International Journal of Advance Research in Science and Engineering
Volume No.06, Issue No. 12, December 2017
www.ijarse.com
IJARSE
ISSN: 2319-8354

Survey on User Response Prediction Models and Metrics in Computational Advertising

Avila Clemenshia P[1], Dr. Kalyani P[2]

1(*PhD Scholar, Sri Ramakrishna College of Arts & Science, India*)

2(Department of Information Technology, Sri Ramakrishna College of Arts & Science, India)

ABSTRACT

In recent era of increased device utilization and enhanced online user activities, web advertisements (ads) have a major responsibility in enriching the user experience also increasing ad revenue. Predicting the user response plays a vital role in presenting relevant ad to the right audience. This survey is to discuss the various computational advertising scenarios used to predict the ad usage trends in growing web access population in order to benefit both the consumer and the ad network. It involves the impact of the metrics in measuring the ad effectiveness. The paper also provides insight to the research areas which are to be focused for better serving the web usage community in future where social media and mobile app users are increasing drastically every day.

Keywords – Ad Revenue, Click Through Rate, Computational Advertising ,Conversion Rate

I. INTRODUCTION

Web advertising is a form of propaganda that uses the internet and World Wide Web for the major cause of rendering promoting messages to attract customers. Online advertisements (ads) are delivered by an ad server. Online Advertising has become a major medium for advertisement industry to the web user population. It is also called internet advertising or online market. Ad network is a company that connects advertisers to web sites that want to host advertisement. The key function of an ad network is aggregation of ad space supply from publishers and matching it with advertiser demand. The extremely fundamental distinction between customary media promotion systems and online advertisement systems is that online advertisement systems utilize a focal advertisement server to convey notice to purchasers, which empowers focusing on, following and detailing of impressions in ways impractical with simple media choices.

The inevitable advantages in online advertising have extended the scope of advertisement in the industry. The online advertisement industry runs by the revenue generated by the clicks made on the advertisement. Hence it becomes essential to predict the number of clicks made on particular advertisements over the impressions. This will help the advertisers in analyzing the requirements and the interests of the customers and enhance their advertising approach and ad campaign strategy. Thus it becomes essential to predict the number of clicks gained by the advertisements.

Revenue of many internet companies is driven by advertising.

Characteristics of Online Advertising:

- Targeting potential customers: With Online advertising, ads can be shown to the people who are interested in specific brands and services. With more targeting options available like age, demographic location, language, devices used etc. one can more precisely reach out to specific group of people.
- Budget Control: The maximum amount to spend on the campaign and the pattern in which it should be spent can be set. If during the day the traffic is more the ad can be set in a way to be shown more during the day so as to bring more clicks.
- Easy to make changes and updations in Ad Campaign: Based on the ad campaign performance, one can make desired modifications so as to achieve improved outcomes. For instance if an Ad Words campaign has been framed where the specific keyword is showing better result the other keywords can be changed on the basis on the one which is showing better result.
- Exposure to larger group of audience: With more people accessing internet, online advertising targets large number of audience. Ad campaign run online, can reach the audience worldwide. Businesses can get huge number of potential customers in short span of time as the number of targeted audience is large.
- Specific campaign for specific audience: Ad campaigns are designed keeping in mind the audience to be targeted. Depending on the interest of the group of user it is possible to run specific campaigns. There is an option of customizing the campaign wherein the age-group, gender, location of the target user can be selected.
- Wide Coverage: It provides many professional knowledge, skills, and case analysis on online advertising. It can reach people all over the world via internet all around the clock, without restrictions to domain and time.
- Interaction and direct sense experience: Carriers of online advertising are multimedia and hypertext documents. If consumers are interested in a specific product, they can further know more detailed and vivid information about it by just clicking the mouse so as to experience in person the product, service and brand. If new technologies and applications such as virtual reality can be used in online advertising, it allows customers to really experience goods or service which can greatly strengthen online advertising.
- Strong advertising strategy to decide upon the logic and planning behind an advertisement, research the target audience and develop strategy to reach them.
- Creative use of the media: Deciding the best place or means to deliver the advertisement, choose the outlets that will efficiently and effectively reach the target audience, reaching many people for little money.
- Effective ads: Advertisements that deliver the message that the advertiser intended and consumers respond as the advertisers hoped they would gain attention which creates a positive impression for the brand and separate the brand from the competition in the mind of consumers. Influence people to respond in the desired ways.
- Economical: The concern of economy is present in every kind of expenditure in the business and advertisement. The budget spent on advertisements should prove economical with its wide spread message generating good results in the form of more and more demand and sales. Optimum promotion is possible only by economical use of the resources meant for it.

The main goal of successful online display advertisement is to present advertisements and messages to the people who are the visitors of the site. Publishers present the ad to the user by displaying the advertisement in

International Journal of Advance Research in Science and Engineering
Volume No.06, Issue No. 12, December 2017
www.ijarse.com
IJARSE
ISSN: 2319-8354

their web pages. The website or application the user visits greatly impacts the CTR of the ad. The placement of ad in a specific slot plays a major role in increasing the response that the ad could get. There is an ad server for advertisers, publishers and ad network that involves in administration of the ads and the way it is distributed in web sites. The ad server positions them in various websites once they receive the ad files. Selection of ad in display advertisements is decided by the ad exchanges and it is a tedious task.

II.COMPUTATIONAL ADVERTISING AND RESEARCH BREAKTHROUGH

Computational advertising is a rapidly emerging subfield that is aimed at finding the best match between a user and the relevant ad in a context. Computational Advertising enables to achieve personalized service. In advertisement industry Behavioral Targeting is important which predicts potentially profitable users who will click target ads. The prediction of users that are likely to click the ads is done by using user's clicking/web browsing information and presenting the ads they would prefer. Multiple Criteria Linear Programming Regression (MCLPR) [1]prediction model is generated. The experiment data Sets are provided by a leading Internet company in China, and can be downloaded from track 2 of the KDD Cup2012 data sets. In this paper, Support Vector Regression (SVR) and Logistic Regression (LR) are also used for comparison. The advantages include reduced cost, increase in the positive impact on the user experiences, improved web economy and so on. It is necessary to utilize computational advertising strategies as all the players in web ads have their own benefits. Advertisers expect increased ROI. Users expect relevant ads. Publisher expects maximum revenue for each impression or access and ad network needs increased revenue.

1. Literature Review

The various research works on the different types of advertisements are analyzed and they are categorized as sponsored search, contextual ads, display ads, mobile ads and social media ads ad presented in this section.

1.1 Sponsored search

Sponsored Search is a type of web advertisement which displays online advertisement along with the search results when a user enters a query in the web search engine. The keywords act as a basis for this type of advertising. The main problem in Sponsored Search Advertising [2] is of keyword suggestion. The advertisers tend to bid for the keywords with more search volume but that was also expensive. Later on improved models were generated that suggest the related long tail keywords for advertisers with low volume and inexpensive. Experiments conducted have proved that this approach is better than existing keyword suggestion method.

A new Bayesian click-through rate (CTR) prediction algorithm [3]that is based on probit regression model is used in Microsoft's Bing search engine for Sponsored Search. This algorithm maps input features that is discrete or real-valued to probabilities. Gaussian evaluations are done and the scalability of the particular algorithm is ensured through a principled weight pruning procedure and a relatively close parallel implementation.

To predict an ad's CTR, historical click information is required. The new ads lack in historical data and is challenging to make predictions. Bayesian network (BN) [4] is used to generate accurate prediction model by identifying and inferring dependencies and uncertainties among variables. In this model Bayesian network of the keywords to describe the ads in a certain domain is developed first. Next an algorithm is generated to find similar keywords with the new ads. The similar ads are found from this and the CTR of the new ad is estimated more accurately.

1.2 Contextual ads

Relevance of the displayed ads to the page content is provided by scoring the match between the individual ads ad the page content. In Contextual advertising [5] additional parameters are added to improve this match by a logistic regression model on the page words and ads. The logs containing ad impressions and clicks with shrinkage estimators are being used to combat sparsity. A Hadoop framework is used. Experimental evaluation is provided showing improved click prediction over a holdout set of impression and click events from a large scale real-world ad placement engine. This model achieves a 25% lift in precision relative to a traditional information retrieval model which is based on cosine similarity, for recalling 10% of the clicks in the test data.

1.3 Display ads

In display advertising conversion rate plays a vital role in estimating the effectiveness of ad campaign is its the proportion of users who take a predefined action on the advertiser website, such as a purchase. Ad impression can be predicted with this conversion rate using machine learning. Delayed feedback [6] in which conversions occur at a considerable period up to a month after an impression is a difficult scenario. This issue is handled in this paper by proposing a model for evaluating the conversion delay. It enables to identify if the user is in negative sample.

Presenting the relevant ad to the user involves identifying the adslot in which the relevant display ad is to be placed. This enables to achieve high conversion. Micro and macro approaches [7]where used to match between the adslots, ads and users. Micro level methods require rich attribute set on users, ads and adslots. Certain cases may not provide all such attributes. Later on research using a macro approach was carried out for mining new adslots for each ad where attribute information are pre-evaluated offline. The matrix factorization technique is applied to the history matrix of adslot performance. This has helped to predict the performance of the target adslots which yields high conversion rates.

1.4 Mobile ads

Mobile advertising has recently seen dramatic growth, as a result of the global proliferation of mobile phones and devices. This makes the task of predicting ad response more crucial for increasing the ad revenue. The ad response data change dynamically and are subject to new startup situations in which scarcity in history makes the reliable prediction task a complex one. It is necessary to have a robust regression [8] estimation for increased and highly accurate prediction and also efficient ranking to identify the varied impacts of different ads. The development of Hierarchical Importance-aware Factorization Machine renders an effective generic latent factor framework. Relative analysis results in stating that HIFM outperforms the contemporary temporal latent factor models. This shows that the HIFM's importance-aware could improve the prediction accuracy and hierarchical learning is better for new scenarios.

By 2019, mobile ad spending will rise to $65.87 billion, or 72.2% of total digital ad spend [9] which is illustrated below in Fig 1 & Fig 2.

US Mobile Ad Spending, 2013-2019

	2013	2014	2015	2016	2017	2018	2019
Mobile ad spending (billions)	**$10.67**	**$19.15**	**$28.72**	**$40.50**	**$49.81**	**$57.78**	**$65.87**
—% change	120.0%	79.5%	50.0%	41.0%	23.0%	16.0%	14.0%
—% of digital ad spending	24.7%	37.7%	49.0%	60.4%	66.6%	69.7%	72.2%
—% of total media ad spending	6.3%	10.8%	15.3%	20.4%	23.9%	26.3%	28.6%

Note: includes classified, display (banners and other, rich media and video), email, lead generation, messaging-based and search advertising; ad spending on tablets is included
Source: eMarketer, March 2015
186582 www.eMarketer.com

Fig 1: US Mobile Ad Spending statistics.

Mobile ad spending has shown a drastical rise in the above table given by eMarketer and will continue to grow compared to that of desktop ads. In 2014, US adults spent an average of 2 hours, 51 minutes with mobile devices each day, up from 2 hours, 19 minutes in 2013. At the format level, mobile display dollars will continue to increase, outpacing mobile search. [9].

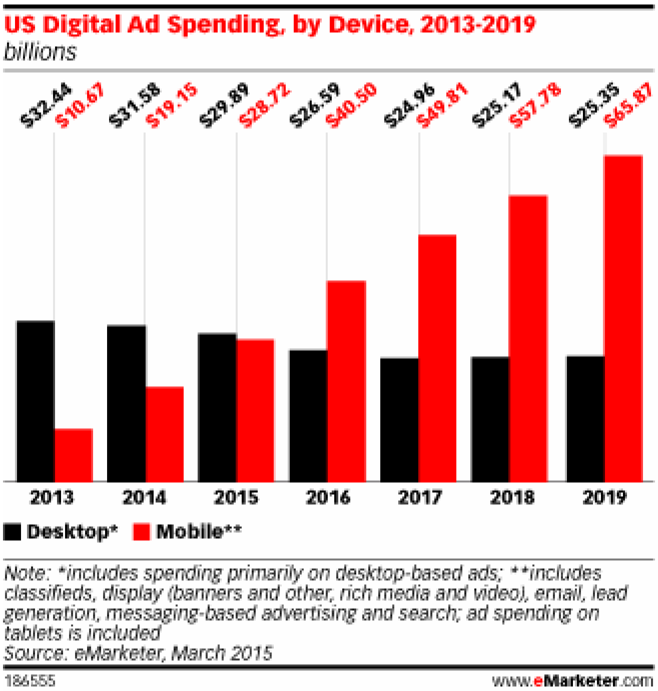

Fig 2: Digital Ad Spending by various devices.

There are ample of opportunities in the arena of mobile advertising for advertisers to only bid for ads [10] and pay for evaluated user responses, such as clicks on ads. This makes click prediction systems as core to most mobile advertising systems. The usage of mobile users' dramatic growth have made predicting clicks on Mobile ad a challenging machine learning task.

1.5 Social Media Ad

Various industries constantly work to improve customer relationships and customer attention. A new systems architecture has been proposed that combines the textual content messages on social media with product information such as the summarizing catalogue descriptions for delivering marketing campaign recommendations. User profiles are built based on buying histories and other customer-specific information is different from social media users. In this research the problem of targeting individual social media messages without personalized profile information has been carried out by combining two disparate computational toolboxes for text analytics, natural language processing and machine learning to select social media users for whom to target with relevant topic-specific advertisements. The context of social media messages is analyzed by natural language. Machine learning is used to analyze product information in order to match social media messages with products and generating ranks for potential advertisements. The framework has been developed for tourism industry using Twitter® social media platform [11]. Social media is growing everyday with increased user population within its network. From the perspective of advertising, focus is on the social speech from which psychological and personality of a person can be assessed and rich profiles of users are derived based on the type of the language used in the social speech. Also the advertising models cluster people based on their click interests. Using machine learning patterns are extracted to make the data informative for decision making. This customer profiling provides the ad network with the capability to target.

III. RESPONSE PREDICTION METRICS - CTR & CVR

The two main estimations that are essential for prediction tasks in display advertising are clickthrough and conversation rates. A framework [12] of machine learning using logistic regression was created. It was modeled to handle specifically the type of display advertising. The resulting system possess many characteristics such as easy implementation and deployment, high scalability and have been trained on terabyte volume of data, and it provides models with state-of-the-art accuracy.

The click-through rate (CTR) has been used as a key measure to evaluate the [13]effectiveness of ad campaign in on-line search and display advertising. The author found that an alternative that has been used for measuring direct profitability that has become increasingly popular is the conversion rate (CVR). Various user action like buying, registering, downloading rather than simply page browsing shows that they desire the ad and this proportion of users are measured using the CVR. This research is focused on the post-click conversion (PCC) problem which is the analysis of conversions after a user click on a referring ad. The solution puts forward an approach to measure the attribute relevance.

Click prediction model is generated by using multimedia features [14]which enables to generate pricing models such as cost-per-click (CPC) and cost-per-action (CPA). This resolves the challenge faced by new ads that lack in historical data for prediction. The similarity in content and aesthetics of other ads is used for predicting clicks for new ads in this work.

Ad revenue and reputation is improved and the advertising performance is optimized for advertisers with the help of click-through rate (CTR) prediction. Prediction accuracy has to be improved by facing the challenge due to the imbalanced distribution of the advertising data and real time bidding. In this paper, the user features, the historical CTR features, the ID features, and the other numerical features are used to develop a novel online CTR prediction by user profile system construction. A novel CTR prediction approach [15] is presented to address the imbalanced learning sample distribution by integrating the Weighted-ELM (WELM) and the Adaboost algorithm.

IV. CONCLUSION & FUTURE WORK

Web advertisements are becoming close to consumers day by day and their experience becomes desirable towards ads if the relevant ads are presented for their perusal. This can be achieved by applying computational advertising strategies and evaluating metrics to generate the best user response prediction model. The trends are emerging so as to maximize the mobile and app usage which requires research to be contributed on the activities of mobile users. This enables to reach even larger group of users enhancing the ad revenue and better user experience. The mobile user profile system, app usage history, frequency of a particular user visit through a particular device, multimedia features also the neighboring ad features of the ad clicked by the user can also be analyzed to extract more features. Also social media advertising is becoming more popular and advertising can be targeted to users by analyzing their preference to attain increased revenue.

REFERENCES

[1] Fang Wanga*, Warawut Suphamitmongkola, and Bo Wanga, (2013), *"Advertisement Click-Through Rate Prediction using Multiple Criteria Linear Programming Regression Model"* Elsevier B.V, Procedia Computer Science 17 p. 803 – 811.

[2] Ms. Rinki Talwar, Dr. Shuchita Upadhyaya, (2017), "*Long Tail Keyword Suggestion for Sponsored Search Advertising*", International Research Journal of Engineering and Technology (IRJET), e-ISSN: 2395-0056 Volume: 04 Issue: 07 | July -2017, Page 1407.

[3] Thore Graepel, Joaquin Quinonero Candela, Thomas Borchert, and Ralf Herbrich, (2010), "*Web- Scale Bayesian Click-Through Rate Prediction for Sponsored Search Advertising in Microsoft's Bing Search Engine*", Proceedings of the 27th International Conference on Machine Learning, ICML 2010.

[4] Zhipeng Fang, Kun Yue, Jixian Zhang, Dehai Zhang, and Weiyi Liu, (2014), "*Predicting Click-Through Rates of New Advertisements Based on the Bayesian Network*", Mathematical Problems in Engineering Volume 2014, Article ID 818203, 9 pages, doi:10.1155/2014/818203.

[5] Deepayan Chakrabarti, Deepak Agarwal, Vanja Josifovski, (2014) "*Contextual Advertising by Combining Relevance with Click Feedback*" Proceedings of the 31 st International Conference on Machine Learning, Beijing, China, 2014. JMLR: W&CP volume 32, 2008, Beijing, China. ACM 978-1-60558-085-2/08/04.

[6] Olivier Chapelle, (2014), "*Modeling Delayed Feedback in Display Advertising*", Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining, pages: 1097-1105.

[7] Kazuki Taniguchi, Yuki Harada, Nguyen Tuan Duc, (2015), "*Adslot Mining for Online Display Ads*",

IEEE International Conference on Data Mining Workshop (ICDMW), pages: 355 – 364

[8] Richard Jayadi OENTARYO, Ee Peng LIM, Jia Wei LOW, David LO and Michael FINEGOLD,(2014), "*Predicting Response in Mobile Advertising with Hierarchical Importance-Aware Factorization Machine*", WSDM'14: Proceedings of the 7th ACM International Conference on Web Search and Data Mining, pages: 123-132.

[9] https://www.emarketer.com/Report/US-Ad-Spending-eMarketer-Forecast-2017/2001998

[10] M. Sree Vani, (2016), "*Prediction of Mobile Ad Click Using Supervised Classification Algorithms*", (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 7 (2), pages: 623-625.

[11] S. Buckley , M. Ettl, P. Jain, R. Luss , M. Petrik, R. K. Ravi and C. Venkatramani, (2014), "*Social Media And Customer Behavior Analytics For Personalized Customer Engagements*", IBM Journal of Research and Development Volume: 58, Issue: 5/6, pages: 7:1 - 7:12.

[12] Olivier Chapelle, Eren Manavoglu and Romer Rosales, (2015), "*Simple and Scalable Response Prediction for Display Advertising*", ACM Transactions on Intelligent Systems and Technology (TIST) - Special Sections on Diversity and Discovery in Recommender Systems, Online Advertising and Regular Papers, Volume 5 Issue 4, Article No: 61.

[13] Rómer Rosales, Haibin Cheng and Eren Manavoglu, (2012), "*Post-click conversion modeling and analysis for non-guaranteed delivery display advertising*", Proceedings of the fifth ACM international conference on Web search and data mining, February 08-12, 2012, Seattle, Washington, USA [doi>10.1145/2124295.2124333].

[14] Haibin Cheng , Roelof van Zwol , Javad Azimi , Eren Manavoglu , Ruofei Zhang , Yang Zhou and Vidhya Navalpakkam, (2012), "*Multimedia features for click prediction of new ads in display advertising*", Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining, August 12-16, 2012, Beijing, China [doi>10.1145/2339530.2339652].

[15] Sen Zhang, Qiang Fu and Wendong Xiao, (2017), "*Advertisement Click-Through Rate Prediction Based on the Weighted-ELM and Adaboost Algorithm*", Scientific Programming Volume 2017, Article ID 2938369, 8 pages.