

Machine Learning

機械学習



機械学習の種類

1. 教師あり学習
2. 教師なし学習
3. 半教師あり学習
4. 強化学習



教師あり学習VS教師なし学習

教師あり学習の問題点

- データのラベル付けは高コスト
 - しかし、教師データの数が少ない場合にはデータに偏りが生じやすく、偏ったデータから検出対象クラスの物体の特徴を学習することは困難です。
 - 負例が全くない、人によってデータのラベルの信頼度が異なる、ラベルの内容が広すぎる, etc.
 - 即ち、データを得る者は天下を取り! e.g. Google, Amazon, Alibaba, etc.

教師なし学習の問題点

- 学習成果の妥当性を評価することは困難である。
 - その故、教師なし学習の研究は教師あり学習の研究より少ないし、教師あり学習の研究を支援するために研究された感じがする。

教師あり学習の改善



Few-Shot Learning -> Meta Learning

Few-Shot Learning

- データ: (x, y) (few)

x はデータ(ベクトル)、 y はラベル

- 目的:

事前知識を用いて、少ない教師あり情報から、 x から y への写像(関数)を学習

- 例: Siamese Network(シヤムネットワーク)、等々

詳しく説明: <https://qiita.com/ell/items/9e9de65521c8b935d28f>



弱教師付き学習(Weakly Supervised Learning)

Weakly Supervised Learning

- ・ データ: (x, y) (ラベルには様々な問題点がある: 弱いデータ)

x はデータ(ベクトル)、 y はラベル

- ・ 目的:

弱いデータを用いて、 x から y への写像(関数)を学習

- ・ 例: PU、PNU、Pconf、補ラベル学習、等々

詳しく説明: <https://lionbridge.ai/ja/articles/weakly-supervised-learning/>

各社の機械学習事例



オンライン広告

スマートフォン、パソコンで表示される広告

1. ディスプレイ広告
2. ビデオ(動画)広告
3. 検索連動型広告

日本国内市場規模: 2018年1兆7589億円

<https://dentsu-ho.com/articles/6572>

ディスプレイ広告

動画 知恵袋 地図 リアルタイム 求人 一覧▼

Q 検索

書籍購入で最大50%戻ってくる » 自宅でソフトバンクのiPhoneをお得に申し込み

コンテンツ スポーツ 国内 国際 IT・科学 地域

、88歳 **NEW** ♫ 401
うれしい ♫ 2172
り警告 **NEW** ♫ 227
波」不安 **NEW** ♫ 275
ーと明暗 ♫ 1053
でも明暗 ♫ 1120
減カラー **NEW** ♫ 140
い 原監督 ♫ 499
クス一覧

金色の道
11/14(土) 18:54
毎日新聞

PayPayで買うと **超回祭**
最大 25% 相当※
戻ってくる!!
※付与条件・上限があります。詳細は遷移先をご確認ください。
25%は15日(日)のPayPayモールが対象です。※画像はイメージです。

PayPay モール YAHOO! JAPAN ショッピング

Yahoo!ニュース」をク

stText

スタリングしようとした。
微妙な結果となってしまいました。
タリを用いて、Yahoo!ニュース記事のクラスタリングを行

ソースの自然言語処理ライブラリです。

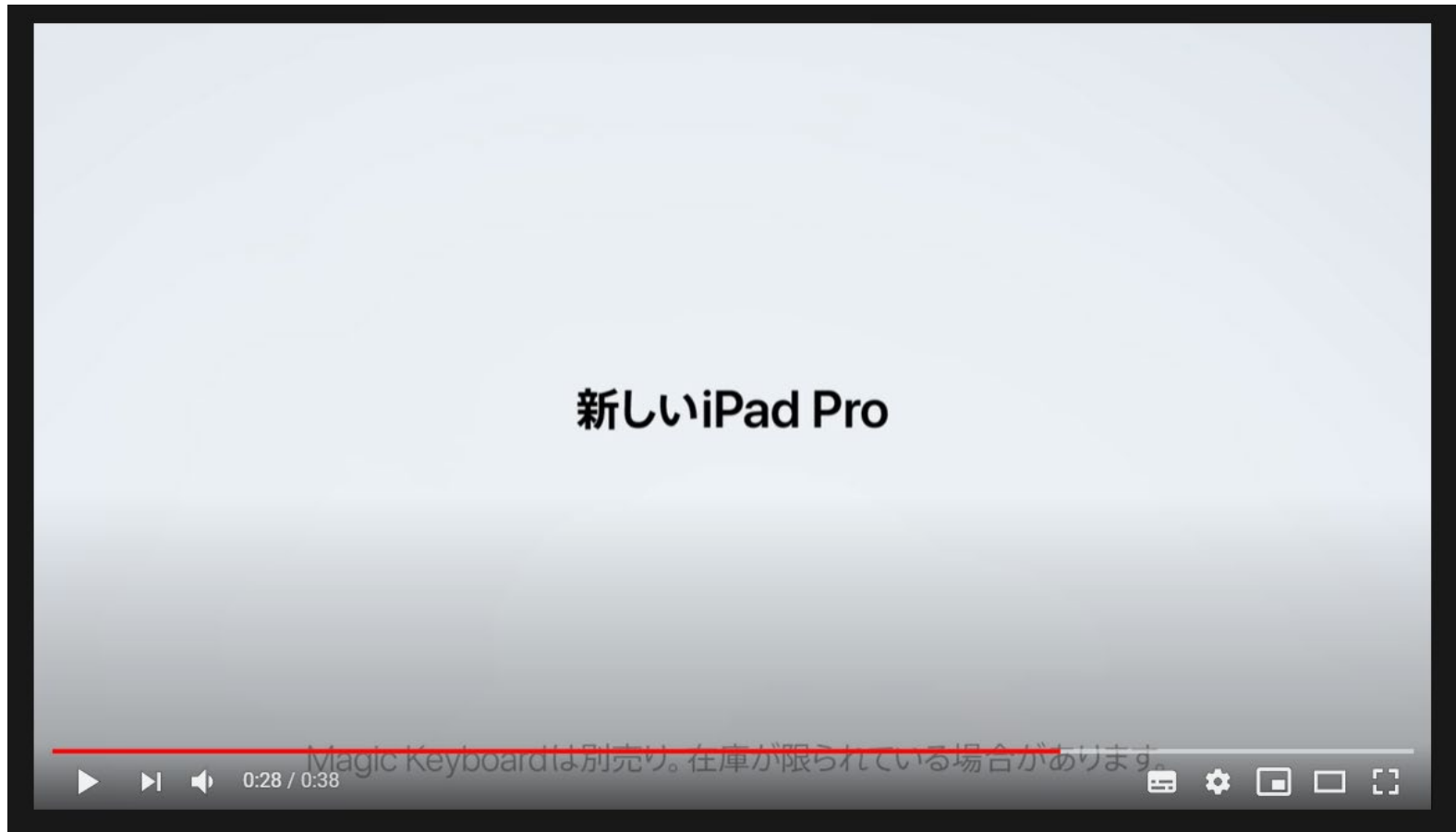
企業があなたを奪い合う!!
11.18
転職ドラフト開催!!
二> 転職 DRAFT

こだわり抜きました。PHP
専門の
プログラミングスクール

【18〜32歳限定】「非常識な方法」で学ぶ
未経験から23日間で、PHPエンジニアに!
dt30.net

開く

ビデオ(動画)広告



検索連動型広告

Google

引っ越し

× | 🗣️ 🔍

🔍 [すべて](#) 📰 [ニュース](#) 📍 [地図](#) 🖼️ [画像](#) 📺 [動画](#) ⋮ [もっと見る](#) ⚙️ [設定](#) 🛠️ [ツール](#)

約 130,000,000 件 (0.59 秒)

広告 · hikkoshi.kakaku.com/引っ越し/簡単見積もり ▼

【価格.com】安い引っ越し業者探し - 引っ越し見積もり比較サイト

たった3分で無料**引っ越し**見積もり！さらにご契約成立で今なら10万円が当たるキャンペーン実施中。**引っ越し**料金相場や**引っ越し**業者のクチコミ、料金、満足度ランキングなど役立つ情報満載！
価格.comの**引っ越し**比較・最安値を一括見積もり・最大50%オフ。

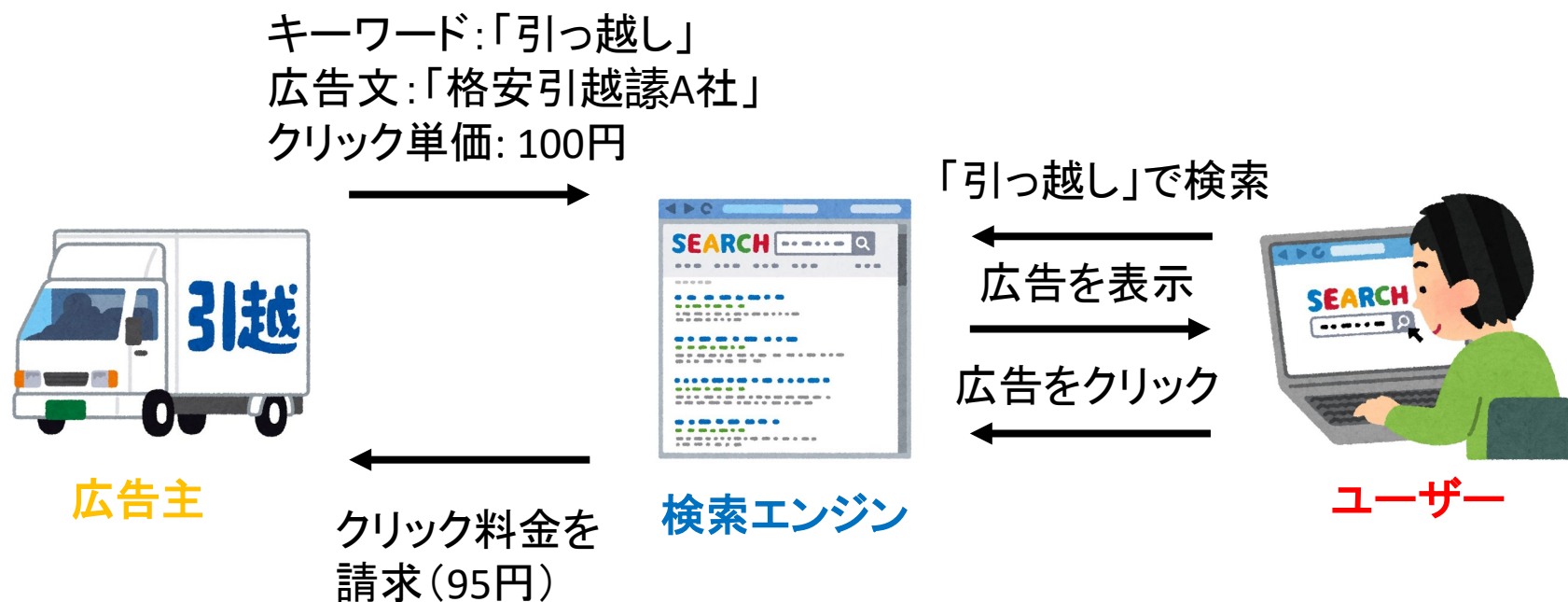
引っ越し業者ランキング 引っ越し業者選びで迷った方へ みんなのランキングをぜひ参考に	単身の引っ越し相場料金 単身の引っ越し相場を知って 無料一括見積もりでお安く引っ越し
引っ越し料金相場 早見表 単身、家族など引っ越し料金相場が 一目でわかる！平均相場の一覧表	家族の引っ越し相場料金 家族の引っ越し相場を知って 無料一括見積もりでお安く引っ越し

広告 · www.ohikkoshi.net/ ▼

11～12月 1 番安い引っ越し業者は？ - 2020年／引っ越し料金一括比較

大手業者から地域密着業者まで約100社厳選。同条件で比較するから安い**引っ越し**業者が見つかる。複数社見積もりで**引っ越し**料金費用を最大50%オフに出来るかも。利用者限定特典が充実。24時間ネットで一括見積り・最大"半額"も費用を節約。

検索連動型広告の基本



検索連動型広告とオークション



キーワード:「引っ越し」
広告文:「**格安**引越**誘**A社」
クリック単価: 100円



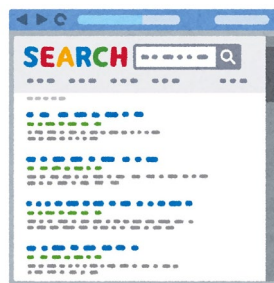
広告主A



キーワード「引っ越し」
広告「**引越**なら**B社**」
クリック単価: 120円



広告主B



検索エンジン

オークションによりランキング

ランキング外の広告は表示されない



キーワード「引っ越し」
広告「**引越**見積もり**無料**」
クリック単価: 95円



広告主C

検索連動型広告とオークション

- ・ クリック単価オークションの問題点:

1. 収入はクリックされた時のみ発生
2. 全くクリックされない広告が上位に表示され得る

- ・ 平均収入によるオークション:

平均収入 = クリック単価 × クリック率

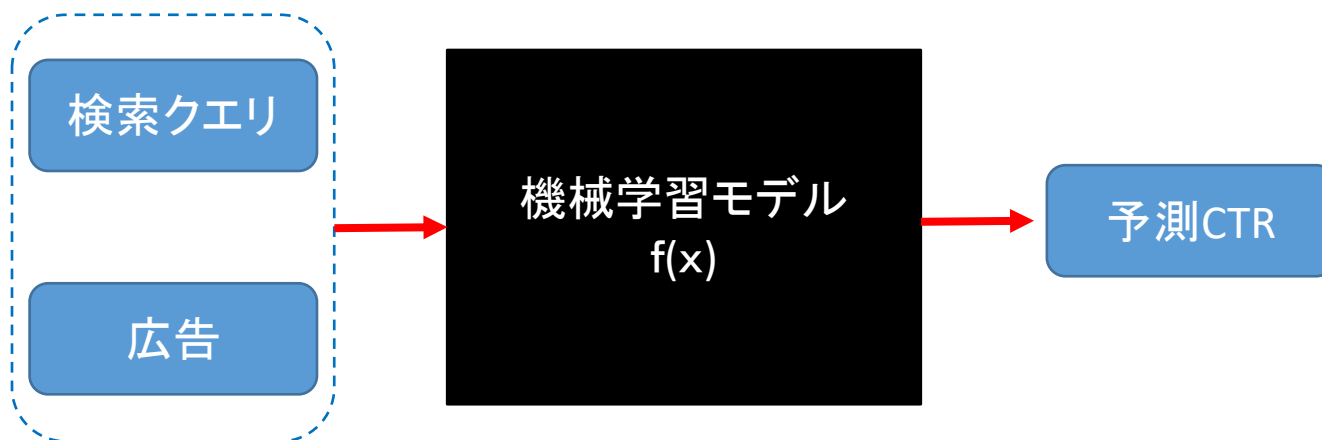
(クリック率はCTR (Click through rate)と呼ばれる)

CTRの予測

- 真のCTRは直接は観測できない
- 過去に表示された広告のクリック有無からCTRを予測する
- 予測CTRを用いてオークション

機械学習によるCTRの予測

- ・ 次の機械学習モデル $f(x)$ をデータから学習(教師あり学習)
- ・ 入力 x : <検索クエリ、広告>
- ・ 出力 y : 予測CTR



最適化問題としての機械学習

- N個の入力データ: $\{x_1, x_2, \dots, x_n\}$ (ベクトル)
- N個の出力データ: $\{y_1, y_2, \dots, y_n\}$
- 次の損失を最小にする $f(x)$ を求める

$$\sum_{n=1}^N \text{loss}(f(x_i), y_i)$$

- $\text{loss}()$ (損失関数) の決め方が重要
- $f()$ としてはニューラル・ネットワークがよく用いられる

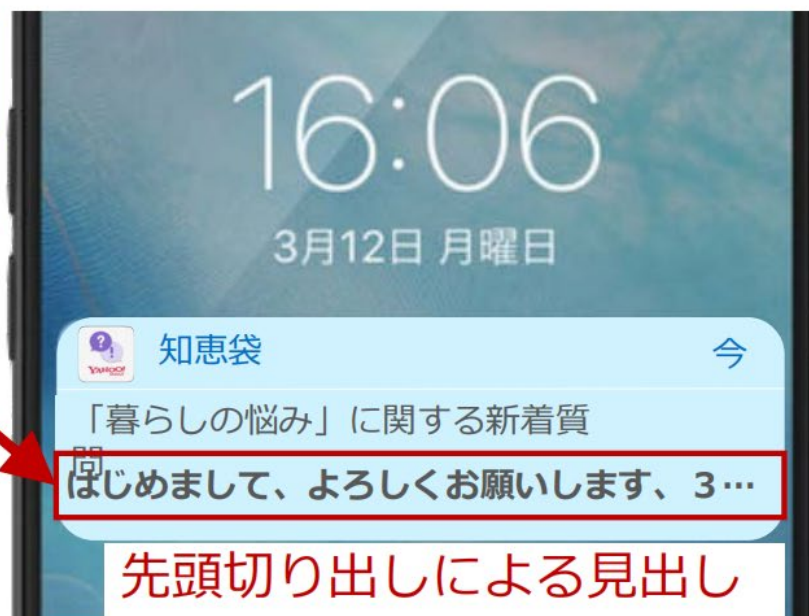
Yahoo!知恵袋の見出しの課題

- Yahoo!知恵袋の見出しは質問文の内容を表していない場合がある

質問投稿文




スマートフォンの通知画面



Yahoo!知恵袋の見出しの課題

- 1. 見出し候補となるスニペット見出し候補の一覧を抽出
- 2. ランク学習を利用して見出し候補一覧から適切な見出しを選択

 ID非公開

はじめまして、よろしくお願いします、30代の男性です。

アドバイスお願いします。

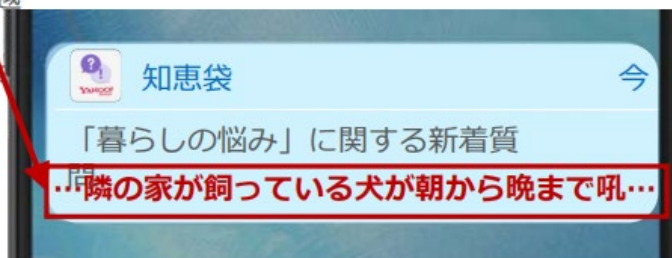
隣の家が飼っている犬が朝から晩まで吠え続けます。

近所の人たちも飼い主に注意を促しているのですが、改善が見られません。

地域一帯が借家ではなく、個人住宅なので引っ越すというわけにもいきません。

しかしこのまま耐えつづければノイローゼになってしまうそうです。

何とかすることはできないのでしょうか？



ユーザの行動からタイミングを最適化

