

# 広告文生成タスクの規定とベンチマーク構築

三田 雅人<sup>1</sup> 村上 聡一郎<sup>1</sup> 張 培楠<sup>1</sup>

<sup>1</sup> 株式会社サイバーエージェント

{mita\_masato,murakami\_soichiro,zhang\_peinan}@cyberagent.co.jp

## 概要

人手による広告文制作の限界を背景に、広告文自動生成に関する研究が盛んに行われている。しかし、分野全体で享受できるベンチマークが存在しないことに加え、モデルの入出力を含め問題設定として十分に既定されていないため、既存研究は手法間の比較が困難な状況であった。そこで本研究では、広告文生成タスクの既定およびベンチマーク構築を行う。具体的には、まず広告文生成を応用横断的なタスクとして規定したうえで、広告文生成における初のベンチマーク **CA Multimodal Evaluation for Ad Text GeneRAtion (CAMERA<sup>2</sup>)** を構築する。さらに、構築したベンチマークを用いた評価実験を通して、広告文生成の現状と今後の課題を示す。

## 1 はじめに

インターネット広告の需要拡大に伴い、人手による広告制作は限界を迎えつつある。特に、検索クエリに関連する広告文を検索結果画面に提示する検索連動型広告(図1)は、前年比約8%前後で増加している膨大な検索クエリの対応に迫られているため、広告文制作の自動化が喫緊の課題となっている[1]。

このような産業的な需要もあり、広告文の自動生成に関する研究が盛んに取り組まれている[2]。事前に定義したテンプレートに対して適切なキーワードを挿入することで広告文を生成するテンプレートに基づく手法[3-5]に始まり、近年では機械翻訳や自動要約などで活用されているエンコーダ・デコーダ型モデルを用いたニューラル言語生成に基づく手法[6-8]など数多くの手法が提案されている。

一方で、広告文生成研究の評価には重大な課題が残っている。英語や日本語など言語問わず分野全体が享受できる共通データセット(ベンチマーク)の欠如により、既存研究はこれまで独自データを用いた個々の検証に留まっており、手法間の包括的な比較ができていない[2]。また、ベンチマークが存在

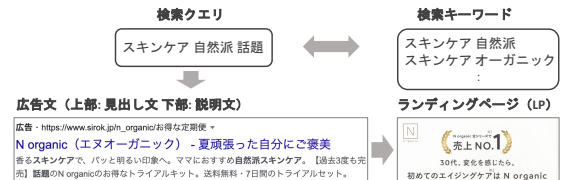


図1: 検索連動型広告の例。

しないゆえにモデルの入出力形式に関しても合意形成が取られていない。検索キーワードを入力に用いる研究[9]もある一方で、既存の広告文[7]や広告文をクリックした遷移先のランディングページ(LP)などを入力に用いる研究[8]も存在する。このことは、分野全体としてこれまで広告文生成の問題設定を十分に規定できていないことを示唆しており、ベンチマークの欠如と併せて広告文生成技術の一般化の大きな妨げとなっている。ベンチマークの設計を通して問題を規定し、研究者間で共有できれば、手法を横並びに評価できる。その結果として、分野全体が取り組むべき本質的に重要な部分が明らかになり、技術の一般化が進むことが期待される。

そこで本研究では、広告文生成の学術分野としてのさらなる発展を目指し、広告文生成タスクの規定およびベンチマークの構築を行う。具体的には、まず広告の目的および要件を整理し、**広告文生成 (Ad Text Generation)** を応用横断的なタスクとして規定する。このとき、従来タスクとの違いや本タスク特有の課題を提示するとともに、研究題材としての学術的な意義について示す。次に、本研究では日本語における広告文生成ベンチマークのための共通データセット **CA Multimodal Evaluation for Ad Text GeneRAtion (CAMERA<sup>2</sup>)** を構築する<sup>1)</sup>。最後に、事前学習済みモデルの種類やマルチモーダル情報の使用有無といった観点で多様な複数の広告文生成モデルを用意し、構築したベンチマークを用いたベースライン実験を通して、提案ベンチマークの有用性および本タスクの現状と今後の課題を示す。

1) 構築したデータセットは今後一般公開する予定である。

## 2 広告文生成タスクの規定

### 2.1 広告文の要件

広告の目的は、ある商材（取引の対象となる物品やサービス）があったとき、広告の受け手である消費者（ユーザ）の態度や行動に影響を与えることである。したがって、広告文生成の目的は、ユーザの購買行動を促す文を生成することになる。ここで、広告文とはどのような要件を満たしていれば広告文として成立するのだろうか。本研究では、タスクを既定するにあたり、まず広告文の基本要件として次の2つを定めた：(1) 商材内容と整合がとれている；(2) 商材内容からユーザの潜在的なニーズに応じて情報が取捨選択されている。

要件1は、広告制作現場で実運用するうえでは必要条件ともいえる。また、より具体的には2つに大別でき、(a) 広告文が入力情報である商材内容に含意される場合は、**忠実性** [10] の観点から整合がとれており、(b) 入力情報に含意されない情報を含むが世界知識や常識的知識などの外部知識に基づいている場合は、**事実性** [11] の観点から整合がとれている。要件2について、広告では限られたスペースで商材に関する特徴や魅力を短い文で表し、即座にユーザの関心を惹き付けることも同時に要求される。このことから、広告文は入力情報から**ユーザに訴求できる情報を取捨選択（情報圧縮）**されている必要がある。

### 2.2 広告文生成タスク

2.1 節で整理した広告の目的および要件を踏まえ、広告文生成タスクを次のように定義する：

#### タスク定義

商材に関する説明文書を  $x$ 、ユーザの潜在的なニーズや関心を反映するユーザ信号を  $a$ 、広告文を  $y$  とする。このとき、広告文生成の目的は  $p(y|a, x)$  をモデル化することである。

広告文生成タスクは、入力文書の内容との整合を保持しながら情報圧縮を行う点で、従来の文書要約タスクと関連が深く、特に文書要約の亜種であるクエリ重視要約（query-focused summarization） [12] とは入力にユーザからのクエリが用いられるために、問題設定として最も近いといえる。しかし、クエリ重視要約ではユーザが明示的に指定した特定のクエ

リ（**顕在化されたニーズ**）に対応する要約を返すのに対し、広告文生成ではユーザの**潜在的なニーズ**に対応する必要がある点で異なる。ここで、広告文生成は広範囲な応用先を想定したタスクであるため、ユーザ信号  $a$  は応用先、すなわち広告の形態に依存して変動することに注意されたい。例えば、検索連動型広告の場合はユーザ信号  $a$  には、検索クエリやキーワードが考えられるが、ディスプレイ広告<sup>2)</sup>やEC上の商品広告であれば、ユーザの購買履歴なども考えられる。

広告文生成特有の技術的な課題の一つは、このようなユーザ信号をもとにユーザの潜在的なニーズを捉え、広告効果に繋がるような訴求性のある文を生成できるかであるが、広告効果は受け手であるユーザの心理的特性に大きく依存する。そのため、より高度な広告文生成の実現には、（究極的には）認知心理学や社会心理学を基礎とした広告心理学 [13] との接続も必要になってくるだろう。このように、広告文生成タスクは自然言語処理技術をよりユーザ中心にするためのよい題材の一つといえる。

## 3 広告文生成ベンチマークの構築

### 3.1 設計

本研究ではベンチマーク設計方針についてまず次の2つを定めた：(1) マルチモーダル情報が活用できる；(2) 業種別に評価できる。1つ目の方針について、検索連動型広告を含め多くの広告形態では商材の特徴や訴求ポイントをユーザにより直感的に伝えるために、テキストだけでなくLPなどの視覚情報を併せて用いている。また、LPと関連性の高い内容をユーザへ訴求することが、クリックを促す重要な役割を果たすと知られている。実際、広告文生成分野においてもLPのレイアウト情報などを考慮することでモデルの生成品質が向上したという報告もある [14]。このように、広告文生成タスクにおいて、今後マルチモーダル情報の効果的な活用を探究することは重要である。

2つ目の方針について、効果的な広告文を作成するためには具体的な訴求を入れることが重要であることが知られている。例えば、価格や割引などからお得感を訴求するには「送料無料」、「ポイント高還元！」のような訴求表現が用いられ、サービスや

2) Web サイトやアプリの広告枠にテキスト形式やバナー形式で表示される広告。

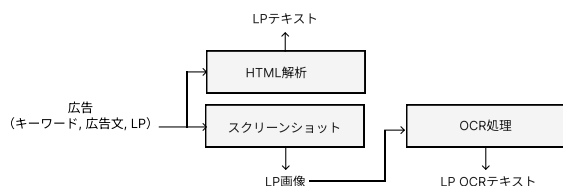


図 2: データセット構築の流れ。

企業の実績を訴求したい場合は、「日本最大級の」、「売上 No.1」などの訴求表現が用いられる。ここで、どのような訴求表現（および訴求軸）を用いれば広告として効果的かは業種ごとに異なることが先行研究 [15] によって報告されている。そのため、より精緻な評価を実施し、頑健なモデルの開発を促進させるためにも業種別に評価できることは重要である。

### 3.2 データセット構築手順

データセット構築の流れを図 2 に示す。元データとして、Web 上にある検索連動型広告を用いた。広告データには「(検索) キーワード」、「広告文」、「LP」がそれぞれ付与されている。ここで、検索連動型広告は、広告見出しと説明文で構成される（図 1）。説明文では見出し文と比較して表示領域が広いので、訴求表現は含まれるものの一般的な自然文で書かれることも多い。一方で、見出し文は表示領域制限に加え、即時的にユーザーの関心を惹き付ける役割を担っているため、人間が容認できる範囲で故意に文法を崩したり圧縮するなど広告特有の言い回しが多く含まれる。例えば、あるサービスにおける無料相談に関して訴求を行う場合、「簡単 30 秒！〇〇で無料相談」のような広告特有の表現が用いられる。このような一見非文にも思える広告特有の表現は人間の理解を妨げるところか、注意を引く効果があることが広告心理学分野で報告されている [16]。本研究では、より広告特有の言語現象に焦点をあてることのできるベンチマークとなるよう広告見出し文のみを広告文  $y$  として抽出した。

次に、広告に紐づく LP から HTML を解析し、Web ページの meta description をその商材に関する説明文書（LP テキスト） $x$  として抽出した。また、設計方針 1 を踏まえ、マルチモーダル情報を活用できるようにするため、LP 全体のスクリーンショット処理を行い、LP 画像を取得した。さらにその LP 画像を用い、Cloud Vision API の OCR 機能<sup>3)</sup>により、矩形領域集合  $R$  に対する画像  $I$ 、レイアウト  $C$ 、テキスト  $\{x_i^{\text{ocr}}\}_{i=1}^{|R|}$  を取得した。

3) <https://cloud.google.com/vision/docs/ocr>

表 1: データセット統計量。

	件数	参照広告文数	業種ラベル
訓練	12,395	1	
開発	3,098	1	
評価	869	4	✓

### 3.3 アノテーション

元データにはすでに運用実績のある広告文（参照広告文）が 1 つ付与されているが、同一商材に対する妥当な広告文は本質的に多岐にわたると考えられる。そのため、本研究では日本語母語話者、かつ広告アノテーションの専門家である 3 名のアノテータにより、追加で 3 つの参照広告文を作成する。評価アノテーション用のデータを収集する際は、設計方針 2 を考慮して、事前に人手で業種ラベルと対応付けられたキーワードを用いて無作為抽出した。なお、本研究で使用した業種ラベルは、人材、EC、金融、教育の 4 種類である。次に、収集した広告データに対して、3 名が独立的に編集し、各入力につき 3 つの参照を作成した<sup>4)</sup>。

本研究で構築したデータセットの基本統計量を表 1 に示す。本データセットは広告文生成のベンチマークを目的に開発しているため、予め訓練/開発/評価に分割しており、訓練（および開発セット）と評価セットの間でデータに重複がないようフィルタリング処理を行った。次に、データセットに含まれる入出力例を表 2 にそれぞれ示す。なお、本データセットには表 2 に記載の情報以外に、マルチモーダル情報（OCR データ・LP 画像）が付与されている。

## 4 実験

本節では、広告文生成の代表的なベースラインモデルを用意し、構築したベンチマークを用いた評価実験を通して、提案ベンチマークの有用性および広告文生成タスクの現状と今後の課題を示す。

### 4.1 実験設定

**ベースラインモデル** 広告文自動生成分野で現在主流であるエンコーダ・デコーダ型モデルに基づく広告文生成モデルを構築した。まず事前学習済みモデルの種類の影響を調査するために、2 つの異なるエンコーダ・デコーダ型の事前学習済みモデル（BART [17], T5 [18]）に基づきベースライ

4) アノテーションガイドラインの詳細は付録 A に示す。



表 2: CAMERA データセットの基本入出力（評価データセットにおける実例）.

LP テキスト (x)	キーワード (a)	広告文 (y)
カードローン診断.com 即日借りられるカードを 診断！今すぐお金を借り られるオススメ 3 社！	オリックス カードローン	1. 【公式】人気のカードローン 3 選 2. おすすめカードローンを簡単診断 3. 即日借りられるカードを診断！ 4. WEB なら融資まで最短 30 分

表 3: 実験結果（**太字の数値**は各列における最高値を表す）.

	全体			人材			EC			金融			教育		
Model	B-4	R-1	Kwd	B-4	R-1	Kwd	B-4	R-1	Kwd	B-4	R-1	Kwd	B-4	R-1	Kwd
BART	14.6	22.8	76.0	21.7	25.3	70.9	12.5	20.1	81.5	12.3	29.2	80.4	9.0	17.1	73.3
T5	13.6	25.6	<b>90.0</b>	18.9	23.9	<b>84.8</b>	12.9	28.6	93.6	12.2	33.5	94.7	6.4	18.5	<b>88.4</b>
+ {o}	17.8	<b>27.5</b>	85.6	23.4	<b>26.4</b>	82.3	15.9	<b>29.6</b>	87.3	18.6	<b>33.9</b>	88.5	10.2	<b>21.8</b>	85.3
+ {o, l}	<b>18.4</b>	25.7	84.4	<b>24.0</b>	25.9	81.4	<b>18.3</b>	26.5	87.3	17.5	31.9	86.1	<b>10.8</b>	19.9	83.7
+ {o, l, v}	16.3	26.0	84.5	19.0	25.2	82.7	17.9	27.5	86.7	<b>19.2</b>	33.0	91.4	8.0	19.7	78.9

ンモデルを構築した. さらに, マルチモーダル情報の有用性を調査するため, 村上ら [14] に従い, T5 に基づくモデルに対し, LP の画像情報を考慮した様々な設定を用意した. 具体的には, LP の OCR 処理済みテキスト (lp\_ocr;o), LP のレイアウト情報 (lp\_layout;l), LP の BBox 画像特徴量 (lp\_visual;v) の 3 種類のマルチモーダル情報を素性として用いた. その他の詳細な実験設定は付録 C に記載する.

**評価尺度** 生成文の品質を評価するために, 生成文と参照文の N-gram の一致率に基づく BLEU (B-4)<sup>5)</sup> [19] および ROUGE-1 (R-1) [20] を使用した. さらに, LP と広告文の関連性を評価するために, 指定キーワードが生成文に含まれる事例の割合を表すキーワード挿入率 (Kwd) [7] を用いた.

## 4.2 結果

表 3 に実験結果を示す<sup>6)</sup>. まず, 事前学習済みモデルの種類の影響を確認するために, BART および T5 の性能に着目する. 全体としては, B-4 では BART が, R-1 および Kwd では T5 が高い傾向にあることを確認した. 一方で, 業種別にみると, 人材では B-4 および R-1 の両方で BART が T5 を上回るなど結果にばらつきが生じている. そのため, 妥当な事前学習済みモデルの選択は一概に言えず, 評価の目的に応じて適切に選択する必要がある.

次に, マルチモーダル情報の有用性について確認する. 全体としては, OCR 処理済みテキスト (+

{o}) や LP のレイアウト情報 (+ {o, l}) を素性としてモデルに加えることで生成文の品質 (B-4 および R-1) が改善することを確認した. しかし, 業種別にみると, EC および金融における R-1 スコアにおいては LP の画像特徴量 (+ {o, l, v}) を素性として加えることで性能がかえって悪化している. 性能低下の要因として考えられるのは, LP 情報を活用することで一貫した性能向上を報告した先行研究 [14] では LP ファーストビュー<sup>7)</sup>のみ用いているのに対し, 本研究では LP 全体を用いているため, LP の画像情報の一部がノイズとして機能してしまった可能性が考えられる. 一方で, 参照広告文の中には, LP ファーストビューの情報だけでは生成が難しい場合も存在する<sup>8)</sup>. そのため, 今後の課題としては, LP 画像などマルチモーダル情報から重要な情報を適応的に取捨選択できるモデルの開発が期待される.

## 5 おわりに

本研究では, 広告文生成を応用横断的なタスクとして既定し, 初のベンチマークを構築した. さらに, 構築したベンチマークを用いた評価実験により, 広告文生成モデルの性能は業種毎に大きく変動すること, またマルチモーダル情報をいかに効果的に活用するかが今後の課題であることを示した. 広告文生成は自然言語処理の有望な応用先の一つとしてだけでなく, 自然言語処理技術をよりユーザ中心にするための重要かつ挑戦的な研究題材でもある. 本研究が整備した研究基盤により, 広告文生成の研究開発が今後さらに促進されることを期待する.

5) <https://github.com/mjpost/sacrebleu>

6) マルチリファレンス評価が一般的な BLEU のみ全 4 つの参照広告文を使用し, R-1 および Kwd は元データに付与された運用実績のある参照広告文 1 つをそれぞれ用いて評価した.

7) LP をスクロールせずに最初に見える範囲

8) 付録 B にその実例を記載する.

## 謝辞

本研究のアノテーションを実施して頂いた、株式会社シーエー・アドバンスのアノテータ諸氏に感謝します。

## 参考文献

- [1] Ogi Djuraskovic. Google search statistics and facts 2023 (you must know). Technical report, First Site Guide, 2022.
- [2] 村上聡一郎, 星野翔, 張培楠. 広告文自動生成に関する最近の研究動向. 人工知能学会全国大会論文集, Vol. JSAI2022, pp. 1P5GS601–1P5GS601, 2022.
- [3] Kevin Bartz, Cory Barr, and Adil Aijaz. Natural language generation for sponsored-search advertisements. In **Proceedings of the 9th ACM Conference on Electronic Commerce**, EC '08, p. 1–9. Association for Computing Machinery, 2008.
- [4] Atsushi Fujita, Katsuhiko Ikushima, Satoshi Sato, Ryo Kamite, Ko Ishiyama, and Osamu Tamachi. Automatic generation of listing ads by reusing promotional texts. In **Proceedings of the 12th International Conference on Electronic Commerce: Roadmap for the Future of Electronic Business**, ICEC '10, p. 179–188. Association for Computing Machinery, 2010.
- [5] Stamatina Thomaidou, Ismini Lourentzou, Panagiotis Katsivelis-Perakis, and Michalis Vazirgiannis. Automated snippet generation for online advertising. CIKM '13, p. 1841–1844. Association for Computing Machinery, 2013.
- [6] J. Weston Hughes, Keng-hao Chang, and Ruofei Zhang. Generating better search engine text advertisements with deep reinforcement learning. In **Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining**, KDD '19, p. 2269–2277. Association for Computing Machinery, 2019.
- [7] Shaunak Mishra, Manisha Verma, Yichao Zhou, Kapil Thadani, and Wei Wang. Learning to create better ads: Generation and ranking approaches for ad creative refinement. CIKM '20, p. 2653–2660. Association for Computing Machinery, 2020.
- [8] Hidetaka Kamigaito, Peinan Zhang, Hiroya Takamura, and Manabu Okumura. An empirical study of generating texts for search engine advertising. In **Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies: Industry Papers**, pp. 255–262. Association for Computational Linguistics, 2021.
- [9] 福田宏幸. キーワード条件つき変分 autoencoder による広告文生成. 人工知能学会全国大会論文集, Vol. JSAI2019, pp. 2L4J903–2L4J903, 2019.
- [10] Joshua Maynez, Shashi Narayan, Bernd Bohnet, and Ryan McDonald. On faithfulness and factuality in abstractive summarization. In **Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics**, pp. 1906–1919. Association for Computational Linguistics, 2020.
- [11] Artidoro Pagnoni, Vidhisha Balachandran, and Yulia Tsvetkov. Understanding factuality in abstractive summarization with FRANK: A benchmark for factuality metrics. In **Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies**, pp. 4812–4829. Association for Computational Linguistics, 2021.
- [12] Hoa Trang Dang. Overview of duc 2005. In **Proceedings of the 2005 Document Understanding Conference**, 2005.
- [13] Walter Dill Scott. **The theory of advertising**. Small, Maynard and Company, 1903.
- [14] 村上聡一郎, 星野翔, 張培楠, 上垣外英剛, 高村大也, 奥村学. Lp-to-text: マルチモーダル広告文生成. 言語処理学会 第 28 回年次大会, 2022.
- [15] Soichiro Murakami, Peinan Zhang, Sho Hoshino, Hidetaka Kamigaito, Hiroya Takamura, and Manabu Okumura. Aspect-based analysis of advertising appeals for search engine advertising. In **Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies: Industry Track**, pp. 69–78. Association for Computational Linguistics, 2022.
- [16] Taifeng Wang, Jiang Bian, Shusen Liu, Yuyu Zhang, and Tie-Yan Liu. Psychological advertising: Exploring user psychology for click prediction in sponsored search. In **Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining**, KDD '13, p. 563–571. Association for Computing Machinery, 2013.
- [17] Mike Lewis, Yinhan Liu, Naman Goyal, Marjan Ghazvininejad, Abdelrahman Mohamed, Omer Levy, Veselin Stoyanov, and Luke Zettlemoyer. BART: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension. In **Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics**, pp. 7871–7880. Association for Computational Linguistics, 2020.
- [18] Colin Raffel, Noam Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, Yanqi Zhou, Wei Li, and Peter J. Liu. Exploring the limits of transfer learning with a unified text-to-text transformer. **The Journal of Machine Learning Research**, Vol. 21, No. 1, 2022.
- [19] Kishore Papineni, Salim Roukos, Todd Ward, and Wei-Jing Zhu. Bleu: a method for automatic evaluation of machine translation. In **Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics**, pp. 311–318. Association for Computational Linguistics, 2002.
- [20] Chin-Yew Lin. ROUGE: A package for automatic evaluation of summaries. In **Text Summarization Branches Out**, pp. 74–81. Association for Computational Linguistics, 2004.
- [21] Diederik Kingma and Jimmy Ba. Adam: A Method for Stochastic Optimization. In **Proceedings of the 3rd International Conference on Learning Representations (ICLR 2015)**, 2015.

## A アノテーションガイドライン

アノテータに与えた主要な指示は次の3点である: (1) 検索キーワードをユーザの意図として考慮; (2) LPに記載の商材内容と矛盾しないように広告文を作成; (3) 広告文の長さは全角15文字以内<sup>9)</sup>.

## B LP情報を考慮しないと生成が難しい例

表4: LPを考慮しないと生成が難しい実例. 赤色下線部はLP上段部(3a)に, 青色下線部はLP中段部(3b)にそれぞれ対応した情報の記載がある.

LP テキスト (x)	キーワード (a)	広告文 (y)
豊富なサービスラインナップ(採用手法)と、業界に精通した貴社専任のプロフェッショナルが、採用ニーズにワンストップ・ワンブランドでお応えします。	デューダ 企業	<ol style="list-style-type: none"> <li>【企業の人事担当者さまへ】 - デューダ 企業</li> <li>企業様向け / <u>最短1日で利用開始</u></li> <li><u>業界最大級のサービス数 doda</u></li> <li><u>業界最大規模会員数 708 万人</u></li> </ol>



(a) LP 上段部

業界最大級のサービスラインナップを誇る、doda。				
求人広告掲載や人材紹介、転職フェア オンラインなどの転職イベント、ダイレクト・ソーシングなどを通じて、貴社の採用ニーズにワンストップ・ワンブランドでお応えします。				
	幅広く 募集をかけたい	直接 ハンティングしたい	効率的に PR したい	ピンポイントに 採用したい
サービス名	doda 求人検索サービス (doda求人広告/DM)	doda Recruiters (ダイレクト・ ソーシング)	doda 転職フェア オンライン	doda 人材紹介サービス
募集可能職種数	1	無制限	ゲームに沿った職種	無制限
母集団形成	マッチ度	-	-	●
	量	●	-	-
選考スピード	-	●	●	-
導入スピード	4日～	1日～	開業日による	-

(b) LP 中段部

## C ベースラインモデル実験設定

本実験で用いた広告文生成モデルは各事前学習済みモデルを CAMERA データの訓練セットを用いて finetuning を行うことで構築した. 具体的には, BART に基づくベースラインモデル (以下, BART) は, 京都大学が提供している日本語版 BART の事前学習済みモデル<sup>10)</sup>における `japanese_bart_base_2.0` を, T5 に基づくベースラインモデル (以下, T5) は, 村上ら [14] と同様に日本語版 T5 の事前学習済みモデル (sonoisa)<sup>11)</sup>を使用した. 次に, 両モデルの学習時のハイパーパラメータを表5に示す. ここで, T5 における最大文長については, LP の BBox 画像特徴量を用いるモデル (+{o,l,v}) のみ 712 に設定し, それ以外のモデルは全て 512 に設定している. また, 開発セットに対する loss が T5 に関しては 3epoch 連続, BART に関しては 5epoch 連続で劣化した場合には早期終了した.

表5: モデル学習時のハイパーパラメータ

ハイパーパラメータ	設定値 (BART / T5)
モデル	japanese_bart_base_2.0 / t5-base-japanese
最適化器	Adam [21]
学習率	3e-4
最大エポック数	20
バッチサイズ	8
最大文長	512 / 712 (T5+{o,l,v}モデルのみ)

9) Google レスポンシブ広告の見出し文に関する規定 (<https://support.google.com/google-ads/answer/12437745>) に従った.

10) [https://github.com/utanaka2000/fairseq/tree/japanese\\_bart\\_pretrained\\_model](https://github.com/utanaka2000/fairseq/tree/japanese_bart_pretrained_model)

11) <https://huggingface.co/sonoisa/t5-base-japanese>