

## FACEBOOK USER ADVERTISEMENT CLICK PREDICTION

*Tehmeem Bukhari<sup>1</sup>, Afrin Khan<sup>2</sup>, Heta Shah<sup>3</sup> & Babita Bhagat<sup>4</sup>*

*<sup>1,2,3</sup>Research Scholar, Department of Computer Engineering, Pillai HOC College of Engineering and Technology, Rasayani, Maharashtra, India*

*<sup>4</sup>Assistant Professor, Department of Computer Engineering, Pillai HOC College of Engineering and Technology, Rasayani, Maharashtra, India*

**Received: 03 Apr 2020**

**Accepted: 10 Apr 2020**

**Published: 17 Apr 2020**

### ABSTRACT

*Social media has become a large platform to extend one's brand's awareness as it bridges the gap between customer and e-commerce. As online advertisement marketing is in demand, many dealers approach social media application owners for marketing their products so that these companies display the advertisements of the merchandise dealers. Such digital marketing methods have poor audiences targeting. This affects either side as customers aren't getting ads they're curious about and also dealers for paying huge price on a daily basis. The machine learning model analyze the given data and using logistic regression algorithm, predicts probability whether the given user supported attributes is probably going to click on or not. As logistic regression algorithm could be a sigmoid relation between data, such that numbers for training is 67 percent and testing is 33 percent. This will increase the productivity up to 200 percent and effective usage of ad revenue. This model is significantly improved version as compared to last model.*

**KEYWORDS:** *Click Prediction, Logistic Regression, Sigmoid*

### INTRODUCTION

Ad is the significant piece of any professional that web based life has become a stage for advanced showcasing. Online commercial is sought after the seller's ways to deal with internet based life and pay tremendous sum for their item ad. But because of poor objective crowd, it is misfortune for the vendors just as it doesn't reach to concerned crowd. Face book, the most famous informal organization has built up a focus on innovation which permits commercials to contact a particular crowd. Face book Promotions is a part within Face book; which is accessible to both clients and business. While posting a promotion through the Face book Promotion Administrator, a publicist is given a lot of qualities that will characterize his objective market. Face book calls this as crowd focusing on. These qualities incorporate land area, sexual orientation, age, work, and relationship status, and intrigue, for example, music, among others. Face book claims that promoters can even tweak their intended interest group dependent on their conduct, for example, buying designs, gadget use, and different exercises. This is the reason Face book clients see promotions on their profile page that are pertinent to their inclination and interests. This permits the promotions to be not so much meddlesome but rather more fruitful in conveying the proper substance to the correct crowd. The notice calculation is likewise equipped for checking execution. So sponsor or Face book advertisers can alter their crowd just as the nature, spending plan, and term of promotions dependent on its presentation.

## Significance of the System

The paper mainly focuses on how machine learning techniques can be applied to predict probability whether the given user supported attributes is probably going to click on Ad or not. The study of literature survey is presented in section III, Methodology is explained in section IV, section V covers the experimental results of the study, and section VI discusses the future study and Conclusion.

## LITERATURE SURVEY

Kushal Dave (2010) <sup>[6]</sup> this model acquires the snap data of uncommon/new promotions from other semantically Related advertisements. The semantic highlights are gotten from the inquiry promotion navigate charts and sponsors Account data.

Abirami, R and Vijaya <sup>[5]</sup> (2012) this paper show the relevance of help vector relapse, an AI method, for foreseeing The stock cost by learning the notable information. The stock information for the time of four years is gathered and prepared with different parameter settings. The exhibition of the prepared model is assessed by 10-overlay cross approval for its prescient exactness.

Fang Wang (2013) <sup>[4]</sup> Major online distributors, for example, Yippee, Microsoft and Google have eagerly grasped this plan of action. The business estimation of ad on the web relies upon whether clients click in the notice. Right now, Vector Relapse (SVR) and Strategic Relapse (LR) are utilized as two benchmark models for examination. The outcomes show that MCLPR is a promising model in a social focusing on undertakings.

Vijaya M.S. (2016) <sup>[3]</sup> the advertisement is shown to the clients when they visit a page or application. The assignment of foreseeing active visitor clicking percentage particularly for show notice is a significant testing issue. Right now relapse (LR), Poisson relapse (PR) and bolster vector relapse (SVR) relapse procedures are utilized..

Suha Kagan Kose (2017) <sup>[2]</sup> The investigation shows Dynamic In significant web indexes, when a client look through inquiry, pertinent promotions are shown by advertisements' snap engaging and offers that sponsors ready to pay when their advertisement is clicked.

Jelena Gligorijevic (2018) <sup>[1]</sup> AI models made extraordinary achievement in anticipating CTR for supported pursuit. The vast majority of the models received in the business depend on an enormous arrangement of all around structured highlights to.

## METHODOLOGY

The console used in Fb user advertisement click prediction is Jupyter notebook. It is divided into two parts: for data visualization Tableau and Matplotlib flash seaborn. Logistic Regression algorithm is used in this project. In logistic regression, Sigmoid function is used. Formula:  $y = \frac{1}{1+e^{-x}}$

This is used to predict whether user will click the particular advertisement or not. Hyper tangent is used in this project. Error function formula:  $2\sqrt{2} \int_0^{\infty} \pi e^{-t^2} dt$ .

500 Data Set will be taken and data will be collected from Kaggle and Face book. The attributes of data are name, emails, country, click, time spent on site and salary.

## ALGORITHM

### Logistic Algorithm

Logistic regression is a supervised learning algorithm which is typically used for binary classification problems. Although “regression” contradicts with “class”, the primary target right here is on the word “logistic” bearing on logistic function which does the category mission throughout this algorithm. Logistic regression is an easy, yet very powerful classification algorithm, so it's usually used for numerous binary class tasks. Customer churn, junk email, electronic computer or ad click on prediction are some example of the regions during which logistic regression offers a strong solution.

$$\text{Sigmoid Function: } y = \frac{1}{1+e^{-x}}$$

The basics of logistic regression is that the logistic characteristic, also observed because the sigmoid characteristic, which takes in any real valued number and maps it to a fee between 0 and 1

$$\text{Formula: } y = \frac{1}{1+e^{-x}}$$

Logistic regression model takes an equation as enter and use logistic characteristic and log odds to perform a binary type mission. Before entering into detail on logistic regression, it's better to review some concepts within the scope probability. In many cases, you'll map the logistic regression output into the answer to a binary classification problem, during which the intention is to properly be expecting one in all feasible labels (e.g., “tapped” or “not tapped”). A later module makes a specialty of it. You'd possibly be questioning how a logistic regression version can make sure output that constantly falls between 0 and 1. As it happens, a sigmoid function, described as follows, produce output having the ones self-same characteristics.

$$\text{Formula: } y = \frac{1}{1+e^{-x}}$$

The sigmoid function yields the subsequent plot:

Figure 1 shows If  $z$  represents the output of the linear layer of a version educated with logistic regression, then sigmoid ( $z$ ) will yield a rate (a probability) between 0 and 1. In mathematical term:  $y' = \frac{1}{1+e^{-(z)}}$

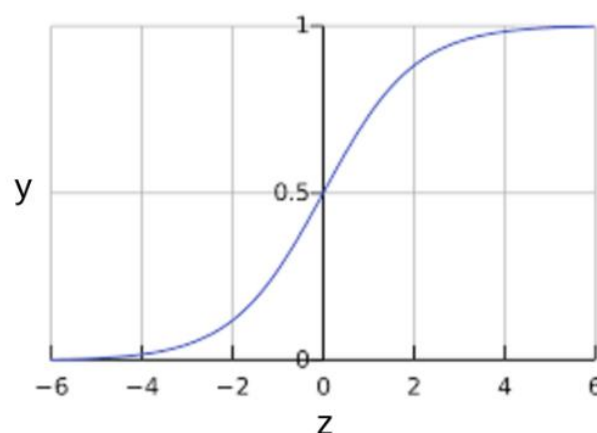


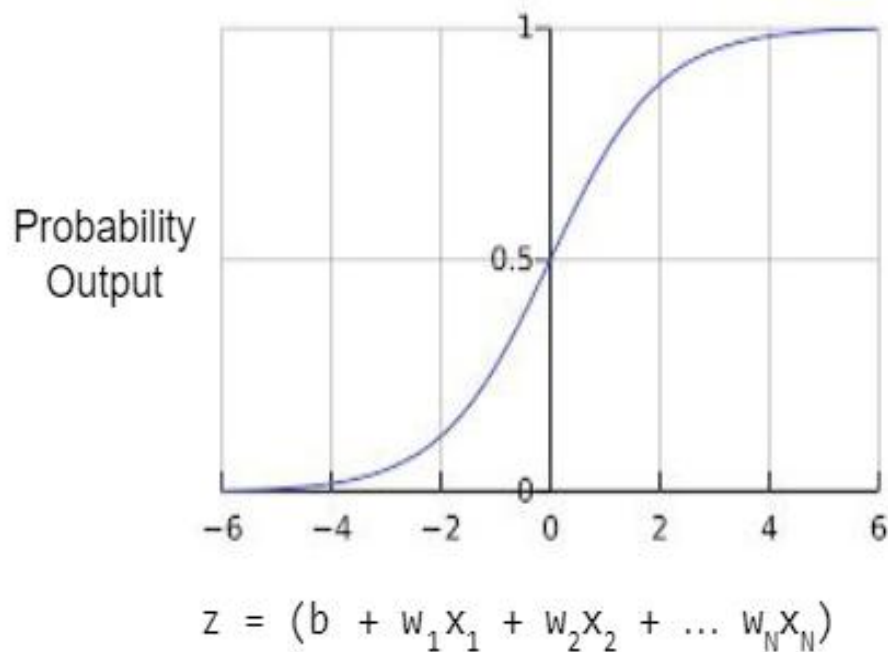
Figure 1

Where:

- Y' is the output of the logistic regression model for a specific example.
- $z = b + w_1x_1 + w_2x_2 + \dots + w_nx_n$
- The w esteems are the model's scholarly loads, and b is the predisposition.
- The x esteems are the element esteem for a specific model.

Note that z is other than passed on up in light of the fact that the log-chances in view of the truth the turnaround of the sigmoid states that z can be described considering the way that the log of the chance of the "1" name (e.g., "Tapped on Ad") isolated with the guide of the probability of the "zero" mark (e.g., "Not tapped on Ad").  $Z = \log\left(\frac{y}{1-y}\right)$

**Here is the Sigmoid Function with ML Labels:**



**Figure 2**

### System Design

The existing system utilized increasingly complex calculations which become an obstacle to create exact outcomes; it may not consider scarcely any highlights during highlight extraction process which brings about less exactness model. In our proposed framework, we will attempt to beat this issue.

The Machine learning model dissects the given information and utilizes calculated relapse calculation, predicts likelihood whether the given client is dependent on ascribes is probably going to tap on or potentially not. As strategic relapse calculation is a sigmoid connection between information, similar to numbers for preparing is 67 percent and testing is 33 percent. This will build the efficiency up to 200 percent and advertisement income will be utilized viably. This model is essentially improved form when contrasted with last model.

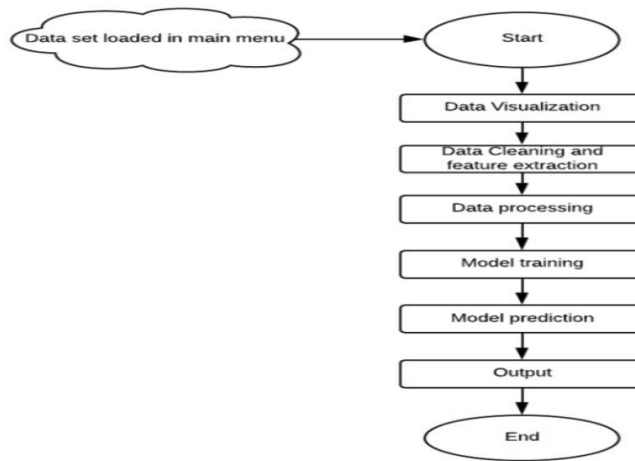


Figure 3: System Design.

EXPERIMENTAL RESULTS

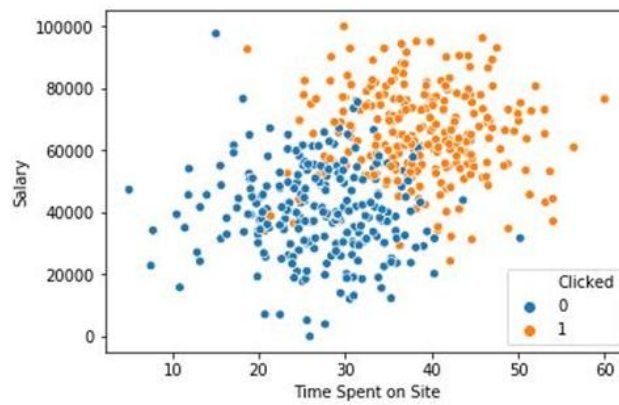


Figure 4



Figure 5

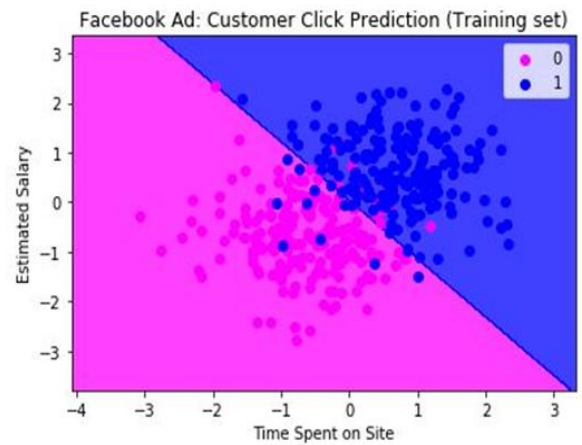


Figure 6

## CONCLUSIONS

This paper exhibits the demonstration and execution of active visitor clicking percentage forecast for commercial show. In the past model was utilized more with complex calculations which become prevention factor to create exact outcome, it may not consider barely any highlights during feature extraction process which brings about less precision model. So, to beat this issue and reach to somewhat exact outcome, we have utilized this machine learning model that dissects the given Information and utilizes strategic relapse calculation, that predicts whether the given client is dependent on credit and is probably going to tap on inclusion or not. This can assist with giving the business well objective crowd, chopping down the expense up to 200 % or somewhere in the vicinity.

For the extension for future research, model can be produced utilizing huge dataset, increasingly precise highlights can be distinguished utilizing KDD procedure to make gathering model to improve the forecast outcome.

## REFERENCES

1. Jelena Gligorijevic, Djordje Gligorijevic *Deeply Supervised Semantic Model for Click-Through Rate Prediction in Sponsored Search.*
2. Suha Kagan Kose *Predicting Click through Rate of new Ads: Considering CTRs as Manual Ratings.*
3. Avila Clemenshi, Vijaya M. S. *Click through Rate Prediction for Display Advertisement.*
4. Fang Wang *Advertisement Click-Through rate prediction using Multiple Criteria Linear Programming Regression Model.*
5. Ling, X., Deng, W., Gu, C., Zhou, H., Li, C., & Sun, F. (2017, April). *Model ensemble for click prediction in bing search ads. In Proceedings of the 26th International Conference on World Wide Web Companion (pp. 689–698).*
6. He, X., Pan, J., Jin, O., Xu, T., Liu, B., Xu, T. & Candela, J. Q. (2014, August). *Practical lessons from predicting clicks on ads at face book. In Proceedings of the Eighth International Workshop on Data Mining for Online Advertising (pp. 1–9).*
7. Abirami, R and Vijaya *Stock price prediction using Support Vector Regression.*
8. Kushal Dave *learning the click through rate for Rare / New Ads from similar Ads.*