

広告クリエイティブ自動生成にむけたマルチタスク学習と Conditional Attention による CVR 予測

北田 俊輔[†] 彌富 仁[†] 関 喜史[‡]

[†]法政大学大学院 理工学研究科 応用情報工学専攻, [‡]株式会社 Gunosy
{shunsuke.kitada.8y@stu., iyatomi@}hosei.ac.jp, yoshifumi.seki@gunosy.com

概要

広告クリエイティブに対するコンバージョン予測はデジタル広告における重要な問題の1つである。コンバージョンされやすい、効果の高いクリエイティブを作成するのは非常に難しいためクリエイティブ作成支援を行いたいが、学術的な取り組みは未だに少ない。そこで既存クリエイティブや入稿時のクリエイティブ評価のために、本研究では広告配信情報は使わず、クリエイティブのテキストや属性データから広告の効果を予測するシステムを構築する。そのために、マルチタスク学習の枠組みと conditional attention を提案する。マルチタスク学習は情報が少なく予測が難しいコンバージョンに加えて、その事前行動であるクリックを組み合わせて予測する。Conditional attention はクリエイティブの属性値に応じた attention を設定する。これらの提案手法に対し実際に配信されたクリエイティブを用いた評価実験を行い、予測精度を確認した。

1 はじめに

デジタル広告における広告クリエイティブ¹は購買対象となる顧客に対して効果的に製品の情報を届ける重要な役割を担っている [1]。広告から成果²が発生することをコンバージョンといい、広告主はコンバージョンが多く発生するクリエイティブを作成したい。広告の配信効果は $\frac{\text{広告による成果}}{\text{広告にかかったコスト}}$ によって定義される。広告におけるコストは一般的に広告がクリックされたときに支払が発生する Cost-Per-Click (CPC) 方式 [2] によって決まる。効果の良いクリエイティブを作るために広告主は試行錯誤を繰り返している。コンバージョンはユーザが広告をクリックしなければ発生しないため、まずは広告の Click Through Rate (CTR) が高く

なるようにクリエイティブを作成、運用する。しかし CTR を高めることに注力した結果、コンバージョンされる割合 (Conversion Rate (CVR)) が低いクリエイティブを生み出してしまうことがあり、効果の高いクリエイティブを作るのは非常に難しくコストがかかる。したがってそのようなクリエイティブの作成を支援することは重要であるが、既存のクリエイティブや入稿されたクリエイティブを正しく評価することは難しい。デジタル広告に対する CTR, CVR を予測するモデルはこれまで数多く提案されている [3, 4, 5] が、これらの先行研究では配信情報からユーザがクリックするか、コンバージョンするかを予測するものであり、本研究の目的であるクリエイティブの評価とは大きく異なる。一般の広告における CTR, CVR 推定はすでに存在する広告との配信データを用いて、効果の高い配信枠の決定やユーザーへの効率的な配信を目指したものである。一方で本研究の目指すクリエイティブ作成支援においては、まだ配信されていないクリエイティブの評価のために、配信データが未知のデータに対する予測を行うため、先行研究とは問題設定が大きく異なる。クリエイティブ作成支援の取り組みは企業側からのプレスリリースとしては報告 [6] があるものの、学術的な取り組みとしては報告されていない。

そこで本研究では配信情報を使わず、クリエイティブのデータのみから広告の効果を予測することでクリエイティブの評価を行うことを試みる。クリエイティブには画像とテキストがあるが、本研究では特にテキストに着目し、CVR の予測を試みる。クリエイティブ作成支援のためには CVR の正確な予測より、コンバージョンが多く集まる質の高いクリエイティブの予測が重要であることを踏まえて評価する。本研究における貢献は (1) 広告クリエイティブから配信情報なしでコンバージョンを予測するモデルの提案, (2) コンバージョン数及びクリック数を同時に学習・予測するマルチタスク学習によるモデルの性能向上の確認, (3)

¹広告においてバナーとテキストのことをクリエイティブという。

²アプリケーションのインストール, 商品の購入などを指す。

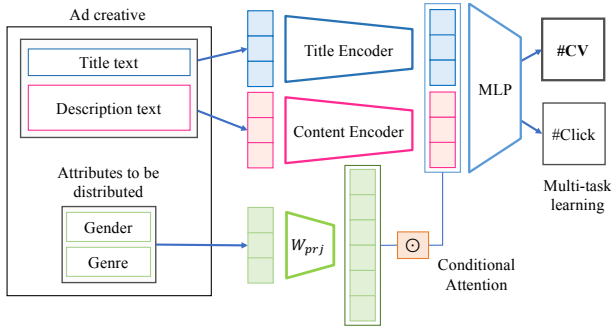


図 1: 本研究で用いるアーキテクチャの概要

Conditional attention による属性値に対する動的な注目単語の設定, およびその可視化の提案である.

2 提案手法

図 1 に本研究で用いるモデルアーキテクチャの概要を示す. 広告クリエイティブのテキストは主にタイトルと説明文によって構成されており, 2つのテキストは情報量を増やすために異なる文脈を持つことが多い. そこで今回使用するネットワークでは, タイトルテキストと説明文で異なるコンテキストを学習することを期待し, それぞれ2つの text encoder を用いてテキストから特徴を取得する. 具体的にはクリエイティブからタイトルテキストと説明文テキストをそれぞれ title encoder と content encoder に入力を行い, 特徴ベクトルを取得する. 本研究では, コンバージョンとクリックを同時に予測するマルチタスク学習と属性値に応じてテキスト内の重要な文字列を予測, 検出する conditional attention の2つを提案する. マルチタスク学習では, コンバージョン自体が少ないクリエイティブに対するコンバージョン予測において, 事前行動であるクリックを用いることで予測精度の向上を期待する. Conditional attention はクリエイティブの属性値に応じて attention を動的に設定可能にする仕組みであり, 配信対象の性別やクリエイティブのジャンルを考慮した attention が設定できることを期待する.

2.1 マルチタスク学習

CVR 予測において, コンバージョン情報は少なく対象広告のコンバージョンを適切に予測するのは難しいタスクの1つである. そこで本研究では図 1 に示すネットワークにおいて, コンバージョン情報よりも多いクリック情報も同時に活用することで, コンバー

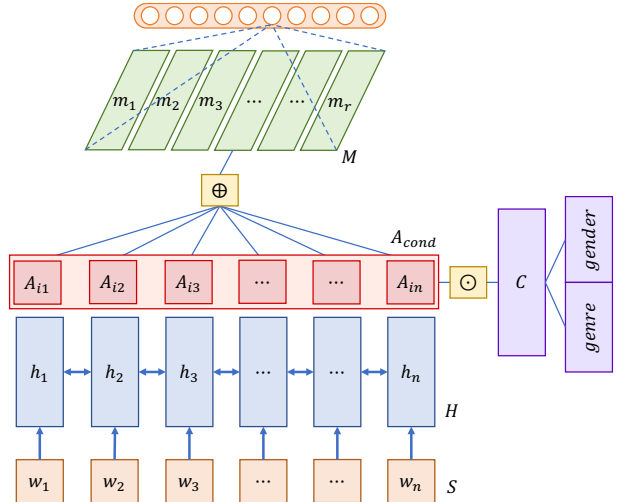


図 2: Conditional attention の概要

ジョン予測の精度向上を測った. このような枠組みはマルチタスク学習 [7] と呼ばれ, 主に自然言語処理において複数タスクに対して重みを共有するネットワークを用いることで精度向上が報告されている [8].

2.2 Conditional Attention

Attention 機構は入力系列に対してどの単語に注視するかを含めて学習可能な機構である. 配信対象の属性や, クリエイティブのジャンルによって注視すべき単語は異なることが予想される. 本研究ではこうした属性値から得られた特徴ベクトルを用いて self-attention [9] を計算する conditional attention を提案する. 単語数 n のクリエイティブテキスト $S = (w_1, w_2, \dots, w_n)$ に対して text encoder を適用し, $n \times u$ 次元の隠れ状態 H を取得する. 性別とジャンル各々 d_{gender} 次元, d_{genre} 次元の one-hot ベクトル表現を結合した d_{feat} 次元の \mathbf{x}_{feat} に対して線形変換を行うことで, 性別およびジャンルを考慮した d_c 次元の conditional ベクトル \mathbf{c} を計算する.

$$\mathbf{c} = \mathbf{W}_{prj} \mathbf{x}_{feat}$$

ここで \mathbf{W}_{prj} は $d_c \times d_{feat}$ 次元のパラメータ行列であり, d_c はハイパーパラメータである. 計算した conditional ベクトル \mathbf{c} と [9] より計算される attention ベクトル \mathbf{a} を用いて conditional attention ベクトル \mathbf{a}_{cond} を計算する.

$$\mathbf{a}_{cond} = \mathbf{a} \odot \mathbf{c}$$

ここで \odot は element-wise product である. r 個のクリエイティブテキストに対して attention を適用する

場合, conditional attention ベクトル \mathbf{a}_{cond} は $r \times d_c$ 次元の conditional attention 行列 A_{cond} となる. したがって埋め込み文ベクトル m は $r \times u$ 次元の埋込文行列 M となる. Conditional attention 行列 A_{cond} と text encoder の隠れ状態 H を乗算し, r 個の重み付き行列を計算する:

$$M = A_{cond}H$$

3 実験と結果

3.1 データセット

評価実験には, グノシー³において 2017 年 8 月から 2018 年 8 月の 1 年間で配信された広告クリエイティブデータを使用した. このデータのうち表示回数が 500 以上のクリエイティブ計 14,000 件を対象とした. 入力に使用した素性はクリエイティブのタイトルテキスト, 説明文, 性別, クリエイティブのカテゴリ属性である. 入力テキストは mecab-ipadic-neologd [10] を用いて形態素解析を行い, 分かち書きをした.

3.2 実験設定

ベースラインとしてサポートベクター回帰 (SVR) と multi-layer perceptron (MLP) を用いた. SVR にクリエイティブテキストを入力する際には, 事前学習済み word2vec を用いて単語ベクトルを得たのちに平均したものを文ベクトルとした. 実験では text encoder として gated recurrent unit (GRU) [11] および bi-directional GRU を利用し, attention を導入しないシンプルなもの (Vanilla), attention を導入したもの (Attention), そして提案手法である conditional attention を導入したもの (Conditional Attention) について比較を行った. 上記に加えて, CVR 予測におけるマルチタスク学習の効果を確認するため, コンバージョン数のみを予測するもの (single) およびコンバージョン数に加えてクリック数を予測するもの (multi-task) について比較を行った. これらネットワークは mean squared error (MSE) 誤差関数を目的関数とし, 誤差逆伝播法によってパラメータを最適化する. パラメータの最適化には Adam [12] を用いた.

3.3 評価方法

グノシーにおいて, 広告はキャンペーンという単位で配信されている. キャンペーンにはその広告がどの

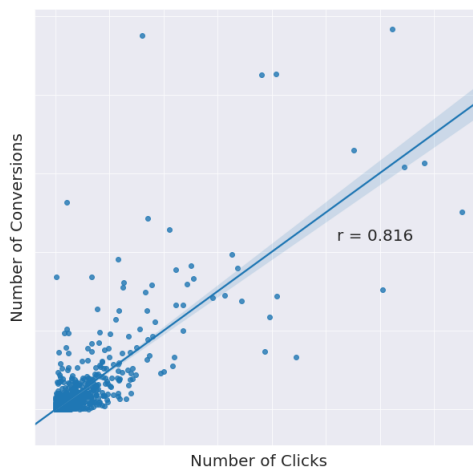


図 3: 使用データセットにおけるクリエイティブのクリック数とコンバージョン数の関係 (相関係数 $r = 0.816$)

ジャンルで, どの属性のユーザに配信するかが設定されており, 複数のクリエイティブが設定されている. 本研究では未知のキャンペーンにおけるクリエイティブテキストに対するコンバージョン数予測を行い, 汎化性能を確認する必要がある. したがって評価時には配信キャンペーンが重複しないような形で 5-fold cross validation を行った.

評価指標として mean average precision (MAP) および normalized discounted cumulative gain (NDCG) を使用した. これらは主にランキング学習の評価指標として用いられる. 一般には MSE や logloss といった指標が用いられるが, コンバージョンの多くが 0 であり, すべて 0 と予測しても評価スコアは良くなってしまふ場合がある. したがってコンバージョンの多いクリエイティブが上位となる予測がされているかを評価する必要がある. 本研究では多くコンバージョンされているクリエイティブを正確に予測するため, MAP 算出時にはコンバージョン数が 1 以上および 10 以上で閾値を設け, NDCG 算出時にはデータすべてに対する結果とコンバージョン数上位 1% に制限したときのスコアを計算した.

3.4 実験結果

実験結果を表 1 に示す. 文脈を考慮することが可能な GRU ベースのネットワークがベースラインの SVR および MLP の精度を上回ることが確認できた. 特に bi-directional GRU を text encoder として用いた場合に予測精度が高く, コンバージョンの多いクリエイ

³<https://gunosy.com/>

表 1: 各モデルに対する性能比

Model	Mean average precision (MAP)				Normalized Discounted Cumulative Gain (NDCG)				
	#CV >1		#CV >10		All		#CV top 1 %		
	single	multi-task	single	multi-task	single	multi-task	single	multi-task	
SVR	0.3469		0.0897		0.9672		0.8373		
MLP	0.3554	0.3624	0.0875	0.0956	0.9668	0.9718	0.8410	0.8297	
GRU	Vanilla	0.3434	0.3613	0.0803	0.0937	0.9654	0.9700	0.7639	0.7851
	Attention	0.3571	0.3615	0.0937	0.0965	0.9676	0.9711	0.8549	0.8300
	Conditional Attention	0.3575	0.3663	0.0965	0.1005	0.9677	0.9720	0.8711	0.8714
Bi-directional GRU	Vanilla	0.3531	0.3603	0.0958	0.0980	0.9680	0.9693	0.8088	0.7995
	Attention	0.3564	0.3619	0.0966	0.1005	0.9683	0.9717	0.8162	0.8325
	Conditional Attention	0.3613	0.3689	0.1005	0.1109	0.9686	0.9721	0.8392	0.8357

タイプを適切に予測できている。実験結果より、クリエイティブテキストにおいて語順を適切に捉えることは重要であると示唆される。

今回、マルチタスクモデルはコンバージョン数のみを予測するシングルモデルよりも高い予測精度を実現できた。通常広告クリエイティブは今回使用したデータセット図3のようにコンバージョン数とクリック数の間に高い相関があるため、コンバージョン数単体を予測するよりも、クリック数も同時に予測するマルチタスクでモデルを学習させることで予測精度が向上したと考えられる。

通常の attention 機構と比較して提案手法である conditional attention が MAP, NDCG とともに優れた結果であることが確認できた。従来の attention 機構では単語間の共起関係のみを考慮した注視単語の学習が行われていたが、クリエイティブ配信時に付随する属性情報を考慮することにより、配信時のコンバージョン数を適切に予測されている。

4 おわりに

本研究では広告クリエイティブ作成支援を目標に、マルチタスク学習を用いた枠組みと conditional attention を提案し、クリエイティブの評価を行った。コンバージョンとクリックを同時に予測するマルチタスク学習をにより、コンバージョンをより正確に予測することが可能となった。加えて conditional attention を導入することにより予測精度が向上し、同時にクリエイティブの属性値を考慮した動的な注目単語への可視化を行うことを可能とした。今後はクリエイティブに付随する画像も同時に使用したシステムを構築し、コンバージョン予測の精度向上を目指す。

参考文献

- [1] O. Chapelle, E. Manavoglu, and R. Rosales, "Simple and scalable response prediction for display advertising," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 5, no. 4, p. 61, 2015.
- [2] M. Mahdian and K. Tomak, "Pay-per-action model for online advertising," *Proceedings of the 1st international workshop on Data mining and audience intelligence for advertising*, pp. 1–6, 2007.
- [3] M. Richardson, E. Dominowska, and R. Ragno, "Predicting clicks: Estimating the click-through rate for new ads," *Proceedings of the 16th International World Wide Web Conference (WWW-2007)*, January 2007.
- [4] J. Chen, B. Sun, H. Li, H. Lu, and X.-S. Hua, "Deep ctr prediction in display advertising," *Proceedings of the 2016 ACM on Multimedia Conference*, pp. 811–820, 2016.
- [5] W. Liu, R. Tang, J. Li, J. Yu, H. Guo, X. He, and S. Zhang, "Field-aware probabilistic embedding neural network for ctr prediction," *Proceedings of the 12th ACM Conference on Recommender Systems*, pp. 412–416, 2018.
- [6] T. Kazuki, O. Kazuhiro, and Y. Kota, "New advertising creative by artificial intelligence," *Journal of the Japanese Society for Artificial Intelligence*, vol. 32, no. 4, pp. 501–508, jul 2017.
- [7] R. Caruana, "Multitask learning," *Machine learning*, vol. 28, no. 1, pp. 41–75, 1997.
- [8] R. Collobert, J. Weston, L. Bottou, M. Karlen, K. Kavukcuoglu, and P. Kuksa, "Natural language processing (almost) from scratch," *Journal of Machine Learning Research*, vol. 12, no. 8, pp. 2493–2537, 2011.
- [9] Z. Lin, M. Feng, C. N. d. Santos, M. Yu, B. Xiang, B. Zhou, and Y. Bengio, "A structured self-attentive sentence embedding," *CoRR arXiv:1703.03130*, 2017.
- [10] S. Toshinori, "Neologism dictionary based on the language resources on the web for mecab," 2015. [Online]. Available: <https://github.com/neologd/mecab-ipadic-neologd>
- [11] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," *CoRR arXiv:1412.3555*, 2014.
- [12] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *CoRR arXiv:1412.6980*, 2014.