*J. Appl. Environ. Biol. Sci.*, 5(12S)417-420, 2015



ISSN: 2090-4274
Journal of Applied Environmental and Biological Sciences
*www.textroad.com*

# Applying Data Mining for Advertisement in Social Networks and Improving CTR

**Mohammad Hajarian**

Department of Computer Engineering, Karaj Branch, Islamic Azad University, Karaj, Iran.

*Received: July 24, 2015*
*Accepted: September 31, 2015*

**ABSTRACT**

In this paper different usage of data mining in social networks has been reviewed and a hypothesis has been define that states by implementing data mining in social networks The CLICK THROUGH RATE (CTR) of their advertisement will improve, which can lead to improvement of the revenue of social network websites. Additionally finding of recent study on social network advertisements which has been conducted will be analyzed and evaluated to validate this hypothesis to show that after implementing data mining on advertisement in social networks CTR of an advertisement will be improved. Moreover methods of the performed research such as schema of database tables that was used in this study and other requirements of this study was stated in this paper and findings before and after implementing data mining was compared in form of bar charts. Additionally other relevant studies regarding this subject has been reviewed and the role of data mining to improve click through rate of advertisement in social network was described.
**KEYWORDS:** Data mining, Advertisement, Association rule, CTR, Social network.

## 1. INTRODUCTION

Social network websites usually has different data about their members, using data mining this data can be converted to the knowledge that social networks can benefit from. The objective of this study is to show that Data mining can be used to show relevant advertisement to website users. And by performing that, the CLICK THROUGH RATE (CTR) of online advertisement in social networks will improve. to reach the objective of this study in this paper first author evaluate findings of other studies which implement data mining in social network websites and based on this findings and the methods used author defines a hypothesis. After that, two round data gathering using observation method will be performed to validate the hypothesis. Finally author will describes findings of the study that shows by implementing data mining to show relevant advertisement to the members of social network website, the CLICK THROUGH RATE (CTR) of advertisement will improve.

## 2. LITERATURE REVIEW

Data mining has been used in social networks for different purposes, Davidson et al. [1] has used data mining to improve the recommendation system of showing relevant videos to the users of YouTube website. In that research they had used association rules of data mining to find out each user is more interested in which type of videos, in terms of videos features such as video length and video received likes. After implementing data mining to show relevant videos to the YouTube users, Davidson et al. [1] were able to validate they framework by calculating new CTR of visited videos and comparing it to the CTR that before the implementing data mining was calculated. Additionally in another research by Ranjan et al.[2] has used association rule of data mining to broadcast users activities only to the friends who might be interested in that topic. In another research, Hafizpour et al. [3] has used Apriori algorithm to predict mobile black hole attacks. All these studies showed that by implementing association rule of data mining it is possible to predict users' interest. Moreover Davidson et al. [1] framework shows that by comparing CTR before and after implementing data mining author can validate its recommendation system.
Based on the findings of Davidson et al. [1] and Ranjan et al. [2] and Hafizpour et al. [3] it can be assumed that by implementing association rule of data mining in to the social network website and matching users features such as age and gender to advertisement types, there is a big chance that advertisement CTR will improve in social networks. To validate this hypothesis researcher designed a research and put this hypothesis to practice.

* **Corresponding Author:** Mohammad Hajarian, Department of Computer Engineering, Karaj Branch, Islamic Azad University, Karaj, Iran. Email: m.hajarian@kiau.ac.ir

## 3. DESIGN RESEARCH

This study has been performed using observation method. Data has been gathered anonymously from gohardasht.com social network website which researcher is its founder and CEO of it. To perform this study researcher has gathered information about users' clicks and advertisement views in addition to advertisement version and users features such as age and gender. All participants has been informed about this research aim and voluntary participate to this research. The results of this study was delivered as quantitative data which are more suitable to show in diagrams and graphs [4]. To calculate the success rate of advertisement, click through rate (CTR) has to be obtained and compared. Researcher gathered data such as number of times which specific type of advertisement has been shown to the users and number of clicks that users has made on different type of advertisements. Table 1 illustrates the formula to calculate the CTR of an advertisement.

*Table 1 1- Calculation of CTR*

| |
|---|
| $CTR = \frac{\text{Clicks} * 100\%}{\text{Impressions}}$ |

To show the improvement of CTR, these data (clicks and impressions) was collected and analyzed before and after implementing data mining to show relevant advertisement to users. Then the findings was compared to each other.

## 4. DATA GATHERING

As it was mentioned before data gathering will be done using observation method. Hence all the data such as users' clicks and advertisement impressions will be saved to the database. Table 2 illustrates the proposed schema to gather all the required information.

*Table 2 2- proposed database tables*

| Ad-click-table |
|---|
| Age |
| Gender |
| Ad type |
| AD ID |
| Ad version |

| Ad-delivery-table |
|---|
| Age |
| Gender |
| Ad type |
| AD ID |
| Ad version |

In addition to the proposed tables, appropriate background programing has to be done to gather the users' clicks and advertisement impressions in addition to all other information. Data gathering process will perform in two cycles, in the round one the "ad version" of advertisement will be "1" and in the round to "ad version" will be "2". In the round one, data will be gathered before performing data mining and in the round two data will be gathered after performing data mining.

## 5. ROUND ONE OF DATA GATHERING

After creating appropriate tables and background programming to gather the required data to the proposed tables, data observation started and advertisement was shown 2248148 times to the users of gohardasht.com in Iran. During this observation 793 click has been saved. Based on the formula in Table 1, the CTR of 3.52 in scale of 10.000 has been calculated before implementing data mining.

## 6. IMPLEMENTING DATA MINING

After gathering enough amount of data in round one, data mining has to be applied to the findings. To perform data mining based on the findings of round one of study, by implementing association rules on ad-click-table, some rules was extracted that shows which type of users might be more interested in which type of advertisements these rules has been used to perform second round of the study. Moreover the program that delivers advertisement to the users has been changed in such a way that if it can find an association rule based on the users features, it will deliver relevant type of advertisement to the user, and in this case the data of the clicks and impressions in addition to all other data that was illustrated in table 2 will be saved in SQL tables. Additionally in this round "Ad version" was changed from 1 to 2.

## 7. ROUND TWO OF DATA GATHERING

The round two of the research was performed after implementing data mining on advertisement. In the second round of data gathering advertisement has been shown to the users for 1569829 times and 834 click has been received. Based on the formula of table 1 the CTR of 5.31 in scale of 10.000 was calculated for advertisement in second round.

## 8. ANALYSIS DATA AND VALIDATION

Figure 1 illustrates comparing findings of round one and round two of this research.

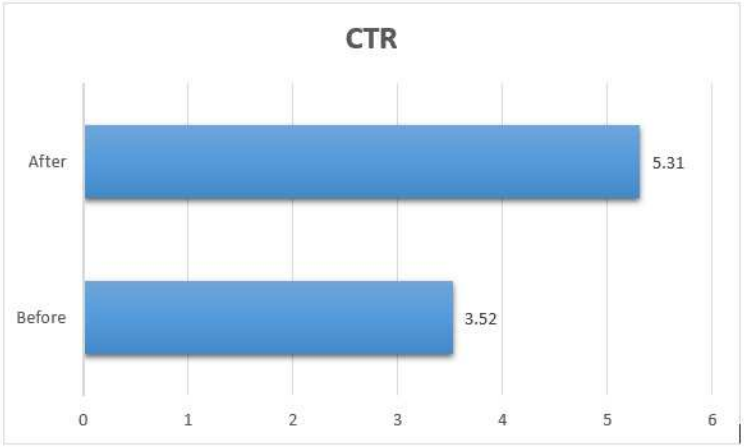


*Figure 1 - 1comparing CTR before and after implementing data mining*

As Figure 1 illustrates, after implementing data mining for advertisement, click through rate (CTR) has been improved from 3.52 to 5.31. Growing of 1.79 in CTR of advertisement validates the proposed hypothesis and shows that by implementing association rules of data mining for advertisement in social networks, click through rate (CTR) of advertisement in social networks will improve and shows that data mining can be used to improve the revenue of social network websites.

## 9. CONCLUSION

In conclusion based on the findings of this research it can be understand that author has reached the objective of this research which is, to show that data mining can be used to show relevant advertisement to the social network users, and by implementing data mining in social networks, CTR of advertisement will increase.

Moreover in this paper author has reviewed different studies regarding data mining in social networks and based on those findings proposed a hypothesis that data mining can improve the CTR of advertisement in social networks.

Additionally it was understand that to implement the data mining for advertisement in social networks based on the literature review association rule of data mining is most suitable technique. The hypothesis was validated by performing observation study before and after implementing data mining in gohardasht.com social network. And results shows improvement of advertisements' CTR. Hence it can be concluded that implementing data mining can improve CTR of advertisement in social networks. For future works, combination of using data mining technique such as association rules and neural network is suggested, researches also encourage to gather on more variables of users for performing data mining on, to get more accurate results.

## REFERENCES

1. Davidson, J., Liebald, B., Liu, J., Nandy, P., Van Vleet, T., Gargi, U., Gupta, S., He, Y., Lambert, M., Livingston, B. and Sampath, D., 2010, September. The YouTube video recommendation system. In Proceedings of the fourth ACM conference on Recommender systems (pp. 293-296). ACM.
2. Ranjan, R., Vyas, D. & Guntoju, D. P. (2014) Balancing the trade-off between privacy and profitability in Social Media using NMSANT. In Advance Computing Conference (IACC), 2014 IEEE International.) IEEE, pp. 477-483.
3. Hafizpour, H. and Mirabedini, S., 2013. Using Apriori algorithm to prevent black hole attack in mobile Ad hoc networks. Management Science Letters, 3(1), pp.351-358.
4. Creswell, J.W. and Clark, V.L.P., 2007. Designing and conducting mixed methods research.
5. Kothari, C.R., 2004. Research methodology: Methods and techniques. New Age International.
6. Creswell, J.W., 2002. Educational research: Planning, conducting, and evaluating quantitative. New Jersey: Upper Saddle River.