Research on Personalized E-Commerce Recommendation Platform

Li Qiong
School of Computer Science and Technology
Hankou University
Wuhan, Hubei, China, 430212

Chai Fang
School of Computer
Central China Normal University
Wuhan, Hubei, China, 430079

Chen Li*
School of Computer
Central China Normal University
Wuhan, Hubei, China, 430079

Abstract—**With the deep integration of Internet and E-Commerce, using network technology to analyze and explore users' interest, and providing personalized services for them are becoming a popular application of online transactions. In order to solve the problems of low recommendation quality and poor real-time performance existing in personalized E-Commerce recommendation platform, this paper tries to establish a multi-hidden layer artificial neural network learning model to deeply explore the potential interest of users, so as to improve the recommendation quality. Meanwhile, this paper adopts the cloud computing technology to parallelize CPU clusters for improving recognition speed and realizing real-time demand.**

Keywords—Internet; E-Commerce; Personalized recommendation; Deep learning

I. INTRODUCTION

In the age of cloud computing and big data, it is nothing new to provide users with massive amounts of information. However, how to perceive and provide personalized information to users is always a difficult problem. Traditional information retrieval system is a passive information acquisition method based on user request, which cannot reflect the personalized needs of users, and the users also cannot perceive their personalized information actively. Thus, under the environment of cloud and big data, it is an urgent problem for some current user information service systems, to build intelligent E-Commerce recommendation systems for meeting users' personalized needs.

The personalized E-Commerce recommendation platform consists of three parts: user modeling, recommendation object modeling and recommendation algorithm. As one of the core technologies, user modeling is the basis and key to determine the performance of a recommendation system. In user modeling, the identification accuracy of user interest points directly affects the quality of user interest model and the recommendation quality of personalized recommendation system in the later stage. The existing methods of user's interest point identification can solve the problem of user's interest

point identification, but the training and learning process is generally time-consuming and tedious, and the accuracy rate of the interest point identification is not high [1]. How to quickly and accurately identify the interest points which can describe the characteristics of users is the key factor to improve the recommendation quality and speed of personalized recommendation system. The idea of deep learning and the technology of parallelization can solve the problems, such as the inaccuracy and comprehensiveness of user's interest point identification, and the poor real-time ability of identification recommendation, etc. The deep neural network model has a good effect on discovering user's interest point.

Therefore, this paper conducts a study on the identification of user interest points in the personalized E-Commerce recommendation system based on deep mining, and explores how to improve the recommendation quality and recommendation speed through deeply mining user' interest points.

II. RESEARCH STATUS AT HOME AND ABROAD

In recent years, domestic and foreign scholars and industry personnel in the field of personalized recommendation system have been paying attention to the deep mining and identification of user interest points. Various data mining technologies, parallelization technologies and optimization algorithms have been taken as the theoretical basis of their research to conduct in-depth mining and identification of user potential interest points.

The foreign scholars study extensively, the method is also quite practical. Agarwal D et al. proposed to use the user's purchase time, user attributes, text tags and other content information as auxiliary information to guide the user's interest in the clustering projection process. Preeti P V et al. proposed the recommendation system ARM based on association rule mining, and provided corresponding portfolio Suggestions to investors by mining association rules between stocks [2][3].

In China, the research on personalized recommendation system is relatively late, and the research on user interest modeling is still in its infancy. Existing academic research mainly focuses on the following aspects [4][5] :

Interest data collection based on user behavior. For example, Chen Shuran introduced hyperbolic model and Logistic model to analyze user browsing time, drag and roll bar times to collect user interest data. Yin Chunhui et al. used Web Monitor to monitor the user browser to collect user interest data. Yang Jiping can obtain users' subscription, refresh, browse, play and download behaviors to collect user interest data by mining the user behavior log.

Representation of user interest model: For example, Yang Jing proposed to use the modeling method of vector space model to describe user' interest with personalized tags, which reduced the dimension of user interest model, avoided the complex process of analyzing tag semantics, and could accurately represent user interest model from the perspective of user. Kang Haiyan proposed an anonymous representation method of user interest model based on the combination of differential privacy and p-link technology, and solved the contradiction between the potential user privacy security and the improvement of the performance of personalized information retrieval in the personalized retrieval technology in big data.

Learning and updating of user interest model. For example, Ma Haibing proposed to combine explicit update and implicit update to update user interest model. Song Yanjuan proposed the TFIDF algorithm based on document structure and web page interest weight, and gave the update algorithm of user interest model. The user interest model based on the algorithm can improve the user interest capture. Gong Lingling proposed to build a user interest model of weighted keywords based on domain ontology. The model can accurately describe the dynamic change process of user interest and distinguish between long term and short term interest.

Although great achievements have been made in the research of personalized recommendation system at home and abroad, it still needs to be further studied and explored to consider and practice the identification of user interest points of personalized recommendation system from the perspective of in-depth mining.

Therefore, in this paper, we use the thought of deep learning to build multi-hidden layer artificial neural network machine learning model, for layered abstract user characteristics, deeply excavating feature information that can reflect the characteristics of the user interest points, identifying potential users interests, creating the user interest model, improving the user interest identification accuracy and generalization ability, and improving the quality of the recommendation of personalized recommendation system.

At the same time, combined with cloud computing technology, through data, model or CPU cluster parallelization, the training time of deep learning is reduced, the recognition speed of user interest points is improved, and the real-time requirement of personalized recommendation system is satisfied.

III. RESEARCH ON PERSONALIZED E-COMMERCE RECOMMENDATION PLATFORM BASED ON DEEP LEARNING

A. Research Content

This paper mainly studies the identification and analysis of user interest points in the personalized E-Commerce recommendation platform based on deep learning. Research contents include: user information collection, deep neural network construction method and parallel computing method. The main technologies involved include: data preprocessing technology, user interest point depth mining technology, parallel computing technology and visualization technology.

User information preprocessing: It mainly includes word segmentation and feature extraction.

- First, word segmentation: user information is processed by word segmentation algorithm to obtain a vector model representing the user's initial characteristics.

- Second, feature point extraction: feature point extraction is to construct an evaluation function to evaluate the information feature of each user. The feature values are queued on the score level, and the user features with the highest predetermined number of scores are selected. The vector of user feature information obtained by word segmentation algorithm is usually too high in dimension, which brings great influence on later mining analysis. Therefore, it is necessary to combine feature extraction algorithm to reduce dimension, reduce machine learning time and improve the mining speed of the system.

Deep analysis of user interest points: The deep neural network method based on deep learning is used to mine user demand, identify user interest points and create user model.

At present, BP artificial neural network belongs to the shallow machine learning method. With the help of a large number of training samples to learn statistical laws, it analyzes and mines the attribute features of unknown objects, so as to predict the unknown events. The theoretical analysis of this shallow machine learning method is very difficult, and the training method also needs a lot of experience and skills.

Professor Hinton proposed the deep learning method: by constructing deep neural network learning model and a huge number of training samples to learn feature information that can describe the nature of user characteristics.

Compared with shallow learning, deep learning method can map the characteristics of the samples in the original space transform into a new feature space by layered characteristics of transformation, so that the classification or predict is becoming easier [6].

Parallelization calculation: Use cloud platform to realize parallel computing. At present, cloud computing technology, based on the Internet, integrates the characteristics of distributed processing and parallel processing, and can provide massive information storage and super computing capacity. The distributed parallel computing environment based on cloud computing is built on the cheap PC, with low cost, which greatly reduces the cost of data processing.

In addition, under the condition of parallelization, PC nodes can be easily added and the system has strong fault tolerance. The deep mining and identification of user interest points can realize the parallelization of identification process by means of cloud platform, accelerate the mining speed of user interest points, and solve the real-time requirement of personalized recommendation system.

Visualization: Adopting computer graphics and image processing technology to convert the recommended results into graphics or images, this will be displayed on the screen to the user, so as to realize human-computer interactive process.

B. Study Scheme and Feasibility Analysis

The research scheme and feasibility analysis adopted in this paper are as follows:

Automatic word segmentation is the first step of text preprocessing and the foundation of Chinese information processing. This paper adopts ICTCLAS, a Chinese word segmentation system developed by Chinese academy of sciences, for Chinese word segmentation [7]. The accuracy rate of the system is as high as 97.58%.

Feature selection is a preprocessing step in the deep mining model of user interest points. The goal of feature selection is to achieve the best recognition effect and improve the relevance of features with the least features. In this paper, LDA model is adopted to reduce the dimension of user information feature vectors. This method can effectively use the hidden semantic correlation information in user description information for feature extraction, effectively reduce the dimension of user information feature vectors and improve the accuracy of user feature representation.

The in-depth mining model of user interest points is one of the core modules of the recommendation system. In this paper, Deep Neural Networks (DNN) model is used to realize user interest point mining analysis. DNN learns user information by building an artificial neural network with multiple hidden layers; it belongs to deep machine learning. Compared with the traditional shallow neural network, DNN has a superior feature learning ability, and the acquired features have a more essential description of user attributes and categories, which is conducive to visualization or classification.

Parallelizing deep mining process: DNN needs to learn massive user information data to train a high-quality user point recognition model. The parallelization computing technology can effectively reduce the tedious and time-consuming process of training and meet the real-time requirements of the recommendation system. In this paper, we use cloud computing platform to realize parallelization of deep mining of user interest points, so as to improve modeling speed and shorten training time.

Construction of recommendation system architecture: this paper intends to use the combination of offline mining and online recommendation. The offline part mainly preprocesses the user information and deeply digs the hidden needs of users. The online part mainly provides users with effective personalized recommendation services. The separation of offline mining and online recommendation can better adapt to

the mining analysis and recommendation of massive data, and improve the scalability and real-time response speed of personalized recommendation services.

Technical route: according to the above research content, project and feasibility analysis, the technical route adopted in this paper is shown in figure 1.
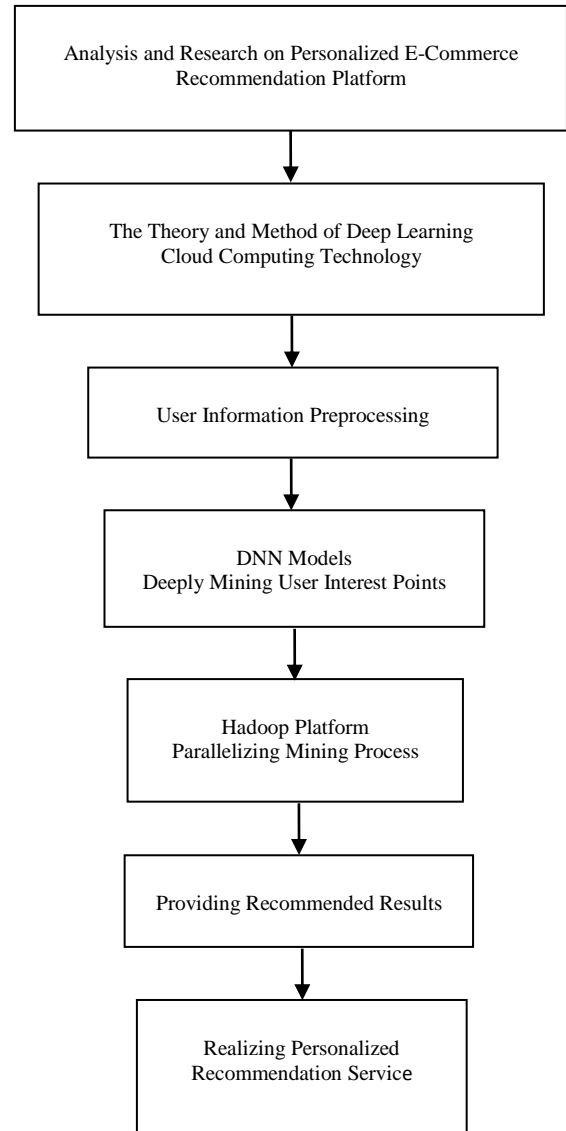

Fig. 1. Technical route.

Experimental method: Using Usenet Newsgroups standard database to do test experiments. After achieving the expected results, we will extend the study scheme to actual E-Commerce websites: designing personalized recommendation system architecture, building a deep neural network model and designing parallelization scheme.

IV. CONCLUSION

Under the environment of cloud computing technology and big data, constructing the intelligent E-Commerce recommendation system that can meet users' personalized needs and provide customers with personalized information services and decision support is becoming a hot application field of online transactions at present.

This paper first analyzes the social significance, economic value and the problems and shortcomings of the existing recommendation system. Then, it puts forward the idea and method of creating personalized E-Commerce recommendation system based on deep learning and cloud computing technology. The proposed method applies deep learning to personalized E-Commerce recommendation system. By building a deep neural network model, user information features can be abstracted layer by layer, user interest point features are mined in depth, and user interest point recognition accuracy is improved, thus improving the recommendation quality of personalized recommendation system. At the same time, parallel computing technology is adopted to accelerate the identification process of user interest points, improve the creation speed of user interest model, and meet the real-time requirements of personalized E-Commerce recommendation system.

ACKNOWLEDGMENT

REFERENCES

[1] Chen Qingzhang, Tang Zhongze, Wang Kai. Study on Automatic Recommendation Technology of Data Mining [J]. Journal of Chinese Information Processing, 2012, vol. 26(4), pp.115-121. (In Chinese)

[2] Kazuki Y, Naruaki T, Yuhei A, et al. An Introduction of A Tag Ratio Model and the Classification Examination for Recommender Systems [J]. Procedia Computer Science, 2013, Vol.22, pp. 606-614.

[3] Mustansar A G, Adam P B. Leveraging Clustering Approaches to Solve the Gray-Sheep Users Problem in Recommender Systems [J]. Expert Systems With Applications, 2014, Vol.41(7), pp. 3261–3275.

[4] Rajaraman A, Ullman J D. Mining of massive datasets [M]. Cambridge University Press, 2012.

[5] Sun Yusheng, Liu Wei, Qiu Rongrong. Research Progress of User Interest Modeling in China [J]. Journal of Intelligence, 2013, Vol. 32(5), pp. 145-149+165.

[6] Dun Yuqing, Chen Li, Liu Jing, Chen Qiang. BP Neural Network Feature Selection Method Based on Sensitivity Analysis [C]. International Conference on Knowledge Discovery and Data Mining, 2009.

[7] Li Qiong, Chen Li. Study on Multi-Class Text Classification Based on Improved SVM [C]. ISKE2013, Press: Springer-Verlag, 2014.