

# Prediction of CPC Using Neural Networks for Minimization of Cost

Pradeep Chaudhari, Ravindra Dingankar, and Roshan Chaudhari, *Member, IACSIT*

**Abstract**—In this paper we have proposed a hybrid model of Neural Networks which will predict the optimum cost for keywords in online advertising of Cost Per Click (CPC) model. Neural network is best known for finding hidden relationship, pattern or prediction of unknown variables; we have combined machine learning approach of Pattern Recognition Model and Error Correction Model in Neural Networks to calculate appropriate values for CPC. The first phase, pattern recognition model, stores keyword and its respective derived numeric value, we call it as efficiency of a keyword. The second phase, error correction model, considers all numeric features plus derived variable from first phase for prediction of CPC. In error correction model, with supervised learning weights are adjusted depending upon difference between actual value and calculated value.

**Index Terms**—Optimization, neural network, pay per click, prediction machine learning.

## I. INTRODUCTION

The advent of revolution in technology and internet has caused increase in marketing platform for advertising companies. There are three models for determining the cost of advertising. They are Cost per Impressions (CPM), Cost per Click (CPC) and Cost per Action (CPA).

CPM denotes Cost per 1000 Impressions (frequency of the ad display). If the advertisement is shown 2000 times the cost will be equal to 2 CPM price.

CPC is Cost per Click. Google Ad Words made this model popular. Advertising companies typically bid on keywords, phrases pertaining to their market with search engines. Websites usually charge a fixed price per click rather than using a bidding system. In this advertising model, you have to pay for number of clicks you get on your ads and not for the number of impressions it takes to generate those clicks. Cost-per click is important because it is the number that is going to determine the financial success of your paid search advertising campaign.

CPA advertising tracks the person clicking on an ad and determines if that person then also creates a desired transaction on the destination site. If that transaction (a sale, a completed form) occurs then the advertiser pays the publisher for that positive result.

Your return on investment, whether you are over or under paying for each action, will be determined by how much you

are paying for clicks. It also depends on what kind of quality you are getting for that investment. Thus, it is important to think about Cost-per Click in terms of both cost and value.

It is necessary to identify and target clicks that are both inexpensive and valuable. Advertising companies and marketing firms are putting lot of efforts in targeting right users on such platforms. The main challenge is to get conversion or lead for their product. This depends upon factors like appropriate ad placement, ranking, CPC, relevant landing page etc. Getting all these odds in line is majorly affected by the bid (in other terms, CPC) set by an advertiser for their keywords. If CPCs are not appropriately set then advertising companies may lose revenue. This paper addresses the issue of setting CPCs for keywords used by advertising companies for their ads using Hybrid Neural Network. The goal of this approach is to moderate the cost and increase the marginal revenue.

The paper is organized as follows. Section II covers some of the past work in pattern recognition as well as error correction model. Section III discusses the architecture. Then in section IV, we will discuss the overall system/model containing memory based as well as error correction model, their input and output variables, how they are calculated. In Section V, we will find out why hybrid model is better than single memory based model or error correction model. Lastly, paper concludes with comments on enhancement of the proposed solution in other domains and some conclusions.

## II. LITERATURE SURVEY

The memory based model for Neural Network can be effectively used to adapt to dynamic environment. The agents should acquire new knowledge as also learn to modify old knowledge as environments can change dynamically for them. Authors, Seiichi Ozawa and Kenji Tsumori, have proposed a neural network model consisting of four modules: resource allocating network, long-term memory, association buffer, and environmental change detector. The model is applied to a simple dynamic environment where target functions to be approximated are changed turn by turn [1].

A cheaper solution for correcting OCR engine output can be found with the help of machine learning. Authors, Ahmad Abdulkader and Matthew R. Casey propose the following solution. Error probability of each word from ground-truth data should be assessed by an error estimator neural network. The output probability error estimates from outputs of multiple OCR engines decide whether the word is inspected by a human. The estimator is trained to optimize the area under the word error ROC improving human correction process. Similar words are brought together during correction process which reduces the cost significantly. The paper also

Manuscript received September 30, 2012; revised January 22, 2013.

Pradeep Chaudhari is with the Tibco Software, Pune, India (e-mail: chaudharipradeep9@gmail.com).

Ravindra Dingankar is with John Deere Technology Center India, Pune, India (e-mail: chaudharipradeep9@gmail.com).

Roshan Chaudhari is with the Imagination Technology, Pune, India (e-mail: rgc183@gmail.com).

illustrates the use of active learning techniques to improve efficiency of error estimator [2].

Artificial neural network (ANN) can be applied for forecasting short-term load for a large power system. Authors, K. Y. Lee, Y. T. Cha and J. H. Park, have considered Weekday and weekend-day patterns. They have proposed a nonlinear load model and have tested many structures of an ANN for short-term load forecasting. Past loads and load forecast for the particular day were the inputs and outputs respectively, of the ANN. The network with one or two hidden layers was tested with various combinations of neurons, and results were compared in terms of forecasting error. The neural network, when grouped into different load patterns, gave a good load forecast [3].

### III. SYSTEM ARCHITECTURE

Our hybrid model shown in Fig. 1 is composed of two models: Pattern Based Matching, Error Correction Model. The derived efficiency of a keyword (numeric) obtained from first model is fed to Error Correction model along other numeric feature inputs to calculate optimum value of CPC for the keyword.

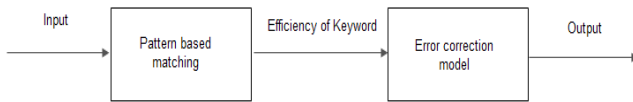


Fig. 1. System architecture

### IV. MODEL IMPLEMENTATION

This section discusses detailed implementation of both the models. First one is memory based model and second is error-correction model.

#### A. Pattern Based Matching

Pattern based Model considers two factors such as quality score, search volume. The quality score of a keyword is calibrated by regarding various factors such as relevance of ad, position of ad, rank of the ad. Search volume is one of the major factors in determining popularity of the keyword. Higher the search volume more is the popularity. By considering these relationships we are calculating efficiency of a keyword by formula:

$$\text{Efficiency} = k \times \text{Search Volume} \times \text{Quality Score} \quad (1)$$

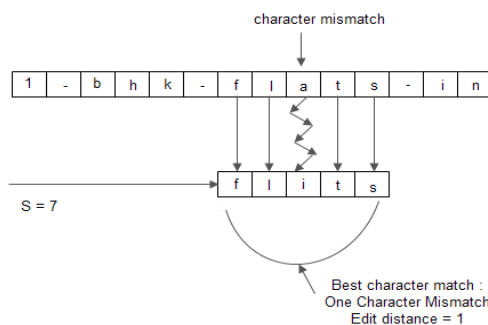


Fig. 2. String edit distance example

Now, we are training our Pattern based neural network with set of keywords and their calculated efficiency factor.

Our Pattern based model will store training data and whenever it is given a new keyword, it will apply pattern based searching and will select a keyword with best match and so its efficiency factor as shown in Fig. 2. [4] This model will chose a keyword and efficiency value from trained data having minimum edit distance for new keyword. [5]

#### B. Error-Correction Model (With Supervised Learning Approach)

This model predicts CPC based on given inputs. The inputs for error correction model are as follows:

- Clicks - No. of clicks received for the particular ad.
- Impressions - No. of times ad shown.
- CTR- Click through Ration which is no. of clicks per thousand impressions of an ad.
- Average Position - Position of ad in display network (Lower the position, Higher the rank of an ad)
- Efficiency of Keyword - Derived variable from first phase

The model in Fig. 3 contains 3 layers; input, output and a hidden layer. The input layer contains 5 neurons respectively for clicks, impressions, CTR, average position and efficiency. The hidden layer contains 3 neurons. The last layer has one output neuron that is for CPC.

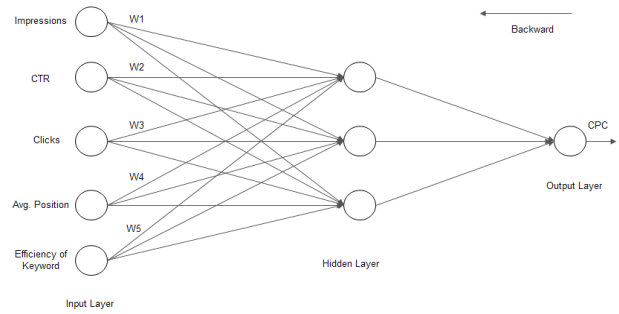


Fig. 3. Neural network

The number of neurons in hidden layers is restricted between no of input neurons and no of output neurons for optimal performance. The model is trained with historic data. [6] Initially all the weights are initialized randomly. The final output is given by the formula:

$$\text{Output} = W_j \times X_m \times W_i \quad (2)$$

where

- $W_j$  is the weight vector at hidden layer.
- $X_m$  is the input vector at input layer.
- $W_i$  is the weight vector at input layer.

For the neuron in the first hidden layer, the formula can be expanded as

$$Hx = W_1 \times \text{impressions} + W_2 \times \text{CTR} + W_3 \times \text{clicks} + W_4 \times \text{avg. position} + W_5 \times \text{efficiency} \quad (3)$$

This computed value will be used in next layer and finally Output value (CPC) is computed. Now, we are applying nonlinear regression technique here. We compute the error at any particular node of a model. This error, standard metric is the sum of squared errors. [7]

$$E_i = \sum_{d \in D} (t_{id} - o_{id})^2 \quad (4)$$

The ultimate goal is to minimize the output error at any particular node by adjusting the weights. We calculate the difference between actual value of CPC and calculated value. The difference is used to modify the weights and hence error in next inputs. The quality of model depends upon the training set data. More is the training data provided to neural network, more the accuracy and less is the error in value of CPC. [8]

$$W_{ij} = -\eta \frac{\partial E_i}{\partial W_{ij}} \quad (5)$$

In this equation  $\eta$  is the learning rate. Hence our new weight is given by formula:

$$W_{\text{new}} = W_{\text{old}} - \eta \frac{\partial E_i}{\partial W_{ij}} \quad (6)$$

Once error-correction networking is fully trained, we can combine both the networks to predict values for new terms/keywords coined by providing expected details by the network.

For example: We can build similar kind of a Memory-based learning model like keyword v/s derived variable models for other properties e.g. keyword-conversions etc. using this kind of model, we can predict new input values which will be used for our CPC prediction models.

## V. WHY HYBRID MODEL

We can define the success of an advertisement by percentage of conversions, indirectly by defining number of leads for the sales. Success of an advertisement does not totally depend upon Impressions or no. of clicks made but it also depends upon the quality of your product/service and the quality of relevant landing page. For e.g. Let us consider keyword, 'flats in region x'. Suppose it will lead to an advertisement say, y. On click of this advertisement, one will go to landing page, depending upon quality, relevance of page with respect to advertisement, the person may or may not stay on it and eventually fill the form (If the person fills the form, this will be considered as lead). The training set can belong to different domains, and each domain keywords will have different features/characteristics data. For this proposal, we are considering specific (particular) domain instead of generic solution. Here we are going to train the neural network using historical data. This historical data obtained from various sources can have different kind of hidden strategies of setting CPC for different domains. To handle such complex relationship of keywords with numeric data we have divided this model into two learning approaches. The first approach is Pattern Based Model and the Second Approach is Error-correction Model. By applying these two approaches basically we tried keeping separate model for keyword and rest of the numeric data.

## VI. FUTURE ENHANCEMENTS

This paper addresses only specific domain in online advertising. The neural network can be trained for advertisements of various domains. This model with more enhancements can be used to predict CPC in real-time using large set of training data distributed over timeframe. This will

empower users to predict values based upon social data e.g. holidays, Christmas period, Diwali seasons, as also the time during the day when users are mostly active. It's obvious that such kind of model needs lot of diversified training data, but this can be implemented with systematic data analysis and hybrid neural network.

## VII. CONCLUSION

This paper surveyed the application of neural networks in online advertising for predicting CPC values for keywords; it demonstrated how we can integrate the two or many network models. As CPC values would be closer to those of similar keywords and since these values will be further refined by second backward propagation based on neural network, cost would be reduced by avoiding high CPC values of keywords and vice versa, thus increasing efficiency. Although neural networks are not perfect in prediction, they outperform all other methods in every field for minimizing error in results in dynamic environment.

## REFERENCES

- [1] S. Ozawa, "A memory-based neural network model for efficient adaptation to dynamic environments," *IEEE Fuzzy Systems*, vol. 1, pp. 437-442, July 2004.
- [2] A. Abdulkader, "Low cost correction of ocr errors using learning in a multi-engine environment," in *Proc. IEEE International Conference Document Analysis and Recognition*, July 2009, pp. 576-580.
- [3] K. Y. Lee, "Short-term load forecasting using an artificial neural network," *IEEE Power Systems*, vol. 7, issue 1, pp. 124-132, Feb. 1992.
- [4] G. A. Carpenter, "The ART of adaptive pattern recognition by a self-organizing neural network," *IEEE Computer*, vol. 21, issue 3, pp. 77-88, March 1988.
- [5] A. Mirzaei, H. Zaboli, and R. Safabakhsh, "A Neural Network String Matcher," in *Computer Analysis of Images and Patterns*, Springer, 2007, vol. 4673, pp. 784-791.
- [6] R. H. Nielsen, "Theory of the back propagation neural network," *IEEE Neural Networks*, vol. 1, pp. 593-605, 22 Jun. 1989.
- [7] Y. Singh and A. Chauhan, "Neural networks in data mining," *JATIT*, vol. 5, no. 6, 2009.
- [8] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, Prentice Hall, 2nd Edition, 1995.



Pradeep Chaudhari has pursued his B.Tech from College of Engineering, Pune in Information Technology in 2011. He is currently working with Tibco Software as Associate Consultant since June 2011. His one paper has been accepted for the conference which will held in Hong Kong. That paper is "Optimization tool for PPC advertisement model" which will get published in ASME digital library. His work can be viewed at [pradeepchaudhari.com](http://pradeepchaudhari.com)



Ravindra Dingankar has pursued his B.Tech from College Of Engineering Pune in Information Technology in 2011. He is currently working with John Deere Technology Center India as an Associate SAP Consultant since July 2011. He is working on the warranty module in SAP.



Roshan Chaudhari has pursued his B.Tech from College of Engineering Pune in Information Technology in 2011. He is currently working with Imagination Technologies as a Test Development Engineer since June 2011. His one paper has been accepted for the conference which will held in Hong Kong. That paper is "Optimization tool for PPC advertisement model" which will get published in ASME digital library.