A Novel Intelligent Recommendation Algorithm based on Web Data Mining Technique under the Background of Deep Neural Network

1,2 Changchun Yang, 2 Jun Wang, 2 Min Yuan, 2 Chenyang Lei
1 College of information and mathematical, Changzhou University, Jiangsu 213164, China
2 School of business, Changzhou University, Jiangsu 213164, China

Abstract

The development of the Internet brought us into an era of big data information, give people bring convenient while and also make people ragged when choosing the required information and recommendation system arises at the historic moment, and get the wide attention and applications. Therefore, to enhance the traditional method, we propose a novel intelligent recommendation algorithm based on Web data mining technique under the background of the deep neural network. Firstly, we review the state-of-the-art web data mining algorithms and revise the traditional ones with the parallel data mining algorithm on the multiple processors to perform tasks that will enhance the accuracy and efficiency. Then, we analyze basic neural network model through the inner connection and weight transfer. Later, we introduce the deep network structure to enhance the traditional network. Finally, we combine the revised prior theories into the recommending tasks for enhancement. The experimental analysis show that our algorithm accuracy is enhanced by the extent of 56% and overall time is reduced to the 87% of traditional ones which proves the feasibility. Later, more optimization work will be introduced to modify the current methodology.

Keywords: Recommendation System, Web Data Mining, Deep Neural Network, Topology Optimization, Particle Swarm Optimization.

1. Introduction

In recent years, how to improve the effectiveness of the commercial web sites especially how to use the personalized recommendation technology to realize the electronic commerce personalized service has gradually become a hot topic can cause widespread interest. But at present domestic most of e-commerce recommendation is usually: recommended best-selling products. Recommend related product according to user's browsing history is recommended so to speak, the first two recommended due to fundamental not considering the personality traits of the different users, therefore recommend simply does not have the characteristics of individuation, the third recommend a personalized composition, but most of the site also only stay in only the users against a person's purchase history, just for each user set up a personal purchase records, no transverse to the comprehensive information, so there is no collaboration recommended value which also is unable to realizes real-time comprehensive recommended goods. Personalized technology is the most important of the big data era [1-2].

Recommendation system as a kind of personalized information service system is made up of the information retrieval and information filtering technology development, in the research achievements of the information retrieval and the core information filtering technology on the recommendation system. But the recommendation system in the service
environment and the application form and the two are different [3]. According to the literature review, the modern optimization for the traditional recommendation systems could be summarized as the follows. (1) To strengthen the user control. Most existing recommendation system according to preset automatically generates multiple recommended users’ personal information and requirements, to some extent, limits the user participation and control. The result of the recommendation system should allow users to participate in parameter definition. Recommendation system can be achieved by the relevance feedback mechanism to update the user's real-time demand, for example, by the user to explicitly to the evaluation of the recommended collect user feedback information. (2) To strengthen the interactive interface design. System interface is one of the most important factors affecting the user satisfaction. Existing recommendation system most committed to the improvement of recommendation algorithm, paid little attention to interface problem. Using the multidimensional information visualization technology can make result of the recommended straightforward graphic interpretation can help users understand the causes of recommended. (3) Recommend model support package. Recommendation model based on package, we put forward concept of incremental recommendation system, recommendation process is broken down into several successive steps, system recommended was generated according to the user requirements in each step, by the user decides to join the final plan of object, the user affect the choice of the system in the subsequent steps of recommendations.

![Diagram](image)

**Figure 1. The Illustration of the Modern Recommendation System**

As demonstrated in figure one, we illustrate the modern recommendation system structure. To modify the traditional system, we will combine the Web data mining technique and deep neural network tool for enhancement. With the rapid development of database technology and the wide application of the database management system, behind the huge increase of data is rich in knowledge and though the current database technology can be quickly and efficiently implement the data query, statistics, and other functions, but can’t find huge amounts of data that exist in the relationship, not according to the existing data to predict the future trend of the development. Theoretically, the data mining technique could be categorized as the follows parts. (1) Automatic forecast trends and behavior. Automatic data mining in large databases to find predictive information automatically is proposed to describe the important data model or the data to predict the
future trend of class classify and forecast. (2) Deviation detection. Deviation include many potential knowledge, such as a special case of the abnormal instance, does not meet the rules of classification, the observations and the model prediction deviation, magnitude of change over time, etc. (3) Clustering. Clustering is according to the certain rules could be divided into reasonable data collection, the objects are grouped into multiple classes, or clusters, in the same cluster have high similarity between objects. (4) Correlation analysis. If it is not by using data mining association rules in data mining technology for large trading data mining analysis, data association is a kind of important that exist in the database.

In this paper to enhance overall performance of the traditional recommendation system, we conduct research on the novel intelligent recommendation algorithm based on the Web data mining technique under the background of deep neural network. The reminding of the article is organized as the follows. In the section 2, we review state-of-the-art algorithms proposed in the recent time and analyze the characteristics of the methods. In the section 3, we introduce the primary theory and analysis on the web data mining algorithms. In the section 4, we use the deep structure to optimize the traditional neural network to enhance the robustness of the artificial network. In the section 5, we discuss basic concepts of intelligent recommendation algorithm and propose our revised optimized novel algorithm. In the section 6, we simulate the proposed algorithm and in the section 7, we conclude the research.

2. Review of the Related Work

Recently, a large number of intelligent recommendation systems are proposed. In the [4], Yu proposed the dynamic competitive recommendation algorithm in social network services. They designs and realizes a kind of the intelligent, personalized multimedia recommendation system. System can be used according to the user's habits, use time, using the environment, and recently selected projects were analyzed, and the optimal recommended resources listed after judgment. System through the study of self-learning personal information technology with personalized characteristic analysis technology and content search technology, this technology application in the recommendation system achieve the better interactive features. In [5], Chen’s group conduct research on collaborative filtering recommendation algorithm based on web user clustering. Based on variable precision method to design a collaborative filtering algorithm as first, select the film evaluation data sets, variable precision method are used to get the new data set through the collaborative recommendation gives recommendation results; Second, through the average absolute error index made an evaluation of the accuracy of the result of the recommendation, found that with the improvement of variable precision, recommended were improved, but accuracy needs to be controlled within certain threshold, otherwise more than threshold recommended performance declines.

In [6], Liu’s group conducts research on the immune clustering-based recommendation algorithm. The key technology of this study was to identify organization form of Web pages, which need personalized information page. The personalized information recommendation system based on Web data mining can meet the needs of the future development trend of the Internet. With traditional compared to the page Web information extraction based on the page structure partition information recommendation conforms to the actual situation of particle size. In [7], Xia’s group conducted research on the matching recommendation algorithm for celebrity endorsement on social network. Through the adoption of multi-agent technology to cross system of personalized search system is designed adopting the initial interactive agent initialization single users of the system model, using the dynamic tracking users interested in the agent and historical records retrieval agent to update the user model, and use of the model cross system user model. More models could be found in the [8-15].
3. The Web Data Mining Algorithms

3.1. The General Data Mining Algorithms and the Proposed Methodology

Data mining technology can be divided into categories according to its working process of the data of extraction, data storage and management, key technologies such as data show. (1) The decision tree. Decision tree, also known as decision tree is a kind of similar binary tree or the tree structure which is used to discrete and continuous attributes in the predictive modeling. Decision tree is to use the sample properties as node, with the attribute value as the branch, which is similar to the flow chart of the process, where each internal node said test on an attribute, each branch represents a test output, and each leaf nodes represent the classes or class distributions. (2) Genetic algorithm. Genetic algorithm is simulated biological evolution process calculation model that is a natural genetics and computer combined with each other, mutual penetration and forms the new calculation method. Genetic algorithm has the implicit parallelism, easy and other model combined with the nature and it set foot in the field of data mining. (3) Clustering algorithm. Clustering analysis algorithm is to measure the similarity between the individual which is based on individual data points in judging geometric space distance, the closer distance, the more similar and the easier it is. After the original definition classification, the algorithm will represent points through the calculation and determine the classification of grouping of appropriateness, then try to redefine these groups to create can better represent data classification. The algorithm will cycle through the process, until it can no longer by redefining classification to improve the results so far. In simple terms, clustering is a collection of data objects grouped as the process of multiple classes of similar objects. (4) Neural network algorithm. Neural network algorithm using three layer neurons composed of a multi-layer perceptron network, a network is also called back propagation law of the Delta. These layers are: input layer, optional hidden layer and output layer, in multilayer perceptron network, each neuron accept one or more input to produce one or more of the same output.

![Figure 2. The Standard Procedures of the Data Mining Algorithm](image)

As demonstrated in the figure 2, we show standard procedures of data mining algorithm. After reviewing the related literatures, we plan to adopt the chain distance estimation prior to serve as the fundamental theory of our later discussion. The fundamental goal of data mining is to extract data flow or the main characteristics of time series as the data flow sequence is the main source of characteristic data. The formula one defines the feature vectors.
\[ w(A, B) = \frac{|A \cap B|}{|A| || B|^{e_i}} \]

(1)

Where the \( e_i \) represents the characteristics of the data matching weight vector analysis, the aggregation on the characteristic data transfer and data mining is known by the data feature extraction for the whole process of data transmission is shown as the follows.

\[ M = M_{DB} + M_{Dup} = 2M_{DB} + \sum_{i=1}^{n} \sum_{j=1}^{i-1} Jf(I_{i,j}) \]

(2)

On the basis of the above chain distance estimating pretreatment, in view of the traditional method of characteristic data mining algorithm is used to find the data points in the neighbor area method, when the neighbor area appears from the group of factors, the problem of data mining result is bad for the data mining algorithm is improved.

3.2. The Web Data Mining

Meet the requirements of the user processing diversified information of data warehouses is different from the general database. For general database, if only the query function can also be, however, faced with more and the newer requirements of users, to show its limitations has this aspect the superiority of the data warehouse. Only data warehouse and data mining that combine a new research field, from large amounts of complex data to the user truly valuable information and knowledge. Among various kinds of data warehouse, people's understanding of a lot of knowledge is based on the fuzzy concept. For data of some of the properties we can use to abstract the concept of fuzzy, so that users in information analysis, excavation can be presented in the form of natural language. Recursively to the training set, through calculating the information gain ratio records of the training set, dealing with the unknown attribute value until each subset of records all belong to the category or one kind of overwhelming majority.

Using the parallel data mining algorithm, on the multiple processors to perform tasks at the same time, can improve the system operation speed, enhances the efficiency. Because of the different parallel data mining algorithm, should be used according to actual circumstances, suitable mining method as the final formation of the decision to play a more active role. In the figure 3, we show the procedures of the web data mining tasks.

![Figure 3. The Flowchart and Procedures of the Web Data Mining Tasks](image_url)

Parallel processing method of the downside is that it requires a lot of parallel processors or processor, the hardware implementation is difficult. At the same time, we are in for a parallel method or synthesis method are effective cannot be discussed in
isolation from the data more can't use a lot of pure pursuit of the processor. Because both
the data load on each processor or the results will be comprehensive or output is a
complex and arduous task, they need to consume a lot of time and time consuming of
input/output although initially increase and the decrease in the number of processors, but
the final decision tree to establish the core time of consumption is not a monotonic decline
curve of ideal, but rather a rise after falling first not smooth curve, and the increase in the
number of processors needed to deal with intermediate variable space also increased.
Therefore, to enhance the visibility of the web mining result, we use the subordinating
degree function to reflect the property as the formula 3.

\[
\text{measurement} = \frac{\left(\sum_{i=1}^{m} \nu_i(x_i)\right) \times x_i}{\left(\sum_{i=1}^{m} \nu_i(x_i)\right)}
\]

(3)

4. The Enhanced Neural Network Model

4.1. The Traditional Neural Network

Artificial neural network is a kind of imitating the human brain structure, method
system to work with the distributed parallel processing, nonlinear mapping adaptive
learning and robust features, has been widely used in pattern recognition. It is made of a
large number of wide connections of neurons with nonlinear mapping capability of the
nonlinear dynamic system. In the figure four, we demonstrate the topology and
transmission weight of neural network.

![Figure 4. The Topology and Transmission Weight of the Neural Network](image)

Usually the error between actual output and target output, using gradient descent
method to adjust the weights between nerve cells, reduce the error falling in order to
achieve the desired accuracy, complete the learning process, the weight is obtained by the
results of the study. If the initial weights, error precision, vector and choose appropriate,
it will be limited to time, complete the weights of learning [18-21]. The formula 4 defines
the propagation function.

\[
e = \sum_{j=1}^{N} w_j x_j - \theta, \quad y = f(e)
\]

(4)
Where, the $\theta$ denotes the neuron threshold, $\chi$ is the input signal. In essence, this process is an iterative process, an iterative algorithm is generally closely related with the choice of the initial value, such as initial value selection of good, fast convergence rate, such as initial value selection is not good, slow convergence speed and convergence. When the cumulative value use the algorithm of gradient descent, inevitably fall into the local minimum problems. The optimal condition is to meet with the following situation.

$$(\text{min}) E = \frac{1}{2} \sum_{p} (y_p - \hat{y})^2$$

(5)

Neural network training is actually an unconstrained nonlinear optimization problems, for ANN training such a complex problem, there is no good in optimization theory solution. But if use some existing effective method in nonlinear optimization and modify it according to the characteristics of the feedforward neural network as is expected to get some issue of complex algorithm is effective and superior to BP algorithm training. Here we consider a second order optimization method of the application. Combining the genetic algorithm and neural network in optimization and advantages of nonlinear data processing to apply the encapsulation mode of feature selection, evaluation tools by the BP neural network classifier as the data feature extraction and feature dimension reduction and to retain the original classification data model. For a practical problem for the optimization of genetic algorithm, first of all we need to code, all parameters of problem or are encoded data characteristics, the formation of a finite long string which is usually a binary string as the individuals. In the formula 6~8, we show the revised gradient descent algorithm for the neural network training.

$$k(s+1) = k(s) + Z \left( -\frac{\partial E}{\partial k} \right)$$

(6)

$$\frac{\partial e}{\partial k_i} = \sum_{p=1}^{\tau_i} y_i^p \left( 1 - y_i^p \right) \left( y_i^p - t_i^p \right) h_i^p$$

(7)

$$\frac{\partial e}{\partial k_j} = \sum_{p=1}^{\tau} \left[ h_j^p \left( 1 - h_j^p \right) x_j^p \sum_{q=1}^{\tau} \left( y_q^p - t_q^p \right) \left( 1 - y_q^p \right) k_j^p \right]$$

(8)

It can be solved by using novel initialization by liberty and rights and makes the complex nonlinear mapping model is converted into a set of linear equations to solve the problem, which greatly accelerate and simplify calculation process and time complexity is polynomial.

4.2. Deep Neural Network

Deep learning is the deep structural model is established for the target of machine learning, general contract model contains at least three hidden layer. The network of the hidden layer structure with normal neural network training algorithm, BP algorithm is hard to work, for example, not only because the volume of the sample data, training process is slow, and the parameters are easy to convergence on the local rather than global optimal point, the practical significance is not strong. Characteristics as raw materials of machine learning system, there is no doubt to the influence that the final model. Machine learning algorithms depends largely on the performance of data express or express the choice of features, when the data can be expressed as the well characteristics and even using a simple model can achieve satisfactory precision. In the actual application of
machine learning algorithm, an important step is how to pretreatment the data to get a good characteristics. Deep learning approach has made very good result, and is widely used, but it's actually a bit like a "black box", it is not very full and strict theoretical system to support. So there is a problem: use the deep learning method has achieved good results, but in theory, do not know why. Direction of the study in depth, people more and more depth through the use of unsupervised learning algorithm model and more and more complex, a lot of progress, but for the structure of the deep learning method module, such as single since the encoder, single limit Boltzmann machine research is relatively small.

Figure 4. The Deep Neural Network Structure and Organization

As shown above, the general structure of deep neural network is demonstrated. In addition to the input and output the final layer visible, depth study of the other layers are hidden which contains how many or what kind of information is difficult to parse. Because of the encoder and the restricted Boltzmann machine the general train of thought of the two methods are the consistent, mainly lies in the training process is different, so in the experimental part, focus on the encoder. Determine the depth of the network model as the selection is a linear function as the activation of neurons function. Input through the role of the nonnegative linear is linear function, from lower to upper network transmission, step by step for the feature extraction. Positive linear activation function not only simple calculation and the output of zero make the data characteristics of extracted to have stable sparse representation. So, we choose average reconstruction error and a weight sum of squares to measure cost function which is denoted as the following expression [22-25].

\[
J(W,b) = \frac{1}{m} \sum_{i=1}^{m} J(W, b; x', y') + \frac{\lambda}{2} \sum_{i=1}^{d} \sum_{j=1}^{d} \sum_{l=1}^{d} (W_{ijl})^2
\]

(9)

In the formula 9, the \(J(W, b; x', y')\) represents the network output. Second term uses network weight sum of squares as the cost function, estimate depth network connection weights for consistency between the layers. To minimize the weight sum of the squares can reduce weight number and size to prevent network fitting. Express conversion refers to original expression is converted to a different expression and in this paper we consider new expression dimension is less than the original dimension. In pattern recognition system using feature is very important, often need to convert the original redundant feature high dimensional low dimensional retain the characteristics of the effective information, namely feature conversion. To training the deep neural network, we update the weight following the formula 10.
\[ W' (\text{New}) = W' - \alpha \frac{\partial J(W,b)}{\partial W'} \]

(10)

This part mainly explores the understanding of the characteristics and the encoder is a great way to express the transformation and assessed monolayer respectively from the encoder based on the original pixels and the local descriptor dimension reduction effect and the stability of the power of expression and comparing with the traditional principal component analysis. For the feature extraction, respectively uses the encoder and the principal component analysis to express the transformation and the principal component analysis as a linear feature extraction method, since the encoder as a nonlinear feature extraction method.

5. The Proposed Intelligent Recommendation Algorithm

5.1. The Basic Recommendation Algorithm

Recommendation system can be thought of as a kind of expert system based on network environment, it is able to learn from the limited information of the user's preferences and then from alternative projects make personalized recommendation to the user. Related areas such as the artificial intelligence, data mining, the development of the information resources on the network size increase sharply on network information service also gradually to humanization, personalization and the intelligent direction, recommendation system become one of the hot research topic. The existing methods for the recommendation systems could be reviewed as the follows. Recommendation system based on content, the content-based recommendation system based on the analysis of user tried and loved to recommend other similar object. The content is usually expressed as object of multiple keywords, such as the author, publication date, etc. The primary technique used is the angle cosine vector distance.

\[ u(c,s) = \text{score}(\text{ContentBased}(c),\text{Content}(s)) \]

(11)

In addition to the utility based on the predefined formula the data mining techniques can be obtained from the data model to implement content-based recommendation. Commonly used data mining techniques have Bayesian classifier, decision tree and artificial neural network. Based on the collaborative filtering recommendation system is also the popular method. The less when user rating, user similarity computing error is bigger, therefore recommends for not scoring the preset object evaluation to improve the prediction accuracy as the follows.

\[ r_{c,s} = k \sum_{c' \in C} \text{sim}(c,c') s_{c,c'} \]

(12)

Based on the model of the algorithm using existing user ratings training set a model, and then evaluated using the method of probability and statistics score predicts. In order to make full use of the advantage of different recommendation methods, improve the recommendation accuracy and efficiency, the researchers of content-based and collaborative filtering method to integrate, proposes a hybrid recommendation system. Recommendation algorithm based on ontology is recommended ontology is introduced into system, the resources with the method of ontology described, the structured said at same time with semantic information resources. Using collaborative filtering recommendation model and based on the basic thought, content recommendation model user clustering information by using the user rating matrix instead of the user preference
information as the similarity of user preference information and project information using the properties of the similarity calculation method. Hybrid recommendation algorithm refers to the basic recommend using more than one method recommended by the joint. Due to the collaborative filtering and content-based recommendation algorithm have their respective advantages and disadvantages, so the hybrid recommendation algorithm can be used in a variety of advantages of recommendation algorithm program is recommended. At the same time, because of hybrid recommendation algorithm uses variety of recommendations to recommend at the same time, efficiency is not high as weight of various recommendation algorithms is transferred into the optimization problem.

![Figure 5. The User Classification and Division](image)

**5.2. The Modified Intelligent Recommendation System**

Data quantity is more and more big, gradually exposed his shortcomings in the traditional methods: one is data about data sparseness, namely the site of great quantity and customer purchase or the comment is only a small part of the opposition, sparse data can reduce the accuracy of recommendation algorithm. The one solution is before use specific collaborative filtering algorithm using dimensional reduction technology of the raw data compression, to reduce the dimension of make information loss reduced to the minimum at the same time. Similarity comparison method based on vector strictly match the attributes of the object and namely the user overall score as a single score vector. As shown in figure 6, the Web based data pool and regulation for mining and recommend is illustrated.

```xml
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:foaf="http://xmlns.com/foaf/0.1/
  xmlns:foafrate="http://www.foaf.rate.com/foaf/"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">
  <foaf:Person>
    <foaf:name>Blog A</foaf:name>
    ...
    <foaf:weblog rdf:resource="...com/BlogA/"/>
    <foaf:knows>
      <foaf:Person>
        <foaf:name>Blog B</foaf:name>
        ...
      </foaf:Person>
      <foaf:Person>
        <foaf:name>Blog C</foaf:name>
        ...
      </foaf:Person>
    </foaf:knows>
  </foaf:Person>
</rdf:RDF>
```

**Figure 6. The Web based Data Pool and Regulation for Mining and Recommend**

In the standard particle swarm optimization algorithm, the particle's position and speed to its running track, on the other hand, due to the limited speed and fixed trajectory, the search space is limited which lead to global search ability is poor. Therefore, we joint...
deep neural network and the corresponding particle optimization algorithm to enhance the performance. We revise the traditional particles update equation into the formula 13.

\[
X_{id}(t+1) = p_{id}(t) + \frac{E}{2} \ln \left( \frac{1}{\mu} \right)
\]

(13)

Although quantum particle swarm optimization algorithm based on Delta potential well global search ability is much better than the general particle swarm optimization algorithm, but because the particles in the search process according to learn history of its own optimal location to implement the search, did not consider whether there is the risk of fall into the local optimum, in this process, will inevitably reduce the population diversity, which appears premature phenomenon. Quantum particle swarm optimization algorithm based on Cauchy distribution. Cauchy distribution the biggest feature is its flanks probability characteristics, which can effectively expanding global searching of particle, making it bigger than Gaussian distribution generated the random number distribution area, thus, Cauchy distribution has a greater probability of quick jump out of local optimal point. In view of the above problem, we consider the individual to update particle location at the same time, the optimal position and the global optimal position which has been discussed in the previous sections. In the figure 7, we show the proposed intelligent recommendation system with the prior discussion.

Figure 7. The Demonstration of the Proposed Intelligent Recommendation System

6. Experiment and Simulation

In this section, we conduct numerical simulation on proposed algorithm. Recommendation system in the large data environment due to its huge numbers of new users and characteristics of the large number of the product and in data storage and query will produce a lot of data, if the data is too much can lead to network congestion, unable to process the data quickly and sometimes may result in data loss and this makes the algorithm of data analysis are accurate enough. Line storage is the most commonly used currently recommended system data storage structure, is according to the way to store data, in the recommendation system based on user or products in time in order to they are lined up. Under these background conditions, in the figure 8, we test the efficiency of our
proposed algorithm. The result shows that our method obtains the better efficiency compared with the traditional ones. In the figure 9, we illustrate the test for robustness and the accuracy. The experimental result indicates that compared with other 16 set of the algorithms, our method performs the best. The recommendation accuracy is enhanced by the extent of 56% and overall time is reduced to the 87% of traditional ones.

Figure 8. The Performance of the Recommendation Efficiency

Figure 9. The Performance of the Recommendation Accuracy and Robustness
7. Final Conclusion

Recommendation system as a kind of personalized information service system is made up of the information retrieval and information filtering technology development. In this paper, we conduct research on the novel intelligent recommendation algorithm based on Web data mining technique under the background of deep neural network. Our research is finalized by the three levels. (1) We review the state-of-the-art Web data mining algorithms and revise the selected one for later processing. (2) We adopt the deep neural network model to enhance the performance of the traditional algorithms. (3) We propose the novel algorithm combined with the data mining and neural network theory which is novel and innovative. The experimental result proves the effectiveness and robustness of the proposed approach. The recommendation accuracy is enhanced by the extent of 56% and overall time is reduced to the 87% of traditional ones. In the future, we will focus on the network optimization part to enhance the overall recommendation algorithm.

References


