Keyword Selection on Google Ads

[1]Anushree, [2]Roja R and [3]Anand Kumar B
[1,3] Department of CSE, National Institute of technology manglore, Karnataka, India-575025.
[2]Department of CSE,Assistant professor, AMC Engineering College, Bengaluru, Karnataka,India-560083.
[1]anushreeg.anu95@gmail.com, [2]rojaram@gmail.com, [3]anandkumar.bm@gmail.com

Article Info

Jenitta J and Swetha Rani L (eds.), *International Conference on VLSI, Communications and Computer Communication*, Advances in Intelligent Systems and Technologies,
Doi: https://doi.org/10.53759/aist/978-9914-9946-1-2_20
©2023 The Authors. Published by AnaPub Publications.

Abstract — Google Advertising is a publicity agency that provides marketers with advertisements. By choosing keywords relevant to their ad material, advertisers fit the user's search terms and push advertising. Keywords will decide the type of users being pushed by an advertiser, the efficacy of the ad promotion, and therefore the ad product's sales. The main objective is to automatically choosing keywords that are satisfactory to advertisers from an outsized number of keywords given by Google Advertising. But there's not an excessive amount of time for the framework to make a decision whether keywords are chosen and to pick the proper keywords within the shortest time. Therefore, a model structure which can obtain some helpful keywords for advertisers is built also to accomplish this multipurpose task, an enhanced method of multi-objective particle swarm optimization is introduced. Many technical challenges need to be solved to accomplish this multi-objective mission, such as the issue of mixed language, the problem of data imbalance, the issue of obtaining features from the collection, and so on. The mixture of evolutionary computation, deep learning, machine learning and text processing approaches is used here to solve the issue of keyword selection.

Keywords – Google, Deep Leaning, Machine Learning, Text Processing.

I. INTRODUCTION

Advertising is a marketing transmission to advertise or to sell a product, service or concept using a publicly endorsed, non-personal message. Advertising supporters are usually organizations seeking to advertise their goods or services. An advertising agency is an organization which is devoted to designing, organizing and managing advertising and often other forms of promotion and marketing for its customers, it is sometimes referred to as a creative agency or an ad agency.

In 2000, Google introduced AdWords. AdWords marketers paid for the service monthly at first and Google would set up their campaign and handle it. Google soon launched the AdWords self-service platform in order to satisfy small businesses and those who wished to run their own campaigns. At present, the name Google AdWords has changed to Google Ads.

Google Advertising will drive ads from marketers to websites affiliated with Google and few networking sites, for instance Facebook, Google, and the rest. Via marketers' ads, it is beneficial. Text, photographs and videos are included in the form of the advert. Google Advertisements will respond rapidly to marketers with advertising data concerning clicks, views, cost per click, and afterwards marketers promote their advertisements via Google Ads. Later, some more realistic strategies are tailored to improve the advertisement's competitiveness when the advertisers get the details.

Google Ads system is based partly on cookies and partly on keywords determined by advertisers. Google uses these characteristics to place advertising copy on pages where they think it might be relevant. Advertisers pay when users divert their browsing to click on the advertising copy. Adverts can be implemented locally, nationally, or internationally.

Purpose

The main intention of the proposed framework is to present a method of corpus selection to resolve the mixed issue of different languages in the keywords, word embedding method to figure out keyword presentation, reruns that clear the difficulty of data imbalance, enhanced CNN to solve the problem of classification, and a algorithm of multi-objective particle swarm optimization (MOPSO) is used to get CNN neural structure search so that the result of classification is enhanced and the training time is decreased.

II. RELATED WORKS

Previously so many works carried out on keyword selection, but no research work has been carried out on finding the right keyword for advertisers and also finding the reasons is the main challenging task, and here we provide a detail work from various sources.

The work carried out by A. L. Murphy et al., [1] have focused on comparison between user acquisition and actions during and after a Google Advertising campaign on the Headstrong website. The number of user visits increased by more than five times through a Google Advertising campaign designed to guide men to the Headstrong website. However, the engagement of users who responded to the ad campaign was significantly lower than that of users who visited the website through other methods of acquisition, likely reflecting the ad campaign's non-specific online targeting of men. The general

online targeting of men to encourage the mental health of men seems to have little significance.

Zink et al., [2] have focused on to determine the frequency of pruritus in Germany and to determine what the most common associated complaints are by means of a search volume review by Google. For this research, the Google search engine review provided insight into the surprisingly large number of unmet medical needs of persons suffering from pruritus in Germany, in particular those of unique body localizations.

W. C. Serrano et al., [3] have demonstrated on the effectiveness of using online ads to provide targeted messages of prevention related to indoor tanning and skin cancer. The effect of these ads on health habits, however, remains uncertain. Further studies of this method are needed to examine the features of messages that produce views and clicks, and ultimately to decide whether habits are effectively modified by this form of intervention.

S. M. Rezaeinia et al., [4] presented Improved Word Vectors (IWV), a novel approach that improves the precision of pre-trained word embedding in sentiment analysis. This technique is based on tagging strategies for Part-of-Speech (POS), lexicon-based approaches, word location algorithm, and methods for Word2Vec/GloVe. Via various deep learning models and benchmark sentiment datasets, they have checked the accuracy of this process. The results of the experiment show that Improved Word Vectors (IWV) are very efficient for evaluating sentiment.

Yiping Li et al., [5] suggested the improvement of CNN to address the not balanced classification of text sentiment. By advancing a new over-sampling technique, they are trying to resolve the imbalanced data issue. The proposed sampling method directly generates synthetic texts from word spaces, unlike current over-sampling approaches that produce minority-class samples from numerical feature space. Several tests are performed to check the feasibility of the proposed method of producing lexicons, the learning system, and the method of over-sampling. Experimental results show that the induced sentiment lexicons are interpretable, and for imbalanced and domain-specific text sentiment classification, the suggested model is found to be accurate.

N. V. Chawla et al., [6] demonstrates that a mixture of their method of over-sampling the minority (abnormal) class and under-sampling the majority (normal) class will achieve better output of the classifier (in ROC space) than the majority class only under- sampling. It also suggests that a combination of their method of over-sampling the minority class and under-sampling the majority class will achieve better output of the classifier (in ROC space) than varying the loss ratios in Naive Bayes in Ripper or class priors. The development of synthetic minority class examples includes this process of over-sampling the minority class. C4.5, Ripper, and a Naive Bayes classifier are used to conduct experiments. Under the Receiver Operating Characteristic curve (AUC) and the ROC convex hull strategy, the method is evaluated using the field.

Li et al., [7] and an end-to-end entity classification systembased on the neural network model was proposed. First the comparison model used long short-term memory to describe the forms of and entity referred to from the sentences it comprises, to be precise. Secondly, to compensate for the current structures of entity classification, they suggested a fusion model to fuse the forms of multiple mentions. In the method, the experimental results showed the necessity and efficacy of each module. The author claims that their proposed approach is a strong complement to the current entity classification systems.

F. Li et al. [8] proposed A method of feature reduction based on two-stage clustering of features (TSFC) that is applied to short text classification. Feature terms are replaced by related feature clusters, and the vector space dimension is greatly reduced. To test the efficacy of this strategy, multiple classifiers are used. The results show that the methodology largely addresses the dimensional catastrophe and that the accuracy of short text classification can be greatly improved.

Edo-Osagie et al. [9] proposed an attention-based approach to short text classification that they have developed for the realistic application of Twitter mining to track public health. Its purpose is to automatically filter tweets that are related to asthma/difficulty breathing syndrome. They defined a bi-directional architecture of the Recurrent Neural Network with an attention layer called ABRNN) that allows the network to weigh words differently in a Tweet depending on their perceived value. The author demonstrated the performance of the ABLSTM in the real-world application of public health surveillance and contrasted the results with Public Health England's (PHE) real-world syndromic surveillance data. A strong positive association was found between the surveillance signal for ABLSTM and the syndromic surveillance data for real-world asthma/difficulty breathing. The ABLSTM is a valuable instrument for public health surveillance tasks.

R. Türker et al. [10] proposed A novel Knowledge-Based Short Text Categorization (KBSTC) probabilistic model that does not require any labelled training data to classify a short text. This is done by exploiting broad knowledge-based entities and categories that are further embedded in a standard vector space for which a new embedding model of the entity and category is proposed. The experimental results show that their strategy significantly outperforms the classification approaches that do not involve any labelled data, while the results of the supervised approaches are similar to them.

G. L. F. da Silva et al. [11] introduced a method which uses a deep learning approach in combination with an evolutionary technique to decrease the FP number. In order to increase network efficiency and eliminate the need for manual search, the particle swarm optimization (PSO) algorithm was used to optimize the network hyper parameters in the convolutional neural network (CNN). The results show that the PSO algorithm has high performance potential in defining ideal CNN hyper parameters for candidate lung nodule classification into nodules and non-nodules, increasing sensitivity rates in the FP reduction process of CAD systems.

L. Mostafa et al. [12] focused on the system that recommends a webpage's keywords on the basis of frequent terms. The approach used in this description is term frequency to describe the frequent terms. To verify the system results, an experiment is executed; and the outcome of the new method is compared to Google's adword tool. The accuracy of the approach proposed is 82.4%, which is considered a positive result.

Y. Sun et al. [13] has proposed by removing the restrictions on the number of convolutional layers and pooling layers from the conventional CAE, a versatile Convolutional Auto-Encoder (FCAE) is accessible. By leveraging particle swarm optimization, they have developed an architecture discovery approach that is able to automatically search for the excellent architectures of the introduced FCAE with far fewer arithmetic resources and in the absence of manual intervention. On four widely used image classification data sets, they have evaluated the proposed method. Experimental findings suggest that the proposed solution, including state-of-the-art algorithms, greatly outperforms peer competitors.

Y. Shen et al. [14] proposed a cluster-based method of semantic expansion based on hierarchical agglomerative clustering that incorporates word embedding into cluster embedding effectively to gain more semantic knowledge. The superiority of their proposed approach for short text classification tasks is demonstrated by experimental findings on two benchmark datasets.

III. SYSTEM ARCHITECTURE

The conceptual model that describes a system's structure, actions, and more views is a system architecture. A description of architecture is a systematic description and representation of a system, structured in a manner that facilitates thinking about the system's structures and behaviors. A system architecture can consist of components of the system and established sub-systems, which will work together to implement the overall system. In order to explain system architecture, attempts have been made to formalize languages, collectively called architecture description languages (ADLs) Fig 1 represents the Architecture of System and the details of the system is explained below.

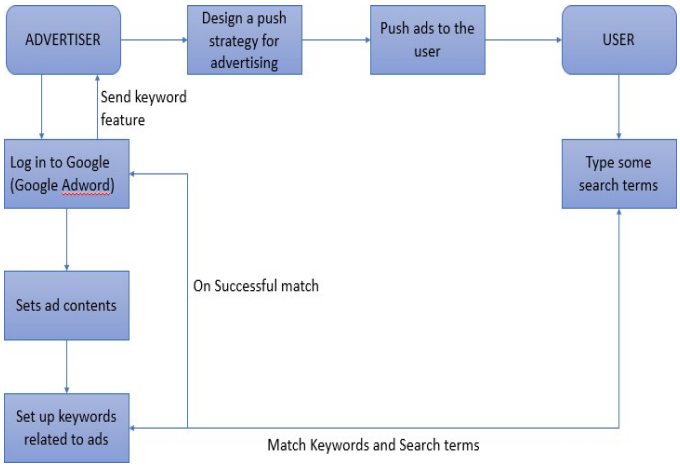

Fig 1. System Architecture.

The advertiser logs in to Google (Google Adwords) and sets the content of the advertisements and set up the watchword which are associated with their content of the advertisements of the system. The set watchwords are balanced by the terms the user searches for in the internet service provider. On successful match, Google Advertisement will post the keyword characteristics to the advertisers, based on the keyword features, advertiser will design the push strategy for advertising, and the ads will be pushed to the more specific users in the event of high-level matching.

The advertisers will modify the plan corresponding to the keyword characters. If the watchwords are improper, (which is considered as negative watchword) then they willbe deleted from Google Ads by advertisers.

IV. ALGORITHM

The methodology is defined composition of reasonably related practices, methods, and processes that identify the best way to plan, develop, control, and deliver a project throughout the continuous implementation process until successful completion and termination.

Step 1: *Initialize*
Randomly generate x in feasible region
$GHA = x$
$PHA_i = x_i$
Step 2: *Select gbest and* $pbest_i$
gbest = MO_selection(GHA)
$pbest_i$ *= MO_selection(* PHA_i*)*

Step 3: *Update $x_i(t)$ according to (1) and (2)*
Step 4: Update PHA_i and GHA
For i=1: popsize
PHA_i (t)=non-dominated-selection(PHA_i (t-1), x_i(t))
if size (PHA_i (t))>maximal number of PHA
Remove the extra particles from PHA_i (t)
end
GHA(t)=non-dominated-selection(PHA(t))
if size (GHA(t))>population size
Remove the extra particles from GHA(t)
end if the termination condition is satisfied
Terminate
or
Repeat Step 2-5
End
Step 6: Select the optimal solution
Output the best particle in GHA by **Final_MO_selection.**

V. IMPLEMENTATION OF MODULES

An Implementation is the realization of a strategy, concept, model, design, specification, standard, algorithm, or policy by an application or execution. Implementation of the system typically benefits from high levels of user engagement and support for management. There are some positive effects from user engagement in the design and operation of information systems. First, if consumers are heavily involved in the design of systems, they transfer opportunities according to their priorities and business requirements to shape the system, and more possibilities to influence the result. Second, they are more likely to respond to the process of transition positively. The incorporation of user knowledge and skills contributes to better solutions. In this Project work, it uses five modules, these modules are explained below.

The use of Word Embedding

The texts translated into numbers are Word Embeddings, and different numerical representations of the same text may be present. In its raw form, several Machine Learning algorithms and almost all Deep Learning Architectures are unable to process strings or plain text. In order to perform some kind of work, whether it is classification, regression, etc., they require numbers as inputs in general terms. And it is imperative to extract information from it and create applications with the enormous amount of data thatis present in the text format.

Also, the issue of handling several artefacts, when considering the problem of NLP. Therefore, methods such as Frequency of Word, Document Frequency, Text Frequency-Inverse Document Frequency etc., is not useful to conventional natural language processing. So therefore, it selects the common method of embedding words in present years to transform keywords into vectors of terms. In this way it is possible to addnot only the keyword feature, but also to extract the keyword characteristics. It should be noted that Word2Vec, Glove, and Fast text are the most widely used methods for word embedding. According to various application needs, distinct word embedding methods are chosen. This contrasts these three methods for the convenience of comparison and selects the finest embedding word approach to plan for the upcoming examination.

Corpus Selection

In Google Advertising, the keywords include the language of more than one country. It converts these watchwords into vectors of words by choosing corpora which comprising various national languages and by utilizing the previous method embeddingof words to solve the problem that different languages occur in the common corpus. In this way, they can be put together to extract features even though a keyword includes several different languages. The keywords which are used here is Chinese and English, so some other corpora that include two languages are checked, such as Sikuquanshu, Weibo, Baidu, Wikipedia, and so on. Additionally, the aim is to address the difficulties of advertising watchwords, therefore, enough advertising information is needed for these corpora. Using other corpora to translate keyword's words into vectors of words, however can arise in the words of watchwords not being used in these collections. In this situation, the random value of the word vectors must be represented. It should be remembered that there is no uniform number of terms found in the keyword, so it is important to set the length to a uniform length. You may use 0 to fill it up if the length does not exceed the specified length. The improved CNN explained below enhances this problem. In addition, various applications need to be chosen by different corpora. To prepare for the next experiment, it matches and chooses the finest collections.

Processing the Imbalanced Data

The amount of keywords that can meet advertisers' needs is very limited, which is already stated in the introduction. However, the keywords which are provided to advertisers by Google Ads, are very large. In this case, direct classification would eventually outcomes in a poor impact due to the imbalance of data. So resampling is used here to enhance this

problem. The three most commonly used problems to address this situation are under-sampling, smote, and borderline-smote. The best approach is used as the method of resampling by dealing with these three approaches separately.

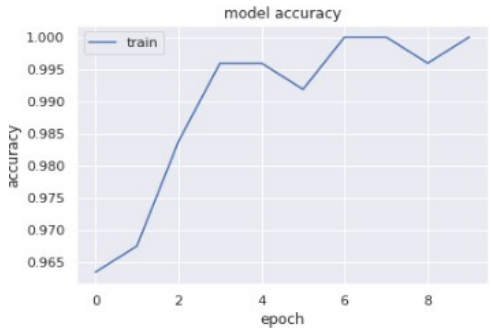

Fig 2. Accuracy of the Balanced data.

The Fig 2 represents the accuracy of balanced data after applying the resampling technique to imbalanced data. X-axis indicates the epoch and Y-axis indicatesthe accuracy. As increase in the value of epoch, accuracy will increase.

The Improved CNN

Due to its capability to automatically train the network and draw out the characteristics, the CNN is commonly used to deal with NLP issues, so here it selects Convolutional Neural Network to choose watchwords. Although, not only do the Watchwords which are referred here need to remove the characteristics of their own language, since they also require to incorporate few of the relevant characteristics of the network keywords, so conventional CNN needs to be enhanced.

The MOPSO based Improved CNN

The collection of keywords can be achieved with the enhanced CNN which is outlined above. However, it has a complicated network structure with several restrictions when CNN is used for keyword choice. The key issue to be solved is how to pick a network layout that can get keywords quickly and more reliably from several limitations. Therefore, MOPSO is used for network optimization.

Fig 3. Training and Validation Accuracy of the Proposed Model.

The Fig 3 shows the training and validation accuracy of the proposed system. X-axis indicates the epoch and Y-axis indicates the accuracy. As increase in the value of epoch, accuracy will increase. The graph indicates the accuracy of the proposed system is 95%.

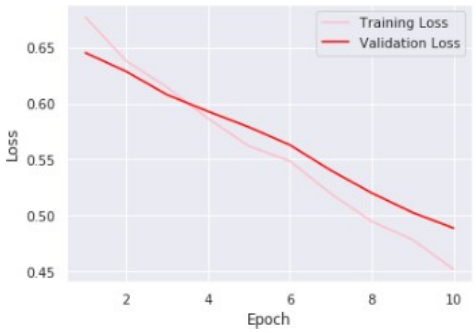

Fig 4. Training and Validation Loss of the Proposed Model.

The Fig 4 shows the training and validation loss of the proposed system. X- axis indicates the epoch and Y-axis indicates the loss. Training loss is calculated as a moving average over 1 epoch, whereas the validation loss is calculated after the learning phase of the same epoch.

VI. RESULTS AND ANALYSIS

In line with the empirical outcomes, the consolidation of machine learning, deep learning, transformative computing, and embedding of text methods, including primary insight for instance reruns, PSO method, and CNN, effectually resolve the problem of struggling in choosing watchwords. Along with a precise intensity, the productivity of choosing watchwords and the standard of culled watchwords have been enhanced in a less period of time, a sequence of obstacles in choosing watchwords for advertising are strongly gets the better results. In the real time, when we select the item name, it searches in the keywords and displays whether that keyword is Positive Keyword or Negative Keyword for gift (best on successful match keyword). When compared to the existing system accuracy of the model in the proposed system is increased from 66.8% to 95%.

VII . CONCLUSION S AND FUTURE ENHANCEMENT

Keyword selection is an issue with the categorization that certain acceptable watchwords for marketers perhaps chosen among several suggested watchwords. Since the amount of terms in a watchword is minor, short text classification methods are used by many experts to solve this difficulty. The mixture of machine learning, deep learning, transformative computing, and embedding of text approaches, including primary ideas for instance reruns, PSO method and CNN, was successful in solving the problem of keyword selection difficulties, according to the experimental results, and to some degree, the effectiveness of keyword selection along with the accuracy of identified watchwords turned out to be boosted in a limited period of time. a sequence of obstacles in choosing watchwords for advertising are strongly gets the better resultsIn this work, there are still some difficulties that required to be overcome in detail. Firstly, the volume of input which are utilized in this framework is relatively small, since data generation requires a lot of capital and time. Next, in a shorter period of time, the main purpose is to pick the correct keywords, so it must be presumed that all influencing variables are the same. Including all these, it has some restrictions. In practice, when selecting keywords, there are several parameters to be considered, for instance, the kind of advertising, the design of the advertisement that the marketer wants, along with that the attributes of the advertising message beneficiary. These variables may as well, in fact, have a major effect on the output of the keywords selected.

In future studies, the complications in choosing watchwords for Google Advertising would be additionally determined and promoted including amount of input rises and the depth of study promoted. This will assist advertisers can avoid a lot of capital investment and brings considerable economic benefits and value to them.

Reference

[1] L. Murphy, S. Peltekian, and D. M. Gardner, “Website analytics of a google ads campaign for a men’s mental health website: Comparative analysis,” JMIR Mental Health, vol. 5, no. 4, pp. 12428, October, 2018.

[2] Zink, B. Schuster, M. Ruth et al., “Medical needs and major complaints related to pruritus in germany: A 4-year retrospective analysis using google adwords keyword planner,” J Eur Acad Dermatol Venereol, vol. 33, no. 1, pp. 151-156, January, 2019.

[3] W. C. Serrano, M.-M. Chren, J. S. Resneck et al., “Online advertising for cancer prevention: Google ads and tanning beds,” JAMA Dermatology, vol. 152, no. 1, pp.101-102, January, 2016.

[4] M. Rezaeinia, R. Rahmani, A. Ghodsi et al.,“Sentiment analysis based on improved pre-trained word embeddings,” Expert Systems with Applications, vol. 117, no. 2019,pp. 139-147, March, 2019.

[5] Y. Li, H. Guo, Q. Zhang et al., “Imbalanced text sentiment classification using universal and domainspecific knowledge,” Knowledge-Based Systems, vol. 160, no. 2018, pp. 1-15, November, 2018.

[6] N. V. Chawla, K. W. Bowyer, L. O. Hall et al., “Smote: Synthetic minority over- sampling technique,” Journal of artificial intelligence research, vol. 16, no. 2002, pp. 321-357, June, 2002.

[7] Q. Li, J. Dong, J. Zhong et al., “A neural model for type classification of entities for text,” Knowledge- Based Systems, vol. 176, pp. 122-132, 2019.

[8] F. Li, Y. Yin, X. Mao, R. Shi, J. Shi, "Method of feature reduction in short text classification based on feature clustering", Appl. Sci., vol. 9, no. 8, pp. 1578, 2019.

[9] Provably Secure Authenticated Key Agreement Scheme for Smart Grid, V. Odelu, A.K. Das, M. Wazid and M. Conti, May: 2018.

[10] R. Türker, L. Zhang, H. Sack, M. Koutraki, "Knowledge-based short text categorization using entity and category embedding", Proc. Eur. Semantic Web Conf.,pp. 346-362, 2019.

[11] G. L. F. da Silva, T. L. A. Valente, A. C. Silva, A. C. de Paiva, M. Gattass, “Convolutional neural network-based PSO for lung nodule false positive reduction onCT images”, Comput Methods Programs Biomed., vol. 162, pp. 109-118, August.2018.

[12] L. Mostafa, "Webpage keyword extraction using term frequency", Int. J. Comput. Theory Eng., vol. 5, no. 1, pp. 174, 2013.

[13] Y. Sun, B. Xue, G. G. Yen, M. Zhang, "A particle swarm optimization-based flexible convolutional autoencoder for image classification", IEEE Trans. Neural Netw. Learn.Syst., vol. 30, no. 8, pp. 2295-2309, August. 2019.

[14] Y. Shen, Q. Zhang, J. Zhang et al., “Improving medical short text classification with semantic expansion using word-cluster embedding,” International Conference on Information Science and Applications, vol. 514, pp. 401-411, June, 2018.