ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 11, November 2019

A Framework to Predict the Advertise View ability by using Machine Learning

Abhishek Gandhi[1], Mohit Gawale[1], Pranav Karmarkar[1], Tanay Gawali[1],

M. V. Sadaphule[2]

B. E Students, Department of Computer Engineering, Sinhgad College of Engineering, Pune, India[1]

Professor, Department of Computer Engineering, Sinhgad College of Engineering, Pune, India[2]

ABSTRACT: Advertisements play a vital role in every industry and they help in the growth of the industry. The advertisements that are published are not viewed properly by the user just because of insufficient clicking. By using the scroll depth process, we are able to predict the viewability of the advertisement based on the scroll percentage and also predict the maximum viewed advertisements. Online display advertising has become a billion-dollar industry, and it keeps growing. Advertisers attempt to sendmarketing messages to attract potential customers via graphic banner ads on publishers' webpages. Advertisers are charged for eachview of a page that delivers their display ads. However, recent studies have discovered that more than half of the ads are never shownon users' screens due to inefficient scrolling. Thus, advertisers waste a great amount of money on these ads that do not bring anyreturn on investment.

KEYWORDS: Computational Advertising, View-ability Prediction, User Behavior.

I. INTRODUCTION

Online display advertising has arisen as one of the most famous forms of advertising. Studies show that publishing advertising has generated earnings of over $98.2 billion in 2018. Internet advertising involves a publisher, who integrates advertisements into his online content, and an advertiser, who provides advertisements to be published. Display ads can be seen in a wide range of formats and contain items such as text, images, flash, video and audio. In display advertising, an advertiser wages a publisher for space on web pages to display a banner during page views in order to impress the visitors who are interested in his products. A page view happens each time a web page is requested by a customer and displayed on internet. One-time display of an advertisement in a page view is called an ad impression, and it is considered the basic unit of advertisement delivery. Advertiser's wages for the ad impressions with the expectation that their advertisements will be viewed, clicked on, or converted by customers (e.g., the ad results in a purchase). Old fashioned display advertisement compensation is mainly based on customer clicks and changes, because they bring direct income to the advertisers. Much research has been done for probabilistic analysis on the click rate and conversion rate, bid optimization and public sales of companies and their products. Certainly, customers like to purchase products from the varieties that they recognize and trust. Display advertisements can make an expressive experience that gets customers surprised about varieties and creates some trust.

To note this problem, another pricing model, which wages advertisements by the number of rendering imitations that a publisher has served, has become popular in the display advertising field. However, a modern study shows that more than half number of the imitations are actually not seen by customers because they do not scroll down a page enough to view the advertisements. Low viewability leads to an inefficient brand advertising. Therefore, a new pricing model is growing: pricing advertisement by the number of imitations that can be seen by a customer, instead of just being used. This eliminates the stress of the advertisers, who concern about paying for advertisements that were served but not viewed by customers. Unsurprisingly, the advertisements located at different page depths have different probabilities of being seen by a customer. Therefore, it is mandatory to judge the probability that an advertisement at a known page

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 11, November 2019

depth will be viewed on a customer's screen, and thus be measured as seen. The vertical page depth that a customer scrolls to is defined as the scroll depth.

1.1 Motivation

- Advertisements play a vital role in every industry and they help in the growth of the industry. The advertisements that are published are not viewed properly by the user just because of insufficient clicking.
- We don't have any platform that can provide us the standard web-page depth that is visited by each and every user that logs in your website.

II. RELATED WORK

In the [1] work, user suggested, "Word representations: A simple and general method for semi-supervised learning". We use near state-of-the-art supervised baselines, and find that each of the three-word representations improves the accuracy of these baselines. We find further improvements by combining different word representations. The disadvantage, however is that accuracy might not be as high as a semi-supervised method that includes task-specific information and that jointly learns the supervised and unsupervised tasks.

In the [2] work, The objective of this paper is to present the design of personalized click prediction models. An essential part of thesemodels is the development of new user-related features. Webase our features on observations over a significant volumeof search queries from a large number of users. Our observations suggest that user click behavior varies significantlywith regard to their demographic background, such as ageor gender. We investigate the click distribution for different users from various backgrounds and design a set of demographic features to model their group clicking patterns.Recognizing that there is still significant variability in demographic groups, we also investigate user-specific features.

In the [3] work, author states, "On the importance of initialization and momentum in deep learning". Stochastic gradient descent with momentum. In this paper, we show that when stochastic gradient descent with momentum uses a well-designed random initialization and a particular type of slowly increasing schedule for the momentum parameter, it can train both DNNs and RNNs (on datasets with long-term dependencies) to levels of performance that were previously achievable only with Hessian-Free optimization. We find that both the initialization and the momentum are crucial since poorly initialized networks cannot be trained with momentum and well-initialized networks perform markedly worse when the momentum is absent or poorly tuned.

In the [4] work, author states that, "The efficient backprop, in Neural networks: Tricks of the trade". The convergence of back-propagation learning is analysed so as to explain common phenomenon observed by practitioners. Many undesirable behaviours of backpropagation can be avoided with tricks that are rarely exposed in serious technical publications. This paper gives some of those tricks, and offers explanations of why they work. Many authors have suggested that second-order optimization methods are advantageous for neural net training. It is shown that most "classical" second-order methods are impractical for large neural networks. A few methods are proposed that do not have these limitations.

In the [8] work, author states that, "The sequential click prediction for sponsored search with recurrent neural networks". Click prediction is one of the fundamental problems in sponsored search. Most of the existing studies took advantage of machine learning approaches to predict ad clicks for each event of ad view independently. However, as observed in the real-world sponsored search system, user's behavior on ads yield high dependency on how the user behaved along with the past time, especially in terms of what queries she submitted, what ads she clicked or ignored, and how long she spent on the landing pages of clicked ads, etc. Inspired by these observations, we introduce a novel framework based on Recurrent Neural Networks (RNN). Compared to traditional methods, this framework directly models the dependency on user's sequential behavior into the click prediction process through the recurrent structure in

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 11, November 2019

RNN. Large scale evaluations on the click-through logs from a commercial search engine demonstrate that our approach can significantly improve the click prediction accuracy, compared to the sequence-independent approaches.

III. PROBLEM STATEMENT

The most important source of income for content publishers or website owners is the marketing of the products or things. As per the Google's payment models the content owners are paid if and only if the advertisement is viewed by the user. We do not have any model that predicts the depth where the advertisement can be viewed by the user.
Solving this issue can benefit online advertisers to allow them to invest more effectively in advertising and can benefit publishers to increase their revenue.

IV. PROPOSED METHOD

To classify the web-page page depth prediction we are going to use Expectation Maximization (EM) Algorithm. The EM algorithm is widely used to solve the maximum-likelihood parameter estimation problem.

- It can be used as the basis for unsupervised learning of clusters.
- It can be used for the purpose of estimating the parameters of Hidden Markov Model (HMM).
- It can be used for classifying the values of depth variables.
- The EM algorithm performs an expectation step (E-step) and a maximization step (M-step) Alternatively.
- Initially, a set of initial values of the parameters are considered. A set of observed data is given to the system that the observed data comes from our model.
- The next step is known as "Expectation"-step or E-step. In this step, we use the observed data in order to compare the variables.
- The next step is known as "Maximization"-step or M-step. In this step, we use the complete data to classify the data.

Architecture

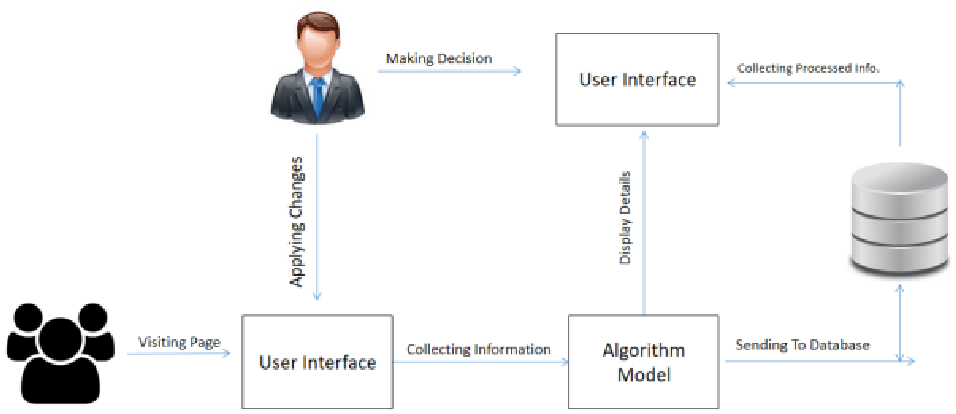

Fig.1 System Architecture

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 11, November 2019

V. EXPECTED RESULT

We will be getting the records of each and every users' web-page scrolling depth so that it can help us to predict the standard web-page depth for our particular website. The data will be recorded of each and every individual for performing operations.

VI. CONCLUSION

By implementing this project, we are providing a platform for various publishers or website owners to predict the content viewability of the page. Solving this issue can benefit online advertisers to allow them to invest more effectively in advertising and can benefit publishers to increase their revenue.

REFERENCES

[1] I. Lunden, "Internet ad spend to reach $121b in 2014, "http://techcrunch.com/2014/04/07/internet-ad-spend-to-reach-121b-in-2014-23-of-537b-total-ad-spend-ad-tech-gives-display-aboost-over-search/
[2] Y. Chen and T. W. Yan, "Improvingaccuracyandcostoftwo-classandmulticlassprobabilisticclassifiersusingROCcurves," in KDD'12, 2012, pp. 795–803.
[3] W. Zhang, S. Yuan, and J. Wang, "Optimal real-time bidding for display advertising," in ACM SIGKDD'14, 2014, pp. 1077–1086.
[4] W. Chen, D. He, T.-Y. Liu, T. Qin, Y. Tao, and L. Wang, "Generalized second price auction with probabilistic broad match," in ACM EC'14, 2014, pp. 39–56
[5] Google, "The importance of being seen", http://think.storage.googleapis.com/docs/the-importanceof-being-seen study.pdf.
[6] M. Mareck, "Is online audience measurement coming of age?" Research World, vol. 2015, no. 51, pp. 16–19, 2015.
[7] S. Flosi, G. Fulgoni, and A. Vollman, "if an advertisement runs online and no one sees it, is it still an ad?" Journal of Advertising Research, 2013.
[8] H. Cheng, E. Manavoglu, Y. Cui, R. Zhang, and J. Mao, "Dynamic ad layout revenue optimization for display advertising," in Proceedings of the Sixth International Workshop on Data Mining for Online Advertising and Internet Economy, 2012, p. 9.
[9] H. Weinreich, H. Obendorf, E. Herder, and M. Mayer, "Not quite the average: An empirical study of web use," ACM TWEB, vol. 2, no. 1, p. 5, 2008.
[10] F. Manjoo, "You won't finish this article," Slate, 2013. Security, vol. 13, 2018