

オンライン広告における CVR 予測モデルの素性評価

田頭 幸浩[†] 小野 真吾[†] 田島 玲[†]

[†] ヤフー株式会社 〒107-6211 東京都港区赤坂 9-7-1 ミッドタウン・タワー

E-mail: [†]{yutagami,shiono,atajima}@yahoo-corp.jp

あらまし オンライン広告はインターネットの経済を支える大きな柱の一つであり、ビジネスと学術の両方から大きな注目を浴びている。本稿ではオンライン広告のコンバージョン率 (conversion rate; CVR) を予測するモデル (CVR 予測モデル) の素性について評価を行う。CVR 予測モデルの精度向上は広告効果の最大化に不可欠である。『Yahoo! ディスプレイアドネットワーク』の配信システムログを用いて CVR 予測モデルの素性を評価し、その有用性を検証した。

キーワード オンライン広告, コンバージョン率, CVR 予測, 機械学習, 統計モデル

1. 導 入

オンライン広告はインターネットの経済を支える大きな柱の一つである。そのため、この分野はビジネス的に、また学術的にも大きな注目を浴びている。オンライン広告の例としては、検索サイトにおける検索連動型広告、ポータルサイトにおけるディスプレイ広告、ニュースサイトやブログ記事のページにおけるコンテキスト広告が挙げられる。

広告主がオンライン広告を出稿する理由にはさまざまな動機が考えられる。例えば、ショッピングサイトを運営する広告主の場合、ユーザーに自社サイトで商品を購入してもらうことを目的として広告を出稿することが考えられる。このような広告主が目標とする特定のユーザー行動をコンバージョンと呼ぶ。コンバージョンの定義は広告主や広告キャンペーンごとに異なり、上記の商品購入以外にもウェブサービスへの会員登録やスマートフォンアプリのインストールを例として挙げることができる。扱う商品が不動産などの高額商品であり、オンラインで決済まで至ることが難しい場合には商品の資料請求がコンバージョンとして定義されることもある。

オンライン広告では広告効果を測るための重要な指標として、コンバージョン数やコンバージョン率 (conversion rate; CVR)、コンバージョンごとのコスト (cost per acquisition; CPA) が用いられている。広告主はこれらの指標を用いて広告効果の測定や最適化を行う。

広告効果を最大化するとはすなわち、限られた広告予算の中で数多くのコンバージョンを得ることと言い換えることができる。従来のオンライン広告では広告効果を高めるために、ターゲティングという仕組みを用いて特定のユーザー集合や時間帯にのみ広告を配信することがされてきた。ターゲティング設定を行うことで、ユーザーの性別や年齢などの属性やオンライン上での行動履歴をもとにした、ある特定のセグメントにのみ広告を配信することができる。これにより CPA が低いユーザー集合にのみ広告配信を行うなどの制御が可能になる。

従来のオンライン広告におけるターゲティングは上記のように、定められた区分の中からユーザー集合を選択するものだった

が、近年ではそれぞれの広告キャンペーンごとに最適なユーザー集合を動的に作成することも行われるようになった。また、real time bidding (RTB) と呼ばれる、各広告インプレッションごとにオークションへの参加や入札額をリアルタイムで決定する仕組みも広く浸透してきた。このような動きにより CVR 予測の重要性は高まっている。CVR 予測の精度を高めることで、各広告キャンペーンは広告効果を最大化するよう、適切に予算を配分することが可能になる。

CVR 予測と似た問題として、広告のクリック率 (click-through rate; CTR) を予測する、CTR 予測がある。現在の広告配信システムでは、広告配信ログをもとに機械学習で CTR や CVR を予測するモデルを構築するのが一般的である。CTR 予測モデルは CVR 予測モデルに先んじて、検索連動型広告やコンテキスト広告に対して多くの研究がされてきた [4], [13]。そのため CVR 予測では CTR 予測で用いられた素性やモデルなどの知識が活かされている。しかしながら、コンバージョンはクリックと比較して数が少なく、また広告キャンペーンごとに定義が異なる。加えて、広告がクリックされるかどうかはウェブページやユーザーと広告自体によって決まる一方、コンバージョンは広告主サイトの内容や構造にも影響を受けていると考えられる。これらのことから、CVR 予測と CTR 予測に対する有効なアプローチには違いがあると考えられる。

本稿ではクリック課金型広告の配信ログを用いて CVR 予測モデルの素性を評価し、その有用性を検証する。クリック課金型広告とは、広告が配信されたページを閲覧しているユーザーが広告をクリックして広告主の設定したページに移動した場合に、入札額に基づき課金される仕組みである。広告クリック時に課金が行われることから、本稿では**広告クリックを経たコンバージョン (post-click conversion)** 確率の予測について扱う。

本稿の構成は以下のとおりである。2 章では研究背景について述べる。続く 3 章では今回用いるデータセットと素性について説明する。4 章では CVR 予測モデルについて述べる。5 章では実際の広告配信ログを用いた評価を行った結果を紹介する。これらを受け 6 章では本稿を結び、将来の展望について述べる。

2. 研究背景

この章では、まずオンライン広告の概要について説明する。続いて関連研究について述べる。オンライン広告の全体像については Broder らの講義資料 [3] や Yuan らの文献 [20], [21] が詳しい。

2.1 オンライン広告の概要

オンライン広告には、広告主、パブリッシャー、ユーザーと呼ばれる三種類の主要なプレイヤーが存在する。広告主は広告を出稿することにより、自社のブランド力向上や製品購入などを目的とする。パブリッシャーはウェブページのオーナーであり、保有するウェブページ上に広告を掲載して収益を得ることを目的とする。ユーザーはパブリッシャーのウェブページを訪れ、そこに掲載された広告を閲覧する、広告主の潜在的な顧客である。広告を閲覧したユーザーは広告のクリックやウェブ検索を通して広告主のサイトを訪れることがある。

ユーザーがオンライン広告に触れて広告主の顧客となるにはインプレッション、クリック、コンバージョンの三段階を経る。インプレッションとは、ユーザーがウェブページに掲載された広告を閲覧することを表す。クリックは、ユーザーが表示された広告をクリックすることを指す。これによりユーザーは広告主があらかじめ指定したウェブページに遷移する。コンバージョンは、広告主が目標とする特定のユーザー行動である。ショッピングサイトを運営する広告主の場合、コンバージョンの例として、自社サイトで商品を購入することが挙げられる。ユーザーから見たインプレッション、クリック、コンバージョンの流れを図 1 に示す。

パブリッシャーはウェブページに広告を掲載する見返りとして、広告主に課金を行う。オンライン広告の課金方式は複数あり、例として、広告表示ごとに課金を行うインプレッション課金や、ユーザーが広告をクリックするごとに課金を行うクリック課金が挙げられる。現在のオンライン広告の一部では、広告表示の機会ごとにオークションが行われている。このオークションでは、対象のインプレッション機会に対して複数の広告キャンペーンがそれぞれ入札額を提示し、良い条件を提示したキャンペーンから順に広告が表示される。これは、各広告キャンペーンは広告表示の機会ごとに、その権利を得るオークションに参加すると言い換えることができる。一般的にパブリッシャーは収益を最大化することを目的とするため、期待収益の高い順に広告が表示される。ここでの期待収益は広告を掲載した際に得られる課金額の期待値のことを指し、インプレッション課金であれば入札額、クリック課金であれば入札額と CTR の予測値を掛け合わせた値である。実際の課金額はオークション方式により異なり、オンライン広告で一般的に用いられる generalized second price auction (GSP) では、次点の広告の期待収益をもとに課金額が決定される。

例えば、クリック課金型の検索連動型広告の場合、広告主は広告掲載の依頼時に特定の検索キーワードに対して入札額を設定する。ユーザーがウェブ検索を行った際、検索結果画面への広告表示に対するオークションが行われ、入力された検索キー

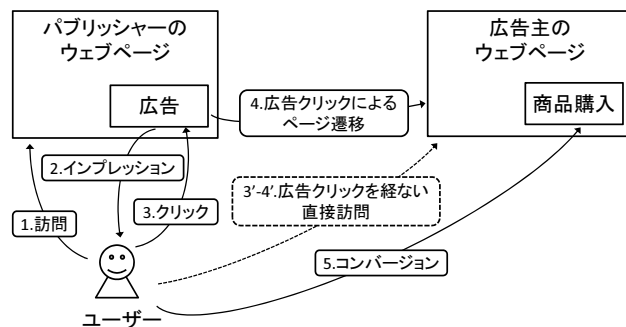


図 1 ユーザーから見たインプレッション、クリック、コンバージョンの流れ。(1) ユーザーはパブリッシャーのウェブページを訪れ、(2) ページ上に掲載された広告を閲覧する。(3) 興味を持ったユーザーが広告をクリックすると、(4) 広告主のウェブページに遷移する。(5) ユーザーが広告主のページ上で、商品購入などの広告主が定義した行動をとるとコンバージョンとなる。(3'-4')のように広告閲覧後、クリックを経ずにウェブ検索などで直接広告主のウェブページを訪れコンバージョンに至ることもあるが、本稿ではこれは扱わない。

ワードに対して入札設定を行っている広告集合の中で競われる。1 章で述べたように、ターゲティング設定を行っていた場合、ユーザーや時間帯などに対する条件が満たされなければオークションへの参加は行われない。一般的に検索結果画面には複数の広告表示枠があり、上で述べたように期待収益が高い順に、表示枠に広告が掲載される。ユーザーが広告をクリックするとパブリッシャーは広告主に課金を行い、課金額は広告主が提示した入札額をもとに決定される。

近年盛んになった RTB では、広告主は上の例のように配信対象と入札額をあらかじめ選択しておくのではなく、インプレッションごとにリアルタイムでオークションへの入札額を決定する。この入札額の決定は、パブリッシャー側から提供される、広告が掲載されるウェブページや閲覧するユーザーの情報をもとに行われる。この仕組みにより、広告主は微細な粒度で入札額を決めることが可能になった。

広告効果を高めるためには対象のインプレッションで得られる広告効果に基づいて、オークションへの入札額を決定することが必要になる。ショッピングサイトを運営する広告主の例では、商品購入につながりやすいインプレッションには高い入札額を提示し、逆に期待が薄いインプレッションの入札額を 0 にすることでオークションに参加しないという手続きが考えられる。つまり、予測した CVR の値によって入札額を変化させることにより、広告効果を最適化することができる。

オークションの流れを図 2 にまとめた。近年では、広告主に代わり広告効果を最適化する demand-side platform (DSP) や、パブリッシャーに代わり収益の最大化を行う supply-side platform (SSP) も見られるようになった。

2.2 関連研究

CVR 予測と関連した問題として CTR 予測を挙げることができる。2.1 節で述べたように、クリック課金型広告の期待収益の計算には CTR 予測が不可欠である。そのため、この課金

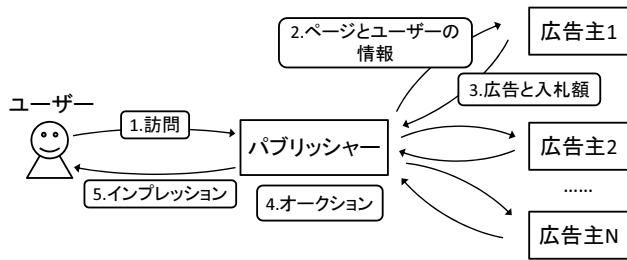


図2 オークションの流れ. (1) ユーザーがパブリッシャーのウェブページを訪問すると、パブリッシャーはそのページとユーザーへの広告表示に対してオークションを開催する. (2) パブリッシャーが各広告主にページとユーザーの情報を送信すると、(3) 広告主はその情報をもとに表示する広告とオークションへの入札額を提示する. (4) パブリッシャーは各広告主が提示した入札額をもとにオークションを行い、それに勝利した広告が表示機会を獲得し、(5) ユーザーの目に触れることになる.

方式が広く使われている検索連動型広告やコンテキスト広告に対して多くの研究がされてきた [4], [13], [17]. 一般的に、広告の CTR 予測は、過去のクリックログを用いて学習された統計モデルによって行われる. これらのモデルの例として、ロジスティック回帰 [5], [6] やプロビット回帰 [9], Gradient boosted decision trees [7], [18] が挙げられる.

近年、DSP や RTB による広告取引が盛んになるに伴い、コンバージョン予測の研究も多くされるようになった. Agarwal ら [1] は階層構造を持つ素性を用いた Log-linear Model for Multiple Hierarchies (LMMH) を提案し、ディスプレイ広告の CTR 予測や CVR 予測に適用した. Rosales ら [16] はディスプレイ広告におけるコンバージョン予測の問題を扱い、クリック予測とコンバージョン予測の比較を行った. Lee ら [10] はディスプレイ広告における CVR 予測の問題に対して、過去の実績 CVR 値を素性としてロジスティック回帰モデルを構築する手法を用いた. Rey ら [15] は予測した CVR を用いて、検索連動型広告の入札額を調整する手法を提案した. Bagherjeiran ら [2] はコンテキスト広告において、広告インプレッションからクリックを経てコンバージョンまで至る、コンバージョンファネルを最適化する手法を提案した. Liu ら [11] はディスプレイ広告において CVR が高いと予想されるユーザーを見つける手法を提案した. Perlich ら [14] はディスプレイ広告の RTB 取引におけるコンバージョン予測の問題に対して、広告主のサイト訪問を用いた転移学習のアプローチを提案した.

3. データセットと素性

この章では、本稿で扱うデータセットと素性について述べる.

3.1 データセット

評価には、『Yahoo!ディスプレイアドネットワーク (YDN)』^(注1) の実際のシステムログを 10 週間分用いた. なお、このデータセットは、広告自体がテキストからなるテキスト広告と、画像からなる画像広告の、二種類の広告配信ログが含まれている.

表1 データセットのサンプル数

	学習データ	バリデーションデータ	テストデータ
正例数	1,803,658	699,437	740,853
負例数	5,236,800	1,902,321	1,902,316
合計	7,040,458	2,601,758	2,643,169

表2 素性グループ

素性グループ	素性詳細
広告主	広告 ID, 広告グループ ID, キャンペーン ID, 広告主 ID, 画像 ID, 画像サイズ, ランディングページの URLID, カテゴリ, ターゲティング条件
パブリッシャー	サイト ID
ユーザー	性別, 年代, 地域, デバイス種別, 興味カテゴリ
その他	商品コード, マッチング種別, ウェブページと広告の類似度, ユーザーと広告の関連度 (2 種)

この広告配信ログの前半 6 週間分を学習データ、続く 2 週間分をバリデーションデータ、後半 2 週間分をテストデータとして扱った. 1 章で述べたように、本稿では広告クリックを経たコンバージョンを扱うため、クリックされた広告のデータを用いる. データ中の各サンプルはクリックされた広告インプレッション一つに対応しており、コンバージョンしたか否かがラベル付けされている. クリックされた後一定期間中にコンバージョンに至ったインプレッションを正例とし、そうでないものを負例として扱った. データは同じ規則にしたがい、フィルタリングとサンプリングを行った. その結果、負例の量は正例のおよそ 3 倍となった. 得られたデータセットのサンプル数は表 1 にまとめている.

3.2 素性

素性は種類に応じてグループ分けを行った. それぞれのグループの詳細は表 2 にまとめている.

広告主に紐づく素性としては広告の階層構造情報に加え、画像広告の場合は画像の情報を用いた. 広告の階層構造は、広告主、キャンペーン、広告グループ、広告の順に粒度が細くなる. それぞれの ID を素性として用いた. 画像広告の場合は画像自体の ID とサイズを追加で用いた. コンバージョンに至るかどうかは、広告クリック後に遷移するランディングページの影響があると考えられるため、その URL の ID を用いた. パブリッシャーに紐づく素性としてはサイトの ID を用いた. ユーザーに紐づく素性としてはアクセスしたデバイス種別に加えて、可能であれば性別、年代、地域情報、ウェブ上での行動履歴をもとに推定した興味カテゴリを用いた. その他の素性として、各広告の商品コードとマッチング種別を用いた. 商品コードは、ビジネス上の広告商品の区分を表すものであり、YDN では、インタレストマッチ、ターゲティング、モバイル、サイトリターゲティングなどがある. マッチング種別は、上記のインタレストマッチ広告において、広告とインターネットユーザーの興味関心の関連度を図る際の軸として用いられるものである. 加え

(注1) : <http://promotionalads.yahoo.co.jp/service/ydn/index.html>

表 3 ID 素性のユニーク数

	学習データ	バリデーションデータ	テストデータ
広告 ID	284,561	159,907	164,007
広告グループ ID	199,354	117,141	120,380
キャンペーン ID	60,630	42,510	43,387
広告主 ID	21,555	15,804	16,158
画像 ID	21,124	14,202	14,714
URLID	155,768	91,750	95,846
サイト ID	2,387	2,074	2,083

て、それらのマッチングで計算されるスコアの中から、ウェブページと広告の類似度と、2 種類のユーザーと広告の関連度を用いた。

データに含まれる各 ID 素性のユニーク数は表 3 にまとめている。また、学習データには含まれなかった ID が、バリデーションデータやテストデータに含まれる割合を図 3 と図 4 に示した。ユニーク数とサンプル数の両方で、日がたつにつれて新規 ID の割合が増えていくことがわかる。特に、粒度が細かい広告 ID の増加率が大きく、2 週間経過するとユニーク ID の割合で約 30%、サンプルの割合で約 25%が未知の ID となった。増加率の小さい広告主 ID でも、2 週間後にはユニーク ID の割合で約 10%、サンプルの割合で約 5%が新規の ID であった。画像 ID は、ユニーク数と比較してサンプル数における未知の割合が少なかった。これは、画像 ID は画像広告でのみ変化し、テキスト広告では常に同じ ID をとるためである。これらの結果から、バリデーションデータやテストデータには学習データに現れなかった素性が無視できない割合で出現することが分かる。

4. 予測モデル

3 章で示したデータセットと素性を用いて CVR 予測モデルを構築した。CVR 予測モデルとしては、以下の式で表現されるロジスティック回帰モデルを用いた。

$$p(c | a, p, u) = \frac{1}{1 + \exp(-c(\mathbf{w}^T \mathbf{x}(a, p, u)))}$$

なお、 $c \in \{+1, -1\}$ はコンバージョンしたか否かを表す変数であり、 $c = +1$ の時にコンバージョンしたことを、 $c = -1$ の時にコンバージョンしなかったことを表す。また、 $p(c = +1 | a, p, u)$ は広告 a 、パブリッシャー p 、ユーザー u が与えられた時に、コンバージョンする確率を表す。 $\mathbf{x}(a, p, u)$ はその 3 つから抽出された素性ベクトルを、 \mathbf{w} はその素性に対応する重みベクトルを表現している。

過学習を避けるため正則化項として $\|\mathbf{w}\|_2^2/2$ を加え、以下の最適化問題を解き $\hat{\mathbf{w}}$ を得た。

$$\hat{\mathbf{w}} = \arg \min_{\mathbf{w}} \frac{1}{2} \|\mathbf{w}\|_2^2 + C \sum_{i=1}^N \log(1 + \exp(-c_i(\mathbf{w}^T \mathbf{x}_i(a, p, u))))$$

ここで $\|\mathbf{w}\|_2$ はユークリッドノルム (L2 ノルム)、 $C > 0$ は正則化パラメータである。 C を変化させて学習データで学習を行い、バリデーションデータでの評価値が高い $\hat{\mathbf{w}}$ を用いてテストデータの評価を行った。

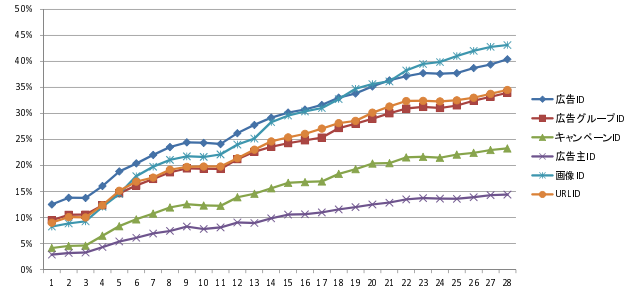


図 3 バリデーションデータとテストデータにおける、学習データに含まれていない新規 ID のユニーク数の割合。横軸は学習データの最終日からの経過日数。

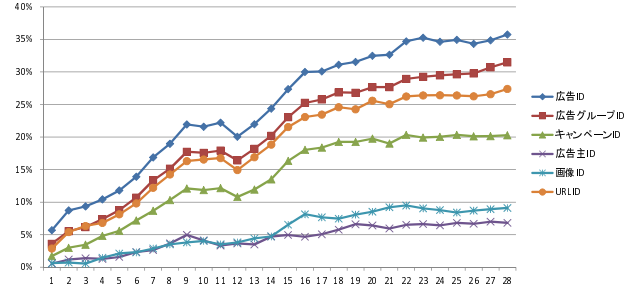


図 4 バリデーションデータとテストデータにおける、学習データに含まれていない新規 ID を含むサンプルの割合。横軸は学習データの最終日からの経過日数。

評価では、Chapelle ら [4] の手法に倣い、素性ベクトル $\mathbf{x}(a, p, u)$ の次元数を hashing trick [19] を用いて制限した。ハッシュ値として 24 ビット整数を用い、素性ベクトルの次元数をおよそ 1,600 万とした。

5. 実験

この章では、用いる素性やモデル、評価指標を変えて実験を行う。なお、3.2 節で述べたように、ID 素性を用いたモデルは学習時に現れなかった素性が評価時に出現し、時間経過によって劣化することが予想されるため、テストデータに対する評価値と合わせてバリデーションデータに対する評価値も併記し比較を行う。

5.1 AUC による評価

まずはじめに、一般的な二値分類の問題設定で評価を行った。各素性グループごとに CVR 予測モデルを構築し、area under the ROC curve(AUC) [12] を評価として用いた。結果が表 4 である。なお、表中の値は以下の式による正規化を行った。

$$\overline{AUC} = \frac{AUC - 0.5}{AUC_{max} - 0.5}$$

結果を見ると、すべての素性を用いたモデルの AUC が一番高かった。素性グループ間での比較を行うと、広告主情報を用いた場合に高い AUC を示した。これは、そもそもコンバージョンの定義が広告主やキャンペーンによって異なるためだと考えられる。ただし、広告主素性を用いたモデルよりも、すべての素性を用いたモデルの評価値がわずかながら高いため、広告系

表4 各素性グループを用いた場合の \overline{AUC}

素性グループ	バリデーションデータ	テストデータ
広告主	0.9819	0.9220
パブリッシャー	0.4471	0.3566
ユーザー	0.3733	0.3256
その他	0.4327	0.5092
上記すべて	1.0000	0.9582

の素性以外にも有用な素性があると考えられる。素性グループ間の比較では広告主素性に次いで、その他グループの素性を用いた場合、パブリッシャー情報を用いた場合の順でテストデータにおける評価値が高かった。単純に考えると、これら二つの序列は「どのサイトで広告が表示されたか」よりも「どのような広告商品か」が予測に強い影響を与えていることを示唆する。しかしながら、それら二つの素性グループは広告主素性と相関があるため、この結果のみでそのように考えるのは性急である。なお、この二つの序列はバリデーションデータでは逆転した。その他の素性グループのテストデータにおける評価値がバリデーションデータの値から向上している一方で、その他の三つの素性グループでは悪化した。ユーザー情報を素性として用いた場合の AUC はふるわなかった。

既に述べたように、広告主やキャンペーンごとにコンバージョンの定義が異なり、その難易度は異なる。例えば、あるキャンペーンでは特定のページへのアクセスをコンバージョンとみなすために、広告をクリックしたユーザーの大多数がコンバージョンに至る一方で、別のキャンペーンでは高額商品をオンラインで購入することをコンバージョンと見なすために、コンバージョンに至るユーザーが非常にまれ、ということがある。このような場合、上記の評価で用いたような、評価データ全てに対する AUC を評価指標とすると、コンバージョンしやすいキャンペーンのサンプルの予測値を高く、コンバージョンしにくいキャンペーンのサンプルの予測値を低くするだけで評価値が高くなる。そこで、次節ではキャンペーンごとの評価指標を用いることにする。

5.2 広告キャンペーンごとの AUC による評価

評価指標として以下の式で定義される、広告キャンペーンごとの AUC をサンプル数で重み付けした値 $wAUC$ を用いる。

$$wAUC = \frac{\sum_{camp} AUC_{camp} \times \#sample_{camp}}{\sum_{camp} \#sample_{camp}}$$

なお、正例と負例の両方を含まないキャンペーンに対しては AUC の値が未定義であるため、 $wAUC$ の計算の候補集合に含めなかった。この手続きはバリデーションデータとテストデータで独立して行った。そのため、バリデーションデータでは $wAUC$ の計算対象に含まれなかったキャンペーンが、テストデータの評価では含まれることもある。学習時には、評価の対象に含まれるか否かには関係なく、学習データ全てを用いて学習を行った。AUC での評価結果と同じく、報告する値は全て以下の式を用いて正規化を行った。

$$\overline{wAUC} = \frac{wAUC - 0.5}{wAUC_{max} - 0.5}$$

表5 各素性グループを用いた場合の \overline{wAUC}

素性グループ	バリデーションデータ	テストデータ
広告主	0.2507	0.2710
パブリッシャー	0.5584	0.4907
ユーザー	0.3622	0.3776
その他	0.2959	0.3061
上記すべて	0.9015	0.8509

表6 広告主素性を用いた場合の \overline{wAUC}

素性詳細	バリデーションデータ	テストデータ
広告 ID	0.2315	0.1923
広告グループ ID	0.1852	0.1680
URLID	0.1662	0.0797
広告主素性すべて	0.2507	0.2710

表7 組み合わせ素性を素性とした場合の \overline{wAUC}

素性グループ	バリデーションデータ	テストデータ
組み合わせ素性なし	0.9015	0.8509
× 広告 ID	0.9577	0.8607
× 広告グループ ID	0.9693	0.8760
× キャンペーン ID	0.9778	0.8968
× 広告主 ID	1.0000	0.9457
× 広告主素性すべて	0.9987	0.9170

$wAUC$ を評価値として用いた場合の結果が表5である。5.1節での評価結果と同じように、すべての素性を用いたモデルの評価値が一番高かった。素性グループ間での比較では、パブリッシャー情報を用いたモデルが高い評価値を示した。これは、広告キャンペーンに着目した場合は、広告の配信面によってコンバージョン率に違いがあるということを表している。次いで、素性グループ間での比較では、ユーザー情報を用いた場合、その他素性を用いた場合、広告主情報を併用した場合と続いた。評価データ全体に対する AUC を評価指標として用いた場合とは異なり、その他グループの素性を用いた場合の評価値はやや低い結果となった。

表6に、広告主の素性グループの各素性のみを用いた場合の結果を示す。なお、 $wAUC$ の評価単位であるキャンペーン ID や、それより粒度の粗い広告主 ID のみを用いた場合は、各キャンペーン内の予測値に差がつかないため、評価から外した。また、画像 ID や画像サイズのみを用いた場合は、テキスト広告の予測値に差がつかず、 \overline{wAUC} の値がほぼ 0 に近い値となったため、表から省いた。広告 ID と広告グループ ID を併用した場合を比較すると、より細かい粒度である広告 ID を併用した場合の評価値が高くなった。広告主のウェブページを表す URLID は、コンバージョンの予測にはそれほど寄与していない結果となった。広告グループ ID のみを用いた場合と、広告主素性すべてを用いた場合の評価値には開きがあるため、それぞれの素性は補完しあっていると考えられる。

5.3 組み合わせ素性を用いた場合の評価

5.2節で導入した $wAUC$ は、広告キャンペーンごとの AUC をサンプル数で重みづけした値であるため、それぞれのキャンペーンごとにモデルを作成することで精度の向上が見込まれ

表 8 L1 正則化による素性選択を行った場合の \overline{wAUC}

素性グループ	バリデーションデータ		テストデータ		ノンゼロパラメータ数	
	L2	L1	L2	L1	L2	L1
組み合わせ素性なし	0.9015	0.8823	0.8509	0.8423	727,786	9,321
× 広告 ID	0.9577	0.9247	0.8607	0.8818	15,516,966	51,901
× 広告グループ ID	0.9693	0.9313	0.8760	0.8822	14,642,865	57,595
× キャンペーン ID	0.9778	0.9526	0.8968	0.8938	11,419,419	64,720
× 広告主 ID	1.0000	0.9746	0.9457	0.9315	6,950,731	65,843
× 広告主素性すべて	0.9987	0.9717	0.9170	0.9059	16,772,878	72,951

る。しかしながら、図 3 と図 4 で示したように、学習データに出現しなかったキャンペーンのサンプルが、テストデータには 15%以上含まれるため、学習データをキャンペーンごとに分割してモデルを作成する単純なアプローチでは、新しいキャンペーンの予測がうまくいかない。そのため、Chapelle ら [4] の手法に倣い、組み合わせ素性を用いたマルチタスク学習 [8] を行うアプローチをとった。5.2 節で評価した、すべての素性を用いた場合に対して、広告主素性とそれ以外の三つの素性グループの間で組み合わせ素性を作成して追加した。これにより、未知の ID を持つ広告の予測にはグローバルなモデルを用い、既知の ID に対してはグローバルなモデルの予測値を ID ごとに補正するようになる。結果が表 7 である。組み合わせ素性を追加することで、評価値の向上が見られた。組み合わせ素性を作成する ID 素性を比較すると、粒度が粗い順に評価値が高くなった。これは粒度が細かい広告 ID などの素性は数が多く、組み合わせ素性の数も膨大になり、過学習を起こしているのだと考えられる。広告主素性すべてを用いて組み合わせ素性を追加した場合と、広告主 ID との組み合わせ素性を追加した場合を比較すると、バリデーションデータではほぼ同じ評価値だったが、テストデータでは差がつく結果となった。

5.4 L1 正則化による素性選択を行った場合の評価

L1 正則化を用いた学習はスパースな解が得られるため、素性選択の手法として広く用いられている。L1 正則化を用いた学習では、以下の最適化問題を解く。

$$\hat{\mathbf{w}} = \arg \min_{\mathbf{w}} \|\mathbf{w}\|_1 + C \sum_{i=1}^N \log(1 + \exp(-c_i(\mathbf{w}^T \mathbf{x}_i(a, p, u))))$$

ただし $\|\mathbf{w}\|_1$ は L1 ノルムであり、この式は 4 章の式の正則化項を置き換えた形となっている。

本節では、L1 正則化を用いた評価を行い、値が 0 でないパラメータの数と予測精度のトレードオフを検証する。結果が表 8 である。L1 正則化を行うことで、多くのパラメータの値が 0 になり、ノンゼロパラメータの数は非常に少なくなり、L2 正則化を用いた場合と比較して、0.3~1.3%程度となった。しかしながら、今回のケースでは hashing trick により、素性ベクトルの次元を 24 ビット空間に制限しており、単精度浮動小数点数の配列で表現した場合、およそ 1,600 万の素性で使用メモリは 64 メガバイトと近年の一般的なメモリサイズと比較してそれほど大きくないため、このメモリサイズの削減によって得られる恩恵は限定的である。それぞれの正則化を用いた場合の評価値を比較すると、バリデーションデータではすべての素性セットで

L2 正則化の評価値が上回っていた一方で、テストデータでは粒度の細かい広告 ID と広告グループ ID で組み合わせ素性を作成し追加した場合に、L1 正則化の方が高い評価値を示した。

6. まとめと今後の課題

本稿では、オンライン広告における CVR 予測モデルの素性評価を行った。CVR 予測は先んじて研究されてきた CTR 予測とよく似た問題だが、コンバージョンはクリックと比較して数が少なく、広告キャンペーンによって定義が異なるという違いがある。そこで、ID 素性を用いた CVR 予測モデルを『Yahoo! ディスプレイアドネットワーク』の配信システムログで検証した。ID 素性を用いたモデルは、時間経過に伴い学習時には現れなかった未知の ID の割合が増え、予測精度が劣化してしまうという問題がある。この問題への対処法の一つが、組み合わせ素性を用いたマルチタスク学習である。ただし、組み合わせ素性は数が多くなるため、学習時の正則化の選択も重要になる。

今後の課題としては、より多くの素性を追加した評価や、異なるモデルを用いた場合の評価を挙げることができる。

文 献

- [1] Deepak Agarwal, Rahul Agrawal, Rajiv Khanna, and Nagaraj Kota. Estimating rates of rare events with multiple hierarchies through scalable log-linear models. In *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining*, KDD '10, 2010.
- [2] Abraham Bagherjeiran, Andrew O. Hatch, and Adwait Ratnaparkhi. Ranking for the conversion funnel. In *Proceedings of the 33rd international ACM SIGIR conference on Research and development in information retrieval*, SIGIR '10, 2010.
- [3] Andrei Broder and Vanja Josifovski. Introduction to computational advertising. <http://www.stanford.edu/class/msande239/>. Accessed: 14/2/2014.
- [4] Olivier Chapelle, Eren Manavoglu, and Romer Rosales. Simple and scalable response prediction for display advertising. *ACM Trans. Intell. Syst. Technol.*, 2013.
- [5] Haibin Cheng and Erick Cantú-Paz. Personalized click prediction in sponsored search. In *Proceedings of the third ACM international conference on Web search and data mining*, WSDM '10, 2010.
- [6] Haibin Cheng, Roelof van Zwol, Javad Azimi, Eren Manavoglu, Ruofei Zhang, Yang Zhou, and Vidhya Navalpakkam. Multimedia features for click prediction of new ads in display advertising. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*, KDD '12, 2012.

- [7] Kushal S. Dave and Vasudeva Varma. Learning the click-through rate for rare/new ads from similar ads. In *Proceedings of the 33rd international ACM SIGIR conference on Research and development in information retrieval*, SIGIR '10, 2010.
- [8] Theodoros Evgeniou and Massimiliano Pontil. Regularized multi-task learning. In *Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '04, pp. 109–117, 2004.
- [9] Thore Graepel, Joaquin Quinonero Candela, Thomas Borchert, and Ralf Herbrich. Web-scale bayesian click-through rate prediction for sponsored search advertising in microsoft's bing search engine. In *Proceedings of the 27th International Conference on Machine Learning*, 2010.
- [10] Kuang-Chih Lee, Burkay Orten, Ali Dasdan, and Wentong Li. Estimating conversion rate in display advertising from past performance data. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*, KDD '12, 2012.
- [11] Yandong Liu, Sandeep Pandey, Deepak Agarwal, and Vanja Josifovski. Finding the right consumer: Optimizing for conversion in display advertising campaigns. In *Proceedings of the Fifth ACM International Conference on Web Search and Data Mining*, WSDM '12, pp. 473–482, 2012.
- [12] Christopher D. Manning, Prabhakar Raghavan, and Hinrich Schtze. *Introduction to Information Retrieval*. Cambridge University Press, 2008.
- [13] H. Brendan McMahan, Gary Holt, D. Sculley, Michael Young, Dietmar Ebner, Julian Grady, Lan Nie, Todd Phillips, Eugene Davydov, Daniel Golovin, Sharat Chikkerur, Dan Liu, Martin Wattenberg, Arnar Mar Hrafnkelsson, Tom Boulos, and Jeremy Kubica. Ad click prediction: A view from the trenches. In *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '13, pp. 1222–1230, 2013.
- [14] C. Perlich, B. Dalessandro, T. Raeder, O. Stitelman, and F. Provost. Machine learning for targeted display advertising: transfer learning in action. *Machine Learning*, pp. 1–25, 2013.
- [15] Benjamin Rey and Ashvin Kannan. Conversion rate based bid adjustment for sponsored search. In *Proceedings of the 19th International Conference on World Wide Web*, WWW '10, pp. 1173–1174, 2010.
- [16] Rómer Rosales, Haibin Cheng, and Eren Manavoglu. Post-click conversion modeling and analysis for non-guaranteed delivery display advertising. In *Proceedings of the fifth ACM international conference on Web search and data mining*, WSDM '12, 2012.
- [17] Yukihiro Tagami, Shingo Ono, Koji Yamamoto, Koji Tsukamoto, and Akira Tajima. CTR prediction for contextual advertising: Learning-to-rank approach. In *Proceedings of the Seventh International Workshop on Data Mining for Online Advertising*, ADKDD '13, pp. 4:1–4:8, 2013.
- [18] Ilya Trofimov, Anna Kornetova, and Valery Topinskiy. Using boosted trees for click-through rate prediction for sponsored search. In *Proceedings of the Sixth International Workshop on Data Mining for Online Advertising and Internet Economy*, ADKDD '12, 2012.
- [19] Kilian Weinberger, Anirban Dasgupta, John Langford, Alex Smola, and Josh Attenberg. Feature hashing for large scale multitask learning. In *Proceedings of the 26th Annual International Conference on Machine Learning*, ICML '09, pp. 1113–1120, 2009.
- [20] Shuai Yuan, Ahmad Zainal Abidin, Marc Sloan, and Jun Wang. Internet advertising: An interplay among advertisers, online publishers, ad exchanges and web users. *CoRR*, 2012.
- [21] Shuai Yuan, Jun Wang, and Xiaoxue Zhao. Real-time bidding for online advertising: Measurement and analysis. In *Proceedings of the Seventh International Workshop on Data Mining for Online Advertising*, ADKDD '13, 2013.