

# 生存時間分析を用いた広告クリエイティブの停止予測

## Stop Prediction for Ad Creatives using Survival Analysis

北田 俊輔 \*<sup>1</sup>      彌富 仁 \*<sup>1</sup>      関 喜史 \*<sup>2</sup>  
Shunsuke Kitada      Hitoshi Iyatomi      Yoshifumi Seki

\*<sup>1</sup>法政大学 理工学研究科 応用情報工学専攻  
Major in Applied Informatics, Graduate School of Science and Engineering, Hosei University

\*<sup>2</sup>株式会社 Gunosy  
Gunosy Inc.

Discontinuation of ad creative is an important process to select high-performance ad creatives. In this study, we develop a prediction framework for ad creative discontinuation; the framework reduces the burden on digital ad operations load and helps to increase business revenue. Our ad creative discontinuation framework consists of deep neural networks take the text, category, image and numerical information of the targeted ad. It estimates the appropriate discontinuation time with a survival prediction strategy. Here, we propose two simple but extremely effective techniques to enhance the prediction performance; (1) sales-based loss function, and (2) two-period estimation. The former considers the importance of ads by weighting the loss function according to the click-through rate (CTR). The latter separately estimates “short-term censoring” as a short-term discontinuation and “long-term wear-out” as a long-term discontinuation, taking into account these different properties. We evaluated our proposal framework using the record of real-world 1,000,000 ad creatives provided Gunosy Inc. with a concordance index (CI). Our proposal multi-modal deep neural network framework performed (0.623 in CI) better than conventional method (0.531). Our two-period estimation largely improved the prediction performance by approximately 20 point on both short-term (0.829) and long-term (0.789) discontinuation. The introduction of sales-based loss further improves performance by an average of approximately 3 point in CI (short: 0.857, long: 0.823).

## 1. はじめに

近年、デジタル広告の重要性が高まるにつれて、広告運用の重要性も高まっている。広告の運用は、広告クリエイティブの作成、入札価格と配信対象の決定、非効率的な広告の停止など、様々なタスクで構成されている。広告市場の成長に伴い、機械学習を用いた広告運用プロセスの支援を行う研究が進められている [Kitada 19, Mishra 19, Maehara 18, Yang 19]。これらは主に配信効果の高い広告クリエイティブの作成支援に焦点を当てたものであり、配信効果の低い広告クリエイティブの停止を支援する研究は著者らが知る限り存在しない。

効果の高いクリエイティブの作成支援を行う研究が増えている一方で、クリエイティブの運用については十分に検討されているとは言えない。運用において、配信された各クリエイティブの効果を確認し、効果が悪い広告を停止するプロセスがある。しかしこのプロセスにおける停止判断の多くは人手で行われている。デジタル広告では効果のよいクリエイティブを探すために運用初期に多くのクリエイティブを入稿するため、その作業は非常に煩雑である。大量に入稿されたクリエイティブの大部分は短期間で配信が打ち切りになる。配信初期で効果が良いと配信が継続するが、これら飽きられていくため日毎に悪化していく。これは一般的に摩耗と呼ばれている。運用者は摩耗する前に次の効果の良いクリエイティブを作らなくてはならない。このような打ち切りや摩耗を予測できれば運用を効率化できるが、そうした部分は運用者の経験に委ねられている。

本研究ではこうした広告運用の効率化を目指して、生存時間予測の枠組みを利用した広告クリエイティブにおける停止の予測を行う。生存時間予測は医療分野で発展し始め、現在では幅広い研究分野で一般的に使われる枠組みとなっている [Dirick 17, Barbieri 16]。近年では深層学習を元にした生存

連絡先: 北田 俊輔, 法政大学大学院 理工学研究科 応用情報工学専攻, 東京都小金井市梶野町 3-7-2 S603, shunsuke.kitada.8y@stu.hosei.ac.jp

時間予測モデルが提案されており、優れた予測精度を実現している [Katzman 18, Gensheimer 19]。深層学習の利用により、画像、テキスト、時系列など様々な特徴から構成される広告クリエイティブにおいて高い性能を発揮することが期待できる。

本研究では、広告クリエイティブの停止に対して生存時間予測による深層学習を元にした枠組みを提案する。私たちの提案する枠組みは“売上を元にした損失項の導入”と“長期と短期をそれぞれ分割して予測する2期間推定法の導入”という2つの大変効果の高い技術から構成される。前者はより価値のある広告クリエイティブが持つ特徴の学習を助ける役割がある。後者は短期の打ち切りと長期の摩耗をそれぞれ別々のモデルで予測を行う。実世界のデータセットを用いた様々な評価実験によって、提案する枠組みの効果を確認した。

## 2. 提案手法

### 2.1 概要

本研究では、深層学習を元にした生存時間予測を用いた広告クリエイティブ停止予測の枠組みを提案する。図1に提案する枠組みを示す。この枠組みは広告クリエイティブを構成する4つの特徴量を使用する: テキスト特徴 (タイトル, 説明文), カテゴリカル特徴 (配信対象の性別, 広告クリエイティブのジャンル), 画像特徴 (バナー画像), 時系列特徴を含む数値特徴 (インプレッション数, クリック数, コンバージョン数, 広告のCAP)。提案する枠組みはこれら特徴量を入力として受け取り、停止する時間を予測した。テキスト特徴や時系列特徴  $\mathbf{b}$  は recurrent neural network (RNN) を元にした long short-term memory (LSTM) によって特徴を抽出した。

深層学習により、広告クリエイティブを構成する様々な要素から特徴を学習可能である。広告を構成する要素から抽出されるテキスト特徴  $\mathbf{x}^{\text{text}}$ , カテゴリカル特徴  $\mathbf{x}^{\text{cat}}$ , 画像特徴  $\mathbf{x}^{\text{img}}$ , 統計特徴  $\mathbf{x}^{\text{stat}}$ , そして時系列特徴  $\mathbf{x}^{\text{time}}$  は結合されて

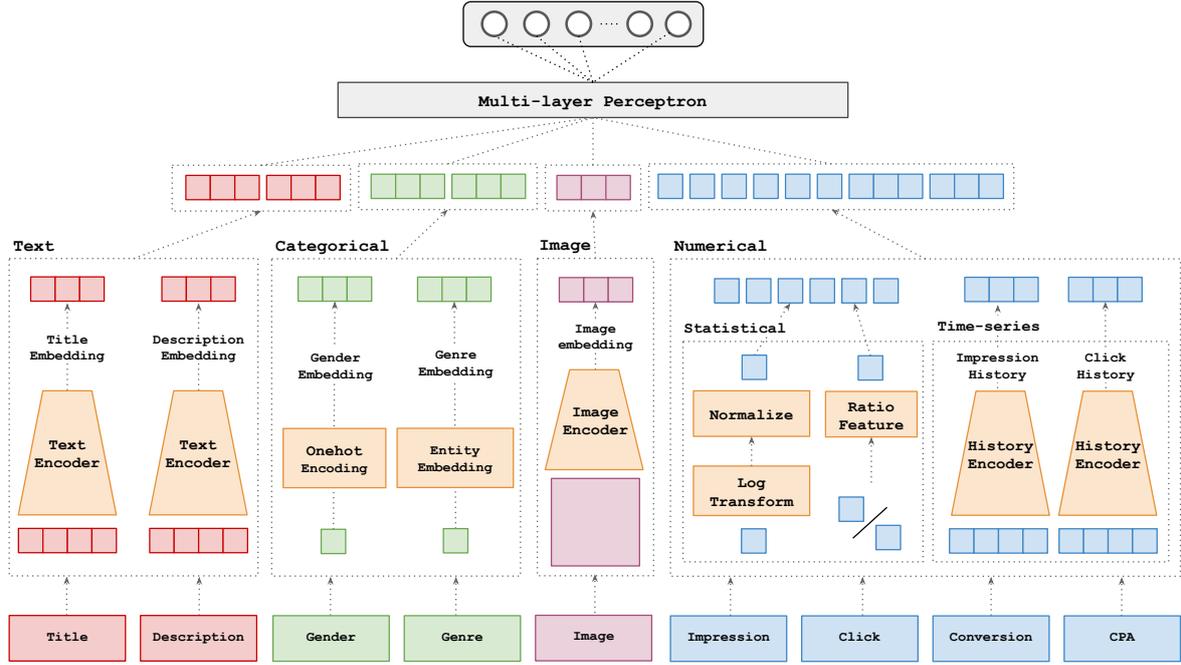


図 1: 提案する生存時間予測を用いた広告クリエイティブの停止予測の枠組み. 入力テキスト特徴, カテゴリカル特徴, 画像特徴, 数値特徴からなる. 出力はあらかじめ定義された期間における, 対象クリエイティブが停止したかどうかである.

多層パーセプトロン (MLP) へと入力される.

$$\hat{\mathbf{y}} = \text{MLP}([\mathbf{x}^{\text{text}}; \mathbf{x}^{\text{cat}}; \mathbf{x}^{\text{img}}; \mathbf{x}^{\text{stat}}; \mathbf{x}^{\text{time}}]) \in \mathbb{R}^{n_{\tau}}. \quad (1)$$

このとき  $\hat{\mathbf{y}}$  は予測される広告クリエイティブの停止確率である. 広告クリエイティブの構成要素に対して, 従来のエンコーディング手法や深層学習に基づくエンコーディング手法を適用して特徴抽出を行った. 学習時には生存時間予測で最先端の結果を報告している [Gensheimer 19] の損失項を利用した.

## 2.2 売上を元にした損失項

一般的に生存時間予測は全てのデータに対して等しく予測を行うが, 今回対象となる広告クリエイティブの重要度は収益度合いによって異なる. したがって, 収益度合いが一般的に高い, click-through rate (CTR) の高い広告クリエイティブの特徴を捉えることは重要である. このような広告クリエイティブに対して正確に予測するため, 各広告クリエイティブの CTR で重み付けをした CTR-weighted loss を最小化する.

$$\mathcal{L}_{\text{CTR}}(\mathcal{T}, \chi) = (r_{\text{CTR}} + 1) \cdot \mathcal{L}(\mathcal{T}, \chi). \quad (2)$$

ここで,  $r_{\text{CTR}}$  は  $r_{\text{CTR}} = \# \text{click} / \# \text{impression}$  によって計算され, 一般的に  $0 \leq r_{\text{CTR}} \leq 1$  の範囲の値を取る. このとき, CTR の低い広告クリエイティブに対する損失が急激に小さくならないように 1 を足すことにより, こうした問題を防ぐ. 私たちはこの損失項が事業上貢献しうるであろう CTR の高い広告クリエイティブの特徴をより正確に捉えることを期待する. 評価のために, 広告の表示回数で重み付けを行う impression-weighted loss と比較する.

## 2.3 短期および長期における停止の予測

広告クリエイティブの停止において短期の打ち切りと長期の摩耗という 2 つの側面があることに着目し, これらを分割し

て予測する 2 期間推定法を提案する. このとき, これら 2 つの期間を予測するモデルを短期モデルと長期モデルと呼ぶ.

短期モデルは短期の打ち切りを予測するための特徴を学習し, 長期モデルは長期の摩耗を予測するための特徴を学習する. 短期モデルにおいて, 時間間隔  $\mathcal{T}_S$  が与えられたとき対数尤度  $\mathcal{L}$  は次のように定義される.

$$\mathcal{L}(\mathcal{T}_S, \chi) = \sum_{i=1}^{N_S-1} \ln(1 - h_i(\mathcal{T}_S|\chi)). \quad (3)$$

ここで  $N_S$  は  $\mathcal{T}_S$  における時間間隔数である. また, 長期モデルにおいて, 時間間隔  $\mathcal{T}_L$  が与えられたときの対数尤度  $\mathcal{L}$  は次のように定義される.

$$\mathcal{L}(\mathcal{T}_L, \chi) = \sum_{i=1}^{N_L-1} \ln(1 - h_i(\mathcal{T}_L|\chi)), \quad (4)$$

ここで  $N_L$  は  $\mathcal{T}_L$  における時間間隔数である. それぞれのモデルでこれら異なる時間間隔で学習させることで, 短期および長期の停止に対する特徴が学習され, 全体を予測するよりも良い予測ができることを期待する.

## 3. 評価実験

### 3.1 実験設定

提案する枠組みを評価するために, 株式会社 Gunosy から提供された 1,000,000 件の実世界における広告クリエイティブデータセットを使用した. このデータセットに対してキャンペーンに基づいた層別分割を行い, 訓練データ 60,000 件, 開発データ 20,000 件, 評価データ 20,000 件を得た.

深層学習を元にした枠組みの効果を確認するために, 従来手法である Cox Regression [Cox 72] および Random Survival Forest [Ishwaran 08] とを比較した. また提案する枠組みに入

表 1: 異なる時間間隔および異なる特徴量における Concordance index の比較. 時間間隔においては短期モデル, 長期モデル, および全体モデルを比較した. 売上上位 25% の広告クリエイティブに対する評価結果も同様に示した. 提案するマルチモーダルな深層学習による枠組みが従来手法よりも 10pt 程度差をつけて最良の結果を示した. 全体モデルと比較して, 提案する 2 期間推定法では短期及び長期モデル共に 20pt 程度の大規模な予測精度の改善を確認した.

Model	Feature				Concordance Index					
					短期モデル		長期モデル		全体モデル	
	Stat.	Image	Text	Time.	All	Top 25% of sales	All	Top 25% of sales	All	Top 25% of sales
Cox Regression [Cox 72]	✓		✓		0.6098	0.7293	0.6574	0.6932	0.5287	0.5320
Random Survival Forest [Ishwaran 08]	✓		✓		0.6213	0.7578	0.6832	0.7294	0.5311	0.5487
<b>Our framework</b>	✓				0.7896	0.8229	0.7535	0.8040	0.5535	0.6040
	✓	✓			0.7962	0.8232	0.7536	0.8045	0.5536	0.6045
	✓		✓		0.7958	0.8262	0.7541	0.8001	0.5874	0.6001
	✓			✓	0.7931	0.8346	0.7880	0.8397	0.5935	0.6101
	✓	✓	✓		0.7972	0.8313	0.7575	0.8113	0.5908	0.6113
	✓	✓	✓	✓	<b>0.8289</b>	<b>0.8640</b>	<b>0.7892</b>	<b>0.8456</b>	<b>0.6225</b>	<b>0.6456</b>

力される特徴量の効果を確認するために, 広告クリエイティブの画像特徴 (バナー画像), テキスト特徴 (タイトルおよび説明文), 時系列特徴 (配信実績) を比較した.

広告クリエイティブ停止の打ち切りおよび摩擦を考慮するため, 予測する時間間隔  $T_S$ ,  $L_L$  をそれぞれ  $[1, 3)$ ,  $[3, 5)$ ,  $[5, 7)$ ,  $[7, 10)$  および  $[1, 10)$ ,  $[10, 30)$ ,  $[30, 60)$ ,  $[60, 90)$ ,  $[90, 120)$  に設定した. これらの設定は訓練データにおける分析で, 殆どのクリエイティブがそれぞれ約 10 日と約 120 日で停止されていることに基づいている. また, 時間間隔を指数的にすると効果的であるという主張 [Gensheimer 19] に基づいている.

短期モデルおよび長期モデルの効果を確認するために, これらのモデルの予測期間を含む全体モデルを構築した. 評価時には, 短期モデルでは評価データの 3 日間の情報, 長期モデルでは評価データの 10 日間の情報を元に停止予測を行った.

### 3.2 評価指標

評価指標として, 生存時間予測で広く用いられている concordance index (CI) による比較を行った. これはモデルが予測した生存時間と, 実際の生存時間の大小関係がどの程度一致しているかを表す.  $CI = 0.5$  の場合はランダムな予測と等価であり,  $CI = 1.0$  の場合は完璧な予測であることを示す.

事業的な観点から売上上位の広告クリエイティブに対して正確な停止の予測を行うことは重要である. 本研究では売上上位 25% のクリエイティブに対して CI による比較を行うことで, 事業に影響のある広告の停止を正確に予測できるか評価した.

### 3.3 実験結果

表 1 に CI による比較を示す. 提案する深層学習を元にした枠組みが, 従来手法と比較して 10pt 以上の差をつけて最良の結果を示した. 広告クリエイティブを構成する特徴量をすべて入力することで予測性能は更に向上した. これらの特徴量はそれぞれ予測に有効であった. 停止を短期と長期に分ける 2 期間推定法により, 全体の期間を予測する全体モデルよりもそれぞれ平均して 20pt 程度の性能向上を確認した. また, 表 2 に売上を元にした損失項である CTR-weighted loss と impression-weighted loss の比較を示す. CTR-weighted loss を導入することにより, 短期モデル, 長期モデル, また全体モデルにおいても予測精度の向上が確認できた. 短期モデルおよび長期モデル共に広告クリエイティブ特徴と時系列特徴を使用

したモデルが平均 4pt 程度差をつけて最良の結果を示した.

### 3.4 考察

**入力する各特徴量の比較.** 提案した予測モデルでは深層学習を元にしたモデルであることを生かして様々な特徴量を扱った. 表 1 に示したように, 同じ特徴量を入力に与えた場合においても, 深層学習を用いた生存時間予測モデルは, 深層学習ではないモデルと比較し高い性能を示した. 深層学習を用いることで予測性能の向上が実現できるだけでなく, 広告クリエイティブを構成する様々な特徴量を入力することが可能である.

本研究では広告クリエイティブの画像, テキスト, そして配信情報を RNN を用いて時系列で扱った特徴量を入力した. 特に配信情報を時系列で入力したときに大きな改善を確認した. 画像およびテキストは短期モデルと長期モデルに分けたときには大きな改善は見られなかった. 一方で, 全体を対象にした全体モデルにおいて改善を示していることから, 予測に一定の貢献があることが示唆された. 時系列のデータは特に配信日数が長い広告クリエイティブの停止の予測において貢献しており, 摩擦を予測することに配信実績が寄与していることを示している. 様々な特徴量によって予測が改善することから, 提案手法には現在人で行われている判断を助ける効果を期待できる.

**短期長期の停止の比較.** 本研究では, 広告クリエイティブの停止において, 短期の打ち切りと長期の摩擦という 2 種類の異なる性質があることに注目した. これを踏まえて, 2 種類の停止を分けて予測する提案手法である 2 期間推定法により, 全体の期間を予測するよりも大規模な予測性能の向上を確認した.

短期および長期モデル共に広告クリエイティブ特徴と時系列特徴を使用したモデルが平均 4pt 程度差をつけて最良の結果を示した. 短期モデルと長期モデルを比較すると, 特に時系列特徴を追加した場合に長期モデルの精度改善が大きく, 画像やテキストなどの特徴は統計データのみの予測と比較してあまり貢献していない. 一方で短期モデルではこれらの特徴が若干ながら精度に寄与している. これは長期モデルでの広告クリエイティブの停止要因が摩擦によるものであり, 広告クリエイティブを構成する主要要素であるテキストや画像がユーザにとって必ずしも魅力的ではなかったからではないと考えられる.

画像やテキストは, 短期モデルと長期モデルを分けなかったときに精度向上に大きく貢献している. 画像やテキストは短期で打ち切られるか長期的に配信されるかといった予測には寄与

表 2: 売上を元にした損失項である CTR-weighted loss  $\mathcal{L}_{CTR}$  および impression-weighted loss  $\mathcal{L}_{imp}$  の比較. ベースラインモデルに対して  $\mathcal{L}_{CTR}$  および  $\mathcal{L}_{imp}$  を導入して学習を行った. CTR-weighted loss を導入することにより, 短期モデル, 長期モデル, また全体モデルにおいてもそれぞれ 3pt 程度の予測精度改善を確認した.

Model	Feature				Concordance Index								
					短期モデル			長期モデル			全体モデル		
	Stat.	Image	Text	Time.	Without	$\mathcal{L}_{imp}$	$\mathcal{L}_{CTR}$	Without	$\mathcal{L}_{imp}$	$\mathcal{L}_{CTR}$	Without	$\mathcal{L}_{imp}$	$\mathcal{L}_{CTR}$
Our framework	✓				0.7896	0.8019	0.8199	0.7535	0.7660	0.7869	0.5535	0.5655	0.5869
	✓	✓			0.7962	0.8076	0.8212	0.7536	0.7652	0.7848	0.5536	0.5654	0.5799
	✓		✓		0.7958	0.8080	0.8270	0.7541	0.7666	0.7827	0.5874	0.5994	0.6124
	✓			✓	0.7931	0.8042	0.8232	0.7880	0.7991	0.8198	0.5935	0.6050	0.6311
	✓	✓	✓		0.7972	0.8084	0.8218	0.7575	0.7694	0.7861	0.5908	0.6022	0.6158
	✓	✓	✓	✓	0.8289	0.8400	<b>0.8567</b>	0.7892	0.8014	<b>0.8225</b>	0.6225	0.6340	<b>0.6548</b>

するが、その中での詳細な予測には寄与しないと考えられる。売上を元にした損失項の効果。私たちは価値が高いクリエイティブの特徴をより正確に学習する CTR-weighted loss を提案した。また比較として、表示回数で重み付けする impression weighted loss を試した。これらはいずれも重み付けのないものよりも良い結果を示した。売上上位のクリエイティブでの精度だけでなく、全体でも精度の改善が実現できており、効果のよいクリエイティブを優先的に学習することが、全体の予測に良い影響を与えていると推測できる。

Impression-weighted loss と比較して、CTR-weighted loss はより良い結果を示した。これは impression の分散が大きいが理由として考えられる。広告における impression は多いものと少ないものの差が極端に大きい。そのため正規化後の impression を掛け算すると不用意に損失が小さくなってしまふ場合がある。また広告クリエイティブのジャンルや配信対象によっても大きく異なるため、正規化が困難である。一方で CTR は一定の範囲に収まるため、このような問題は発生せず、広告の属性によらない重み付けが可能である。これらの結果により、impression で重み付けするより、CTR で重み付けしたほうが、より正確な予測のための訓練を可能だと考えられる。

#### 4. 結論

本研究では、広告クリエイティブの停止予測に対して生存時間予測を利用した枠組みを提案した。広告クリエイティブの停止は配信効果の高いクリエイティブを選択するために重要なプロセスであるが支援する研究は未だ少ない。

私たちは深層学習を元にした広告クリエイティブを構成するさまざまな特徴量を考慮した生存時間予測の枠組みを提案した。この枠組みは“売上を元にした損失項の導入”と、“長期と短期をそれぞれを分割する 2 期間推定法の導入”という 2 つの大変効果の高い技術から構成される。

提案する枠組みは株式会社 Gunosy から提供された 1,000,000 件の実世界における広告クリエイティブデータセットを用いて評価した。提案するマルチモーダルな DNN を元にした枠組みは従来手法よりも高い予測精度を実現した。2 期間推定法では短期モデルおよび長期モデル共に 20pt 程度の大幅な予測精度の改善を確認した。売上を元にした損失項を導入することで、さらに 3pt 程度の予測精度の向上を確認した。

今後の展望として、私たちの提案する枠組みを実際のシステムに導入することで、広告運用業務の改善を目指す。また、対象の広告クリエイティブの停止予測に対して、なぜその期間に停止されるかを解釈可能な形で提供したい。

#### 参考文献

- [Barbieri 16] Barbieri, N., Silvestri, F., and Lalmas, M.: Improving post-click user engagement on native ads via survival analysis, in *Proc. of WWW*, pp. 761–770 (2016)
- [Cox 72] Cox, D. R.: Regression models and life-tables, *Journal of the Royal Statistical Society: Series B (Methodological)*, Vol. 34, No. 2, pp. 187–202 (1972)
- [Dirick 17] Dirick, L., Claeskens, G., and Baesens, B.: Time to default in credit scoring using survival analysis: a benchmark study, *Journal of the Operational Research Society*, Vol. 68, No. 6, pp. 652–665 (2017)
- [Gensheimer 19] Gensheimer, M. F. and Narasimhan, B.: A scalable discrete-time survival model for neural networks, *PeerJ*, Vol. 7, (2019)
- [Ishwaran 08] Ishwaran, H., Kogalur, U. B., Blackstone, E. H., Lauer, M. S., et al.: Random survival forests, *The annals of applied statistics*, Vol. 2, No. 3, pp. 841–860 (2008)
- [Katzman 18] Katzman, J. L., Shaham, U., Cloninger, A., Bates, J., Jiang, T., and Kluger, Y.: DeepSurv: personalized treatment recommender system using a Cox proportional hazards deep neural network, *BMC medical research methodology*, Vol. 18, No. 1, p. 24 (2018)
- [Kitada 19] Kitada, S., Iyatomi, H., and Seki, Y.: Conversion Prediction Using Multi-Task Conditional Attention Networks to Support the Creation of Effective Ad Creatives, in *Proc. of KDD*, p. 2069–2077 (2019)
- [Maehara 18] Maehara, T., Narita, A., Baba, J., and Kawabata, T.: Optimal Bidding Strategy for Brand Advertising, in *Proc. of IJCAI*, p. 424–432 (2018)
- [Mishra 19] Mishra, S., Verma, M., and Gligorijevic, J.: Guiding creative design in online advertising, in *Proc. of RecSys*, pp. 418–422 (2019)
- [Yang 19] Yang, X., Li, Y., Wang, H., Wu, D., Tan, Q., Xu, J., and Gai, K.: Bid Optimization by Multivariable Control in Display Advertising, in *Proc. of KDD*, p. 1966–1974 (2019)