

# A NEW STRATEGIC APPROACH TO SEARCH ADVERTISING RECOGNITION THROUGH FEATURE ENGINEERING

Guide: Ms.HIMA BINDU B

Assistant Professor, Department of CSE, **Malla Reddy Engineering College For Women**,  
TELANGANA, India.

**D.KHYATHI REDDY<sup>1</sup>, G.SHRAVANI<sup>2</sup>, B.HYMAVATHI<sup>3</sup>, K.SWATHI<sup>4</sup>**

B. Tech Pursuing, Department of CSE, **Malla Reddy Engineering College For Women**,  
TELANGANA, India.

## ABSTRACT

Feature engineering is one of the key applied sciences in the lookup of search advertising and marketing recognition. Most of the present search advertising and marketing techniques are chosen in accordance to the prior knowledge, which is too subjective to be popularized. Taking the commercial of Ali search marketing as the lookup object, a function processing technique based totally on the pre-analysis of save and consumer facts is put forward, and then the conversion charge is anticipated with XGBoost (eXtreme Gradient Boosting). Experiments exhibit that in contrast with different priori Feature Engineering, the proposed approach can appreciably enhance the prediction results.

**INDEX TERMS:**Applied Sciences,Advertising,Commercial ,Anticipated.

## 1. INTRODUCTION

Search advertising is a common way of Internet marketing. Businesses purchase specific keywords according to the characteristics of the goods. When users enter these keywords, the corresponding advertising goods will be displayed in the pages that the user sees. The conversion rate of search advertisements is used as an index to measure the effect of advertising transformation, that is, the probability of advertising products being bought by users after clicking. With the rapid development of Internet, search advertising has become

more and more popular in Internet advertising, and has become one of the most important business models in the Internet industry. In Feature Engineering, traditional feature processing methods are linear combination of original features, one-hot coding and so on. It is difficult to improve the recognition rate in traditional ways. This paper, taking Ali search advertising as the research object, proposes a feature processing method based on store and user data pre-analysis, which aims to pre-analyze the features, that is, the first prediction processing of the features of

users and stores, and as a new feature. The results of this experiment take the size of Logarithmic Loss (Logless) as the evaluation standard. In general, we must correctly handle the features and reduce the Logless value as much as possible, which is the next problem we need to solve.

## 2.LITERATURE REVIEW

### 1) Spam Review Detection Techniques: A Systematic Literature Review

AUTHORS: Naveed Hussain 1,2, Hamid Turab Mirza 1,\*, Ghulam Rasool 1, Ibrar Hussain 2 and Mohammad Kaleem 3

We consider the problem of planning the ISS cosmonaut training with different objectives. A pre-defined set of minimum qualification levels should be distributed between the crew members with minimum training time differences, training expenses or a maximum of the training level with a limitation of the budget.

First, a description of the cosmonaut training process is given. Then four models are considered for the volume planning problem. The objective of the first model is to minimize the differences between the total time of the preparation of all crew members, the objective of the second one is to minimize the training expenses with a limitation of the training level, and the objective of the third one is to maximize the training level with a limited budget. The fourth model considers the problem as an

$n$ -partition problem. Then two models are considered for the calendar planning problem.

We consider the problem of planning the ISS cosmonaut training with different objectives. A pre-defined set of minimum qualification levels should be distributed between the crew members with minimum training time differences, training expenses or a maximum of the training level with a limitation of the budget. First, a description of the cosmonaut training process is given. Then four models are considered for the volume planning problem. The objective of the first model is to minimize the differences between the total time of the preparation of all crew members, the objective of the second one is to minimize the training expenses with a limitation of the training level, and the objective of the third one is to maximize the training level with a limited budget. The fourth model considers the problem as an  $n$ -partition problem. Then two models are considered for the calendar planning problem. Abstract We consider the problem of planning the ISS cosmonaut training with different objectives. A pre-defined set of minimum qualification levels should be distributed between the crew members with minimum training time differences, training expenses or a maximum of the training level with a

limitation of the budget. First, a description of the cosmonaut training process is given.

Then four models are considered for the volume planning problem.

The objective of the first model is to minimize the differences between the total time of the preparation of all crew members, the objective of the second one is to minimize the training expenses with a limitation of the training level, and the objective of the third one is to maximize the training level with a limited budget. The fourth model considers the problem as an  $n$ -partition problem. Then two models are considered for the calendar planning problem

Abstract

We consider the problem of planning the ISS cosmonaut training with different objectives. A pre-defined set of minimum qualification levels should be distributed between the crew members with minimum training time differences, training expenses or a maximum of the training level with a limitation of the budget.

First, a description of the cosmonaut training process is given. Then four models are considered for the volume planning problem. The objective of the first model is to minimize the differences between the total time of the preparation of all crew members, the objective of the second one is to minimize the training expenses with a

limitation of the training level, and the objective of the third one is to maximize the training level with a limited budget. The fourth model considers the

## 2. EXISTING SYSTEM

In Feature Engineering, traditional feature processing methods are linear combination of original features, one-hot coding and so on. It is difficult to improve the recognition rate in traditional ways. Businesses purchase specific keywords according to the characteristics of the goods. When users enter these keywords, the corresponding advertising goods will be displayed in the pages that the user sees.

### DISADVANTAGES OF EXISTING SYSTEM:

- The conversion rate of search advertisements is used as an index to measure the effect of advertising transformation, that is, the probability of advertising products being bought by users after clicking.
- Data is the carrier of information, but the original data contains a lot of noise, and the expression of information is not concise enough.

## 4. PROPOSED SYSTEM

This paper, taking Ali search advertising as the research object, proposes a feature processing method based on store and user

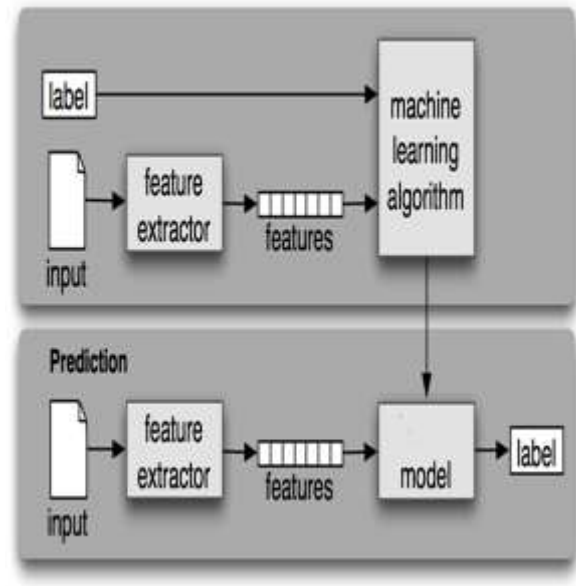
data pre-analysis, which aims to pre-analyze the features, that is, the first prediction processing of the features of users and stores, and as a new feature. Calculate the correlation between each feature and response variable. The commonly used methods in engineering include Pearson coefficient and mutual information coefficient, Pearson coefficient can only measure linear correlation, and mutual information coefficient can measure various correlations well.

**ADVANTAGES OF PROPOSED SYSTEM:**

- The model of a single feature is constructed, and the feature is selected by the accuracy of the model, and then the final model is trained when the target features are selected.
- After feature selection, features are selected again if the user id and user characteristics are combined to obtain a larger feature set and then select a feature, this practice is more common in recommendation system sand advertisement systems.

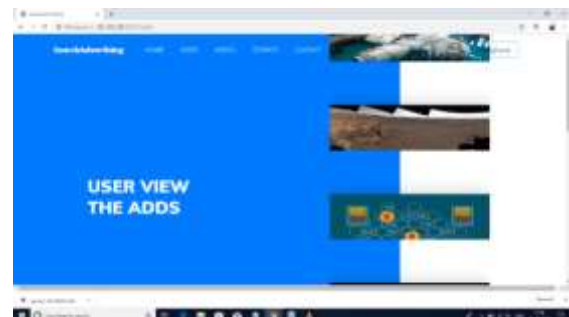
**5.SYSTEM**

**ARCHITECTURE**



**6.RESULTS**

User View Adds



User View Unique Add



### User Search Adds



### User give Rating to Adds



## 7.CONCLUSION

In this paper, the prediction of advertising conversion rate is studied by experiments. For training data, the selection of features has a crucial impact on the prediction performance of the model. In the prediction of advertising click rate, there are many features that can be used, including basic data, advertising commodity information, user information, context information and store information. To achieve a good accuracy rate, the model should be fully excavated as a feature, and the better the combination of these features, the better the performance of the model. The feature learning method proposed in this paper estimates the ad click rate, only considering that the full advertising data is not

considered to show inadequate advertising. In the next work, how to estimate the click rate of sparse advertising from the point of view of characteristic learning is a problem worthy of study, and it is also an urgent problem to be solved at present. At the same time, we should also pay attention to the research of different models integration.

## 8.REFERENCES

- [1] [1] Zeng D S, Huang F L, Pan C D. Feature Engineering for Product Review
- [2] Spam Identification[J]. Journal of Fujian Normal University, 2017.
- [3] [2] Cormack G V. Feature engineering for mobile (SMS) spam filtering[C]
- [4] International ACM SIGIR Conference on Research and Development in
- [5] Information Retrieval. ACM, 2007:871-872
- [6] [3] Li J, Liang X, Ding W, et al. Feature engineering and tree modeling for
- [7] author-paper identification challenge[C]// Kdd Cup 2013 Workshop.
- [8] 2013:1-8.
- [9] [4] Stankevich M, Isakov V, Devyatkin D, et al. Feature Engineering for

- [10] Depression Detection in Social Media[C]// International Conference on
- [11] Pattern Recognition Applications and Methods. 2018:426-431.
- [12] [5] Chen J H, Li X Y, Zhao Z Q, et al. A CTR prediction method based on
- [13] feature engineering and online learning[C]//International Symposium on
- [14] Communications and Information Technologies. IEEE, 2018.
- [15] [6] Chen T, Tong H, Benesty M, et al. xgboost: Extreme Gradient
- [16] Boosting[J]. 2015.
- [17] [7] Chen T, Guestrin C. XGBoost: A Scalable Tree Boosting System[C]//
- [18] ACM SIGKDD International Conference on Knowledge Discovery.

#### AUTHOR

**Ms.Hima Bindu B** Assistant Professor  
Department of CSE(CS) MallaReddy  
Engineering College for Women,  
Hyderabad, himasanagapalli@gmail.com.