baseline	0.244	0.216	0.358	0.412

告文、LP の HTML から構成されており、学習データ 741,506 件とテスト用データ 82,320 件に分割して使用した。

3.3 評価方法

モデルの性能は自社で保有する、RSA の配信データ 200 件を用いて評価した。評価に用いたデータの全体図を図 1 に示す。各配信データは RSA に入稿されているアセット群、RSA に紐づく LP の HTML、及び RSA が紐づく広告グループに登録されたキーワード群からなる。ただし本モデルは一つのキーワード、LP のペアから広告文を生成するモデルである。そこで本実験では、キーワード群の中からコストの高いキーワードを最大 10 個まで抜き出し、それぞれのキーワードと LP のペアで広告文を生成することで、LP ごとに生成結果を取得した。また、定量評価は広告見出し、説明文ごとに行い、評価指標には Pairwise-BLEU, local-distinct_n(n=1,2), recall-BLEU を用いた。

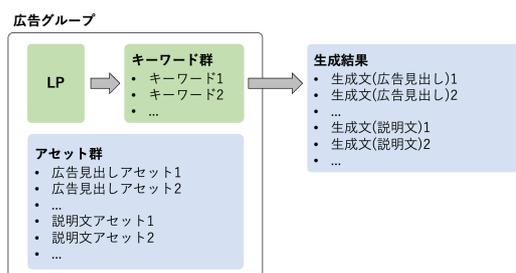


図 1 広告グループ、アセット群、LP、キーワード群、生成結果の関係の概要図

3.3.1 Pairwise-BLEU

Pairwise-BLEU は Shen ら [5] らによって提案された多様性を測る指標であり、生成結果とアセット群間の BLEU ではなく、生成結果間の BLEU を計算した値である。ある一つの生成結果に対し、それを除く全ての生成結果を参照文と見做すことで BLEU を計

表3 各モデルの生成例（広告見出し）

キーワード	エステ体験 痩身 セルライト 痩せる	体験結果 コース
reference	【痩身エステ】夏までに痩せたい 結果を求める女性へ朗報 美脚エステ体験コース一挙紹介 人気フェイシャルエステ TOP5	痩身エステ体験コース口コミは？ 気になるお腹の脂肪には？ セルライト撃退エステ TOP10 夏ダイエット成功への道とは？
baseline	【痩身エステ】夏までに痩せたい 人気結果を求める女性へ朗報 サロン体験予約でセルライト撃退 最新/本気で痩せるエステ TOP5	完全痩せ体験コース一挙紹介 気になる部位から絶対行きたい 施術体験予約でセルライト撃退 今年の夏こそ本気で痩せたいなら

算する。生成結果が多様なほど低い値を取る。

3.3.2 local-distinct-n

local-distinct-n は川本ら [1] によって提案された多様性を測る指標であり、生成結果全体の n-gram のうち、ユニークな n-gram の割合を元に計算した値である。生成結果が多様なほど高い値を取る。また、Pairwise-BLEU とは異なり、同じテキストの中にユニークな n-gram が多い場合も高い値を取る。本実験では n=1,2 とした。

3.3.3 recall-BLEU

recall-BLEU は Hui ら [6] によって提案された妥当性を示す指標であり、実際に入稿されていたアセット 1 つ 1 つに対して、生成結果を参照文と見做して BLEU を計算し、平均を取った値である。生成結果の妥当性が高いほど高い値を取る。recall-BLEU が小さい値となった場合、元の RSA のアセットに含まれていない文ばかりを生成している、すなわち妥当性の低いモデルであると判断することができる。

4 結果と考察

4.1 定量評価

定量評価の結果を表 1、表 2 に示す。多様性を示す Pairwise-BLEU, local-distinct_n(n=1, 2) については、説明文の local-distinct₁ を除く全ての指標で reference が baseline を上回った。特に、広告見出しにおける Pairwise-BLEU では、reference が baseline より約 29% 高い結果が得られた。また、妥当性を示す recall-BLEU においても、広告見出し、説明文のいずれの場合も reference が baseline を上回った。つまり提案手法は、多様性と妥当性を同時に上げられる手法だといえる。これは、重み付けにより、レア度が高く、かつ流暢・妥当性があるトークンと、そもそも文法的にありえないトークンの両者を区別できる

ようになったからだと考えられる。重み付けを行うことで、前者のトークンのみ出力確率を引き上げることができたため、妥当性、多様性が同時に向上したと思われる。

4.2 定性評価

定性評価のためのキーワードと広告見出しの生成例を表 3 に示す。LP はエステサロンの比較サイトを用いた。いずれのモデルも、個々の広告文を見ると広告らしい文が生成できていることが確認できる。reference では「痩身エステ」などの定番の表現に加え、キーワードには含まれないものの、「フェイシャル」、「美脚」などエステの定番部位を含んだものや、「気になるお腹の脂肪には？」といった問いかけ表現を生成することができた。一方 baseline では「痩身エステ」などの定番表現や、「セルライト」などキーワードに含まれる表現を含んだ広告文は生成できたものの、キーワードから連想できるが、キーワード自身には含まれていないその他表現はあまり生成することができなかつた。つまり、もし制作現場で全てのキーワードが網羅できている場合、両モデルに大きな差は出ないが、全てのキーワードを網羅することが難しい場合には、reference モデルを用いた方が、より多様な表現を獲得できると言える。

5 まとめ

本研究では、広告文生成タスクについて、データセットに Perplexity による重みづけを行う手法を提案した。実験の結果、提案手法は既存手法以上の妥当性を持ちながら、多様性を約 29% 上げることができた。本手法は loss の計算時に weight で重み付けするだけで実現できるため、既存の様々な手法に組み込むことが可能である。一方、本モデルは多様な表現の広告文を生成出来るものの、その広告文を実際に配信した時に高い広告効果が得られるかどうかは考慮していない。そこで今後は、多様性を維持しな

がら、より広告効果が高い広告文を生成する手法について研究したい。

参考文献

- [1] 川本峻頌, 張培楠. スタイル制御を考慮した多様な広告文生成. 言語処理学会第 25 回年次大会, 2019.
- [2] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, u. Kaiser, and I. Polosukhin. Attention is all you need. **Proceedings of the 31st International Conference on Neural Information Processing Systems**, pp. 6000–6010, 2017.
- [3] 脇本宏平, 川本峻頌, 張培楠. インターネット広告におけるキーワードに基づく広告文の自動生成. 2020 年度人工 知能学会全国大会, 2020.
- [4] 黒木開, 石塚湖太, 川上孝介, 中田和秀. 複数の指定語句を必ず含むリスティング広告の広告文自動生成. 言語処理学会第 28 回年次大会, 2022.
- [5] T. Shen, M. Ott, M. Auli, and M. Ranzato. Mixture models for diverse machine translation: Tricks of the trade. **Proceedings of the 36th International Conference on Machine Learning**, pp. 97:5719–5728, 2019.
- [6] C. Hui, H. Zhao, Y. Song, and B. Lu. An empirical study on development set selection strategy for machine translation learning. **Proceedings of the Joint Fifth Workshop on Statistical Machine Translation and MetricsMATR**, p. 67–71, 2010.
- [7] C. Raffel, N. Shazeer, A. Roberts, K. Lee, S. Narang, M. Matena, Zhou Y., W. Li, and J. Liu. P. Exploring the limits of transfer learning with a unified text-to-text transformer. **Journal of Machine Learning Research**, p. 21(140):1–67., 2020.