

Research Article

Spring Festival Holiday Tourism Data Mining Based on the Deep Learning Model

Ruijun Wang 

School of Business Administration, Shandong Management University, Jinan 250357, Shandong, China

Correspondence should be addressed to Ruijun Wang; yushu2021@126.com

Received 26 March 2022; Revised 11 May 2022; Accepted 21 May 2022; Published 18 June 2022

Academic Editor: Ahmed Farouk

Copyright © 2022 Ruijun Wang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With the advent of the era of big data, people have entered a situation of information overload, how do users filter out the information they need from a large amount of information. When users browse the website, they will record their search or click behavior, and the recommendation system will mine the data based on these data, and recommend the information they need for each user. With the birth of the recommender system, it has indeed changed the way people obtain information. Instead of relying solely on search engines to obtain information, it can obtain the information they want without people's "consciousness." This shift has made it easier for people to access information. This paper conducts research on travel recommendation during the Spring Festival holiday. The paper introduces deep learning model and data mining technology, proposes that the recommendation system has three important modules, and obtains the corresponding flowchart. The recommendation system was optimized, and a comparison chart of coverage before and after optimization was obtained. Before optimization, the coverage rate of cities and scenic spots was 45.52% and 21.25%, respectively, and reached 55.65% and 49.81% after optimization.

1. Introduction

Nowadays, people have entered an era of big data explosion. The Internet has long been inseparable from lives and has become a part of lives. People can watch movies and TV series on websites such as Youku, listen to songs on music platforms such as QQ Music and NetEase Cloud Music, and shop on shopping platforms such as Taobao and JD.com. How people filter out the information that is beneficial to themselves from a lot of information, it really costs people a lot of energy and time. How will the filtered information be presented in front of people's eyes, and whether this part of the screening work can be converted into computer processing by people themselves.

The birth of the Internet has indeed brought convenience to people's lives, making people's lives more and more colorful. But with the advent of the era of big data, the Internet has also brought unavoidable problems to people-information overload, making people dazzled and unable to choose. In most cases, people are not clear about their goals.

For example, the user wants to travel to a certain place, but the attractions of this place may not be known to him, but the truth is what he wants to visit. At this time, the search engine cannot be displayed, and the results obtained through the search engine are not personalized and are immutable and single. As long as the same keywords are entered, the results presented are the same. With the birth of the recommendation system, these problems are solved, and compared with the search engine, it has more personalized functions.

The innovation of this paper is that the deep learning model is introduced, an improved deep learning framework is constructed, and a schematic diagram of the algorithm flow of the framework is obtained. Data mining technology is introduced, and the remaining parameters are not further compressed by any subsequent steps, and the test accuracy and the results on the dataset are obtained. The tourism recommendation system proposed in the article is optimized, and the comparison data before and after optimization are obtained.

2. Related Work

Regarding deep learning, relevant scientists have done the following research. Litjens et al. outline key medical imaging research concepts and has collected more than 300 papers on the topic, most of which were published last year. He evaluates in-depth research practices in image classification, object analysis, segmentation, recording, and other activities and provides a brief overview of research in each application area, discussing open challenges and directions for future research [1]. Kermary et al. have developed a deep learning framework based on diagnostic tools for screening patients with simple treatments for blinding retinal diseases. The framework uses transfer learning, which uses a small subset of data from traditional methods to train a neural system, using this method to check database visual consistency. The latter tool can help expedite the diagnosis and diagnosis of treatable diseases, leading to earlier treatment and better clinical outcomes [2]. He et al. use a denoising-based predictive communication channel that can learn from the channel structure and evaluate the network against a large amount of training data. He also provides an analytical framework for assessing asymptomatic functioning of the network. According to He et al.'s analysis and simulation results, deep learning is a powerful tool for channel estimation in millimeter wave communications [3]. Weinan and Yu proposed a deep learning-based method for numerically solving variational problems, especially those arising from partial differential equations. The method is naturally nonlinear, naturally adaptive, and has the potential to work in fairly high dimensions. The framework is very simple and well suited for the stochastic gradient descent method used in deep learning. Weinan and Yu illustrate the method on several problems, including some eigenvalue problems [4]. Tom et al. carefully examine key models and techniques used in many NLP projects and describe their development. They also summarize, compare, and create different categories and provide a complete understanding of the past, present, and future of deep learning NLP research [5]. Zhu et al. analyze the challenges of using deep learning for remote sensing data analysis, review recent advances, and provide resources that hopefully make deep learning in remote sensing seem ridiculously easy. Zhu et al. encourage remote sensing scientists to bring their expertise to deep learning as an implicit universal model to address unprecedented, large-scale, and impactful challenges such as climate change and urbanization [6]. Xu proposed a learning-based DOA estimation method for multiple broadband far-field sources. The processing mainly includes two steps. First, a beam-spatial preprocessing structure with frequency-invariant properties is applied to the array output to perform focusing over a wide bandwidth. In the second step, the classification is implemented using a hierarchical deep neural network. Unlike neural networks trained on huge datasets containing combinations of different angles, deep neural networks can achieve multisource DOA estimation with small datasets, since the classifier can be trained in different small subregions. Simulation results show that the method performs well in generalization and defect adaptation [5]. These

methods provide some references for research, but due to the short time and small sample size of the relevant research, this research has not been recognized by the public.

3. Spring Festival Holiday Tourism Data Mining Method

3.1. Deep Learning. Deep learning is the study of the principles and layers of presentation of data samples, and the information gathered during these learning processes is useful for interpreting data such as text, images, and sounds [7]. Its main purpose is to observe and learn from machines, as well as recognize people and data such as words, images, and sounds. Deep learning is a complex mechanical algorithm that analyzes speech and artwork without similar techniques [8].

Simply put, deep learning is a technique that enables computer systems to improve from empirical data [9]. Neural networks, inspired by biomimicry research, simulate how neurons in the brain work, with axons responsible for receiving signals. It is from hundreds of billions of these neurons that make up human brain through complex connections [10]. A perceptron is a linear artificial neuron with a binary classification function:

$$\text{output} = \begin{cases} 0, & \text{if } \sum_v \lambda_v m_v \leq \text{thread} \\ 1, & \text{if } \sum_v \lambda_v m_v > \text{thread} \end{cases}, \quad (1)$$

where $\sum_v \lambda_v m_v$ —the sum of the assigned weights.

Computer vision: the Multimedia Lab of the Chinese University of Hong Kong is the first Chinese team to apply deep learning for computer vision research. In the world-class artificial intelligence competition LFW (Large-scale Face Recognition Competition), the laboratory has won the championship, making the recognition ability of artificial intelligence surpass that of real people for the first time in this field.

For speech recognition, Microsoft researchers cooperated to first introduce RBM and DBN into the training of speech recognition acoustic models and achieved great success in large-vocabulary speech recognition systems, reducing the error rate of speech recognition by 30%. However, there is no effective parallel and fast algorithm for DNN, and many research institutions are using large-scale data corpus to improve the training efficiency of DNN acoustic model through GPU platform.

Artificial neural network is now very popular and widely used. As the basis of deep learning, its characteristics include the following: (1) strong learning ability. The biggest feature of the neural network is that it has a strong ability to extract features. Similar to the brain, it can deal with the transformation of various inputs. When the input changes, it can adjust the features extracted by itself in time and has strong adaptability. (2) Parallelism. The human brain can process multiple things at the same time, which reflects the parallelism of the human brain. The neural network can process information independently by

simulating the human brain, which reflects the same similar parallelism. (3) Nonlinear. Neural network is an important nonlinear system research tool. It can effectively discover the nonlinear relationship model between input and output, and from the outside, the neural network is similar to a black box tool, which hides other parts of the neural network structure except the input and output. (4) Robustness. Since there are many neurons in the neural network, each neuron will share the contribution value, so the influence of each neuron on the overall result is relatively weak. When some data are polluted, they have little effect on the results of the entire network, especially in distributed computing, which can reflect this robustness [11]. Figure 1 shows the network classification performance statistics.

Common deep learning models include:

- (1) Autoencoder. It consists of input layer, hidden layer, and output layer. It models a network with a three-layer neural network structure: the same number of neurons in the input and output layers. Autoencoders are a bit predictive software, image-guided, text-guided. The hidden layers that describe users and objects are learned by recreating relevant information about users and objects, and then these expressions are used to predict user behavior settings.
- (2) Restricted Boltzmann Machines. The Boltzmann machine is composed of visible units and hidden units, and the corresponding variables are called visible variables and hidden layer variables, and they are all binary variables with a value of 0-1. When a neuron is inhibited, its state is 0, and when it is activated, its state changes from 0 to 1.
- (3) The deep belief network is related to the restricted Boltzmann machine, which is composed of multiple restricted Boltzmann machines. It learns layer by layer starting from the underlying RBM.
- (4) Convolutional Neural Networks. Convolutional Neural Networks are multilayer perceptual network models that can be used to process network data such as image data [12, 13]. The main difference between it and a typical multilayer perceptron is that a convolutional neural network can reduce the number of neurons in a sample by combining layer actions. Not only that, the weights of the convolutional neural network can be shared, so that the parameters can be greatly reduced, so that the complexity of the network is reduced, and the generalization ability of the network is improved at the same time. And it also has translation invariance; in addition, the most important thing is that it can directly operate on the picture, and it mainly consists of five parts [14].

The energy model establishes a functional relationship between the energy of a certain state of the system and its probability of occurrence through the energy function and realizes a measure of the probability

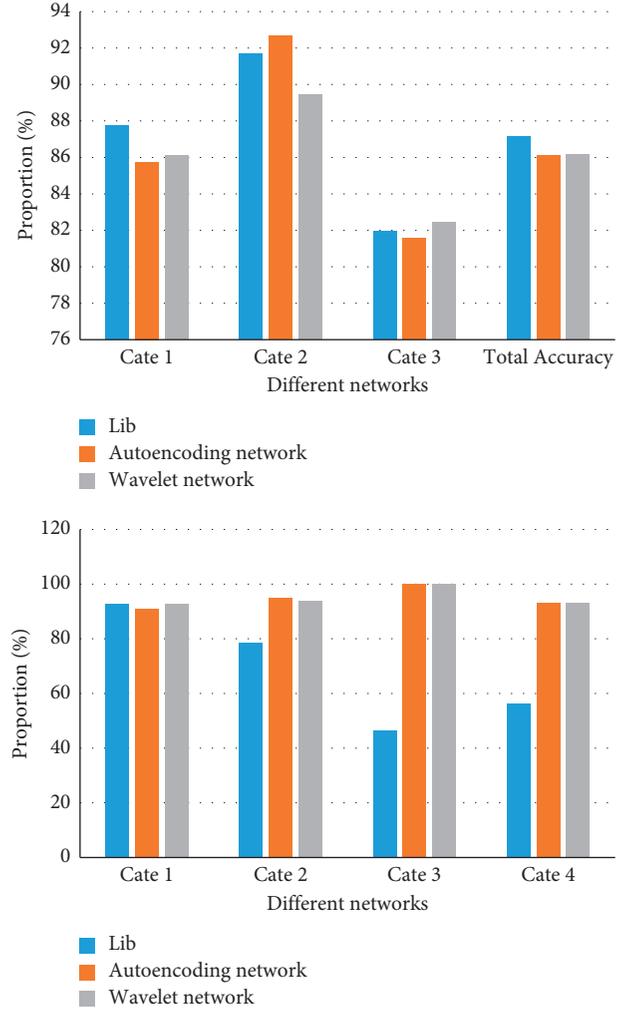


FIGURE 1: Network classification performance statistics.

distribution of a random network. Its energy function is usually recorded as

$$l(m) = \frac{e^{-R(m)}}{C}, C = \sum_m e^{-R(m)}, \quad (2)$$

where $R(m)$ —energy function.

When introducing hidden units:

$$l(m) = \sum_h l(m, h) = \sum_h \frac{e^{-R(m, h)}}{C}. \quad (3)$$

The joint distribution probability function is as follows:

$$l(m, g^1, \dots, g^p) = \left(\prod_{d=0}^{p-2} l(g^d | g^{d+1}) \right) l(g^{p-1}, g^p), \quad (4)$$

where

m —output layer.

$l(g^{1-1}, g^1)$ —joint probability of taking values.

$l(g^{k-1} | g^k)$ —probability under the condition of taking the value.

Its mathematical expression can be written as follows:

$$l(U = u|m, Q, c) = \text{soft max}(Qm + c) = \frac{e^{Q_u m + c_u}}{\sum_v e^{W_v m + c_v}}, \quad (5)$$

where

Q —weight matrix.

c —bias vector.

The loss function is a common objective function for optimizing parameters in a learning model. In the multiclass classification scenario, the negative log-likelihood function is usually used as the loss function, and the formula is as follows:

$$P(\theta = \{Q, c\}, S) = \sum_{u=0}^{|S|} \log(l(N = n^{(u)} | m^{(u)}, Q, c)), \quad (6)$$

$$p(\theta = \{Q, c\}, S) = -P(\theta = \{Q, c\}, S),$$

where

P —loss function.

l —negative log-likelihood function.

For the probability of each sentence, it is expressed by the joint probability of each word that composes the sentence, namely:

$$l(\delta) = l(\delta_1^T) = l(\delta_1, \delta_2, \dots, \delta_T), \quad (7)$$

where

$l(\delta)$ —the probability of a sentence.

δ -T sentence composed of words.

According to the relevant formula, the above formula can be decomposed into:

$$l(\delta_1^T) = l(\delta_1) * l(\delta_2 | \delta_1) * l(\delta_3 | \delta_1^2) \dots l(\delta_r | \delta_1^{T-1}). \quad (8)$$

The specific calculation method of the feature extraction method is as follows:

$$DS(r) = \frac{b_r}{B},$$

$$IDS(r) = \frac{|D|}{j: r \in d}, \quad (9)$$

$$DS - IDS = DS(r) \times IDS(r),$$

where:

b_r —the number of feature words in the current text.

B —the number of all feature words in the current text.

$j: r \in d$ —the number of feature words that the text contains.

$|D|$ —the number of all texts in the corpus.

The likelihood estimate is calculated as follows:

$$l(s|b_i) = l((r_1, r_2, \dots, r_i, \dots, r_n) | b_j) = \prod_{i=1}^n l(r_i | b), \quad (10)$$

where:

r_i —features in text features.

$l(r_i | b)$ —probability of appearing in the text.

$$l(b_j | s) = \frac{l(s|b_j)l(b_j)}{\sum_{i=1}^m l(b_i)l(s|b_j)}, \quad (11)$$

where

s —target text.

The posterior probability is as follows:

$$l(b_j | s) = l(s|b_j)l(b_j) = l(b_j) \prod_{i=1}^n l(r_i | b_j), \quad (12)$$

where

s —calculated text.

r_i —text-based features.

Convex optimization can be expressed as

$$o.p.n_v \left[\frac{\delta}{\|\delta\|} * m_u \right] + \frac{b}{\|\delta\|} \geq \mu, i = 1 \dots N, \quad (13)$$

where

μ —geometry from dataset to hyperplane.

N —the number of samples in the sample set.

δ —weight.

The optimal solution to the original problem is obtained by solving the dual problem:

$$P(a, s, \lambda) = \frac{1}{2} \|\delta\|^2 = \sum_{u=1}^N \lambda_u n_i (\delta m_i + s) + \sum_{u=1}^N \lambda_u. \quad (14)$$

The internal structure of a neuron is as follows:

$$n = f(\delta_1 m_1 + \delta_2 m_2 + \dots \delta_i m_i \dots + \delta_n m_n), \quad (15)$$

where

m_n —input to the neuron.

δ_n —weight of connected edges.

b —bias.

f —activation function.

$$q_k = f_1 \left(\sum_{u=0}^n b_{ku} m_u \right) k = 1, 2, \dots, l, \quad (16)$$

where

n —the number of neurons in the input layer.

l —the number of neurons in the hidden layer.

m —the number of neurons in the output layer.

b_{ku} —weights between the input layer and the hidden layer.

f_1 —the activation function of the hidden layer.

$$n_v = f_2 \left(\sum_{u=0}^l b_{vk} q_k \right) v = 1, 2, \dots, m, \quad (17)$$

where

f_2 —activation function of the output layer.

n_v —the output of the network.

$$T = \frac{1}{2B} \sum_{u=1}^B \sum_{v=1}^m (r_v^u - n_v^u)^2 = \frac{1}{2B} \sum_{u=1}^B T_u, \quad (18)$$

where

B —the number of data.

u —the number of samples.

n_v^u —corresponding network output.

r_v^u —corresponding target output.

T —average error.

Different from traditional shallow learning, the difference of deep learning is as follows:

- (1) The depth of the model structure is emphasized, usually with 5, 6, or even 10 layers of hidden layer nodes.
- (2) The importance of feature learning is clarified. That is to say, through layer-by-layer feature transformation, the feature representation of the sample in the original space is transformed into a new feature space, thereby making classification or prediction easier. Compared with the method of constructing features by artificial rules, using big data to learn features can better describe the rich intrinsic information of data.

By designing and establishing an appropriate amount of neuron computing nodes and multilayer operation hierarchy, selecting appropriate input layer and output layer, through network learning and tuning, the functional relationship from input to output is established. Although the functional relationship between input and output cannot be found 100%, it can approximate the actual relationship as much as possible. Using a successfully trained network model, automation requirements for complex transaction processing can be achieved.

As a new method of machine learning in the development of artificial intelligence technology, deep learning is very important [15]. Based on the power of parallel computing and big data cloud algorithms, we will build a learning network that is closer to the human brain, allowing computers to find ways to solve “abstract ideas” and make computers more intelligent [16]. A deep online learning algorithm is an uncontrollable learning feature. It builds on the analytic hierarchy attributes, by combining and analyzing the attributes behind the buildings and replacing the original top-level abstract attributes with more abstract attributes to create top-level abstract attributes that show work characteristics. Deep networks try to imitate the thinking process in the human brain to interpret information, create deep neural networks with analytical and learning properties, and imitate the way of thinking in the human brain to interpret and analyze data to find scattered information [4].

The unsupervised learning that rises from the bottom is to start from the bottom and train to the top layer by layer. Using uncalibrated data (or with calibrated data) to train the parameters of each layer, layer by layer, this step can be regarded as an unsupervised training process, which is also the biggest difference from the traditional neural network, which can be regarded as a feature learning process. Specifically, the first layer is trained with uncalibrated data, and the parameters of the first layer are learned during training. This layer can be regarded as the hidden layer of a three-layer

neural network that minimizes the difference between the output and the input. Due to the limitation of model capacity and sparsity constraints, the resulting model can learn the structure of the data itself, thereby obtaining features that are more expressive than the input. After learning the $n-l$ layer, the output of the $n-l$ layer is used as the input of the n th layer, and the n th layer is trained, thereby obtaining the parameters of each layer.

Top-down supervised learning is to train with labeled data, and the error is transmitted top-down to fine-tune the network. Based on the parameters of each layer obtained in the first step, the parameters of each multilayer model are further optimized. This step is a supervised training process. The first step is similar to the random initialization initial value process of neural network. Since the first step is not randomly initialized, but obtained by learning the structure of the input data, this initial value is closer to the global optimum so that better results can be achieved. So the good effect of deep learning is largely due to the process of feature learning in the first step.

3.2. Data Mining. Data mining, also known as knowledge discovery in databases, is a hot topic in the field of artificial intelligence and big data and has attracted much attention. Data mining is to analyze and calculate a large amount of data in the database and then reveal hidden, previously unknown data, and this process is of extraordinary significance. Data mining itself is carried out without a clear target, so the results after mining data are also uncertain, which can indicate the future development direction and have a profound impact on the trend of e-commerce.

Commonly used methods of data mining include the following:

Classification: the classification method is mainly to search and classify the general characteristics of one or more data groups in the database according to a specific classification model. Its main purpose is to map the information in the database into specific categories. It can be used to classify customers and predict customer characteristics and characteristics, customer satisfaction, and customer development. For example, a car dealership can rank customers based on their car-related preferences. Then marketers can mail advertisement brochures of different types of new cars to users according to this preference, so through this method, business opportunities can be greatly increased [17].

Regression analysis: regression analysis methods usually describe the value attribute of a specific attribute in a database over time, map the actual value of the generated data column to a function of the predictor variable, and look for dependencies between variables or attributes.

The cluster analysis method is to divide the dataset into different categories according to the difference and probability, and its main purpose is to minimize the similarity and similarity of the data belonging to different categories, as wide as possible.

Association rules: the association rules method is an association that describes the relationship between data

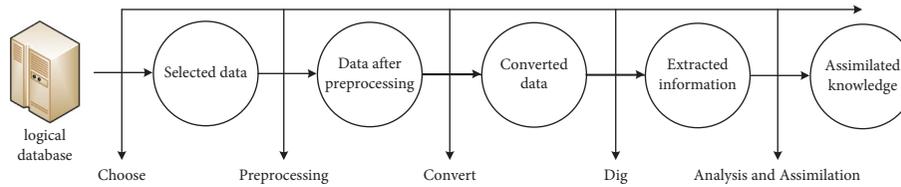


FIGURE 2: The basic process of data mining.

elements in a database. In other words, depending on the occurrence of certain parts of an event, other elements of the same event can be searched for, such as hidden links in the data or links between them.

Feature: feature analysis method is to extract the feature formula of these data from a set of data. These feature expressions represent the general characteristics of the dataset. For example, by extracting the main characteristics of customer churn, marketers can obtain the main reasons and characteristics that lead to customer churn and then avoid a large number of customer churn according to these reasons and characteristics.

Variation and bias analysis: bias includes a large amount of latent and interesting knowledge, such as abnormal instances in classification, abnormal patterns, and observational bias in expectations. Often unexpected rules can hide huge benefits. Once these abnormal rules have found their potential value, the benefits are immeasurable.

Data mining is a step in finding information. Data mining refers to processing algorithms to find hidden information in a large amount of data. With the massive generation of data, big data technology is becoming more and more popular, and big data platform technology is also making continuous progress. Although the data processing at the big data level is not much different from the previous data processing in terms of extraction and algorithm due to the difference in the width and size of big data, the implementation of big data mining will also be somewhat different. The common solution is to redesign the data mining algorithm under the big data platform and realize the expansion of the processing data volume through the dynamic expansion of the cluster [18]. The basic process of data mining is shown in Figure 2.

Data mining, also known as database recovery, is the use of sample computations to extract unknown and potentially valuable data from noisy, inconsistent, large amounts of data, and random data. Data mining can process many types of data, which can be divided into three types: unorganized data (such as video, text, etc.), organized data (such as data stored in communication databases), and unstructured data (such as biotechnological data, data, etc.). Decision makers can find hidden links in data by processing data and identifying ignored factors and data. Query optimization, decision support, information query, process control, etc. all need to use the discovered knowledge. It is a great help in predicting and decision-making behavior. Nowadays, data mining technology is more and more mature and reliable, and it can be applied in more and more scenarios, such as some of the fields listed below:

The field of marketing, in the field of marketing, usually analyzes the actual needs of customers, and according to the customer's consumption habits and consumption characteristics, a simple and direct management is carried out for different customers, hoping to achieve the purpose of smooth product sales and improving the success rate of personal sales. The scope of sales has also developed from the early supermarket shopping to other businesses such as banking and insurance.

In the field of scientific research, in the field of science, scientific research needs to make a lot of experimental tests, and it is necessary to perform complex analysis of experimental data, summarize the reasons for failure, and make adequate preparations for the next experiments. However, the data generated by the experiment are usually huge, so data mining technology is also widely used in the field of scientific research.

Cluster analysis is one of the important research directions in the field of data mining. Fuzzy clustering algorithm is a mathematical method based on fuzzy mathematics, first, a brief introduction to fuzzy mathematics. The revolutionary discipline of mathematics deviates from the absolute black-and-white relationship of classical science. Fuzzy mathematics can accurately analyze and filter complex data with uncertainty. The fuzzy clustering algorithm first describes the properties of the searched objects through a fuzzy table and displays the clusters according to the appropriate membership level. The purpose of clustering is to group small disparate data into one category, and the differences between categories should be clear [19].

Generally speaking, traditional techniques and improved techniques are two distinct branches of data mining theoretical techniques. In addition, data mining objects are mostly variable, with large numbers of samples used to simplify and alter multivariate analysis contained in higher statistics. Factor analysis, discriminant analysis for classification, and group analysis for partitioning groups are especially commonly used in data mining processes.

Artificial neural networks, also known as Neural Networks or known as Connectionist Models, are abstractions and modeling of the basic features of the human brain or natural Neural Networks. Artificial neural network is a physiological research result based on the brain. Its purpose is to simulate some mechanisms of the brain and realize some of its functions.

Decision Tree is a decision analysis method based on various conditional probabilities. By building a decision tree to obtain the probability that the expected value of the net present value is equal to or greater than zero, the project risk

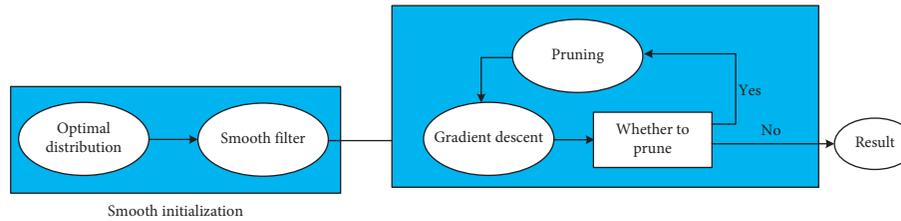


FIGURE 3: Improved deep learning framework algorithm flow.

is assessed and its feasibility is determined. This is a graphical method using probabilistic analysis. In machine learning, a decision tree is a predictive model. It represents the mapping relationship between object properties and object values.

4. The Spring Festival Holiday Tourism Data Mining Experiment

The improved deep learning framework proposed in this paper includes the following elements: (1) using good starting variables and smoothing layers in compressed form to improve network performance. (2) Normal network training and powerful editing methods eliminate unnecessary and useless process variables, simplify network organization, and improve the computational efficiency of the network. The algorithm flow of this framework is shown in Figure 3. First, the optimal distribution is selected to initialize the network parameters, and the initial parameters are smoothed. During the training of the network, after each period of gradient descent training, dynamic pruning is performed on the current network. After several times of dynamic pruning, a new network with significantly simplified structure and basically unchanged accuracy is obtained [20].

The main steps are as follows:

- (1) Determining the business object. Although the results of data mining have unpredictable characteristics, the problems that need to be mined are very clear, and the results of blind data mining will not be successful.
- (2) Data preparation. The selection of data collects all external and internal data information related to the business object, and then selects the appropriate data. In data preprocessing, in order to ensure the quality of data mining results, it is necessary to analyze the quality of the data itself and then prepare for the next process. And according to the purpose of excavation, the type of excavation to be performed is determined. Data conversion, converting data into an analysis model built for mining algorithms, and establishing an appropriate analysis model are the key factors for the success of data mining.
- (3) Data mining. Except for the selection of appropriate mining algorithms, everything else can be done automatically.

- (4) Result analysis. The results of data mining need to be interpreted and evaluated. The analysis method used is generally determined by data mining operations, and visualization techniques are currently used.

The dataset is trained and classified, all parameters in the network are initialized by relevant methods, and the network is recorded as the original network. The test accuracy is shown in Table 1.

The first fully connected layer in the introduced shallow neural network is replaced with a convolutional pooling layer with the same number of channels, and other parameter settings remain unchanged. For ease of identification, this network is named convolutional pooling network. Likewise, the performance of convolutional pooling networks between nonsmooth initialization and smooth initialization is compared, and for different kinds of layers in the network. A typical fully connected layer is shown in Figure 4.

Different smoothing filters make the parameters have different distributions, and the different parameter distributions are reflected in the different precision performances of the neural network. A comparison of different smoothing filters is shown in Table 2.

After deleting unimportant parameters in the neural network, it is sometimes necessary to continue deleting unnecessary neurons (nodes). After parameter pruning, some hidden layer neurons will have no input connection at all or no output connection. This makes these neurons out of the decision-making process of the network, so they need to be deleted to further simplify the structure of the network [21]. Stranded neurons are shown in Figure 5.

Since we use a convolutional pooling layer in the flattening layer, we can prune more parameters in this layer. This shows that using the initialization and pruning method proposed in this paper, the deleted parameters do not make a relatively large contribution to the decision-making process of the network. And the subsequent training process is only a fine-tuning of the network accuracy, rather than relearning the connectivity between the remaining neurons. The pruning results on the dataset are shown in Table 3.

The focus of this paper is on how to prune redundant parameters in the normal training process without affecting the network performance as much as possible, without further compressing the remaining parameters with any subsequent steps. Table 4 shows the test accuracy and the results on the dataset.

TABLE 1: Test accuracy of two hidden layer convolutional neural networks.

Initialization method		Nonsmooth initialization		
Training layer	Convolutional layer	All pass layer	All layers	
Original network	95.54%	95.21%	98.74%	
Convolutional pooling network	94.78%	94.91%	98.78%	
		Smooth initialization		
Training layer	Convolutional layer	All pass layer	All layers	
Original network	88.71%	92.34%	98.24%	
Convolutional pooling network	92.74%	93.73%	98.54%	

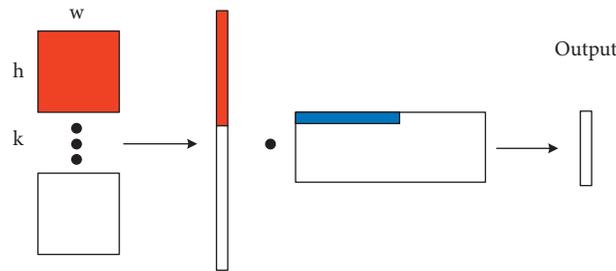


FIGURE 4: A typical fully connected layer.

TABLE 2: Comparison of different smoothing filters.

	Not smooth	3 × 3 Gaussian	5 × 5 Gaussian	3 × 3 Median	5 × 5 Median	
Test accuracy	95.48%	95.62%	65.41%	95.58%	95.62%	
Pruning	Conv1	25 (39%)	30 (45%)	32 (49%)	30 (42%)	28 (44%)
	Conv2	4.5 K	4.5 K	5.1 K	4.5 K	4.4 K
	Conv3	8.7 K	7.8 K	9.4 K	6.8 K	7.2 K
	Precision	95.44%	95.45%	95.32%	95.47%	95.52%

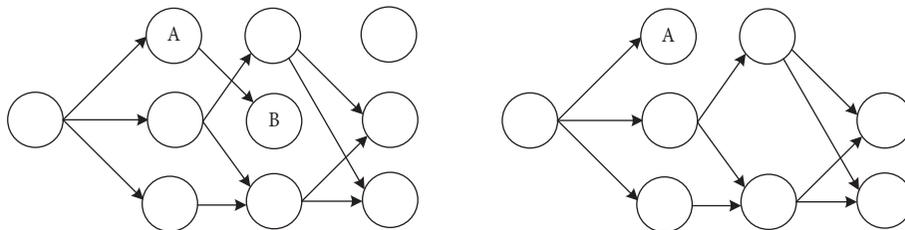


FIGURE 5: Stranded neurons.

TABLE 3: Pruning results on the dataset.

	Convolutional layer 1	Convolutional layer 2	Flattening layer	Output layer	Total
Total number of parameters	0.5 K	25 K	350 K	5 K	421 K
Threshold	0.02	0.05	0.015	—	—
Pruning ratio 1	34%	87%	91%	82%	93%
Pruning ratio 1	30%	74%	95%	—	94%

The number of data in the training set for the experiment is 7494, and the number of data in the test set is 3498. The classification result statistics are shown in Figure 6.

Using the constructed deep learning model to mine user preferences, travel popularity, and other data, a vacation travel recommendation system is constructed. The vacation travel recommendation system has three important modules: user modeling module, recommendation object

modeling module, and recommendation algorithm module. The recommender system model process is shown in Figure 7.

Using a free crawler software to obtain national tourism information from certain tourism official websites. It includes the name, rating, recommendation index, and user comments of a scenic spot. The user's comments mainly obtain the time of the user's comment, and the user's

TABLE 4: Test accuracy and results on the dataset.

Training/pruning method	Test accuracy (%)	Parameter compression ratio
Original network	93.55	1X
GD	92.71	2X
DGC	93.65	3X
Before dynamic pruning	94.12	1X
Dynamic pruning (36%)	93.54	4X
Dynamic pruning (41%)	93.85	7X

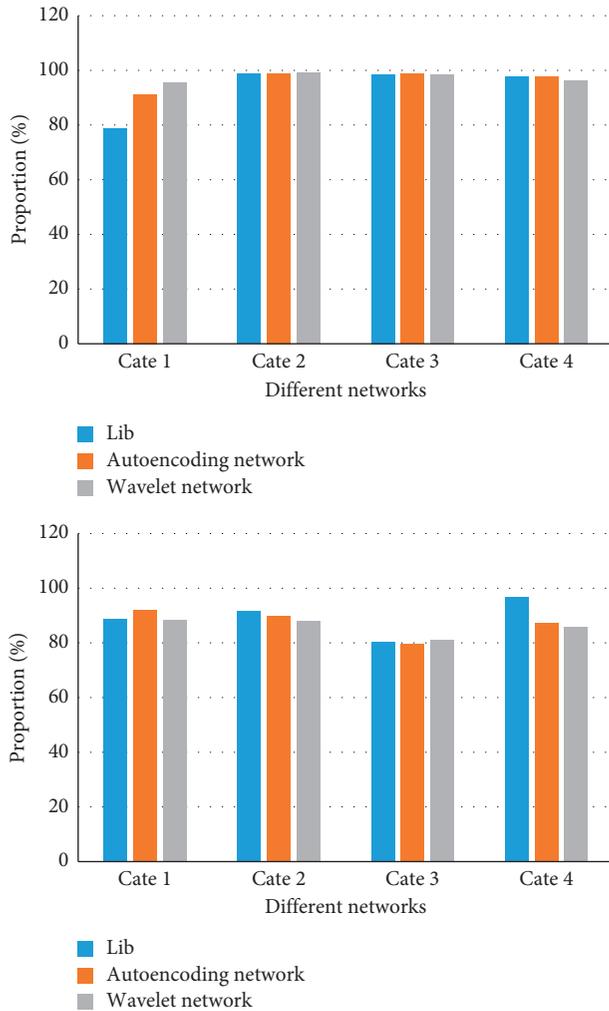


FIGURE 6: Classification result statistics.

comment is used to judge whether the user is a local order or a remote order [22]. Using WEKA to mine the data, the results are shown in Figure 8.

From the analysis, the following conclusions can be drawn: among the tourist attractions in city A, the tourist attractions with local characteristics are more inclined to visit by nonlocal tourists, and the locals of popular attractions are more inclined to visit. Therefore, it can be inferred that tourists from different places prefer different tourist attractions, local tourists prefer popular scenic spots, and nonlocal tourists prefer local special scenic spots. In summer, tourists are more inclined to go to scenic spots with

water sports, while in winter, tourists are more inclined to go to scenic spots with hot springs such as health spas and tourist resorts. Therefore, it can be inferred that the number of orders for tourist attractions will be different in different seasons, especially some attractions with water activities, which are more seriously affected by seasons. It was also found through analysis that couple travelers prefer scenic spots with the sea, and parent-child travelers prefer scenic spots such as safari parks. Therefore, it can be inferred that different types of travelers have different favorite tourist attractions [23].

Due to the variability of Spring Festival travel information, many external factors need to be considered. The so-called external factors refer to factors such as seasons, weather, user preferences, and historical behaviors of users, due to many factors. Therefore, when generating a recommendation candidate set, only relying on a certain strategy cannot solve the user problem very well. Therefore, it is necessary to build a set of combined strategies and carry out at the same time to ensure the rationality of the recommended candidate set for users. After the recommendation candidate set is generated, it is necessary to sort the generated recommendation candidate set according to a certain sorting strategy. The purpose of sorting is that those information will be recommended to users first, and those information will be filtered out because they are unreasonable. Therefore, there are three steps when constructing a recommendation candidate set. The first step is called a recall strategy. This step is to generate a recommendation candidate set according to the user's behavior information. The second step, called a filtering strategy, filters the initially generated candidate set. The third step is called the sorting strategy, which sorts the generated candidate set to generate the final recommendation set. As shown in Figure 9, the recall strategy for generating the recommended candidate set is shown.

The recall strategy for generating the recommended candidate set consists of four recommendation strategies. The first strategy is a strong correlation strategy for user historical behavior, which is mainly for users to browse and collect unpurchased travel products to generate a recommended set. The second strategy is a collaborative filtering strategy based on sights browsing behavior and user real-time search behavior. The third one is a recommendation strategy based on geographic location, and the last one is a substitute strategy, which is a popular city recommendation strategy when users cannot apply the above three strategies.

The system proposed in this paper is optimized, as shown in Figure 10. The left picture is a schematic diagram of coverage before and after optimization, and the right picture is the score of the results of the four recall strategies.

For users, the strong correlation strategy of historical behavior and the collaborative filtering strategy are personalized recommendations for users, and this recommendation is also the best way to express the user's interest. The location-based recommendation and popular city recommendation are an auxiliary recommendation because the historical behavior strong correlation strategy and the collaborative filtering strategy are based on the

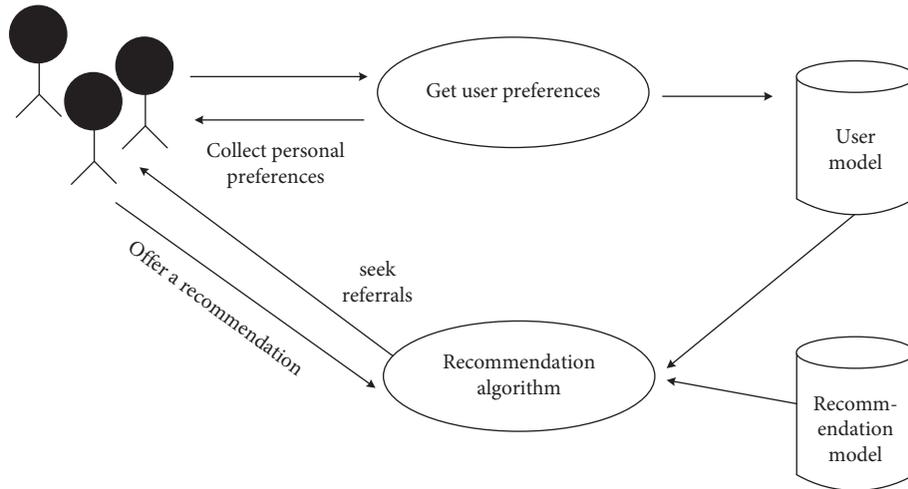


FIGURE 7: A general model for recommender systems.

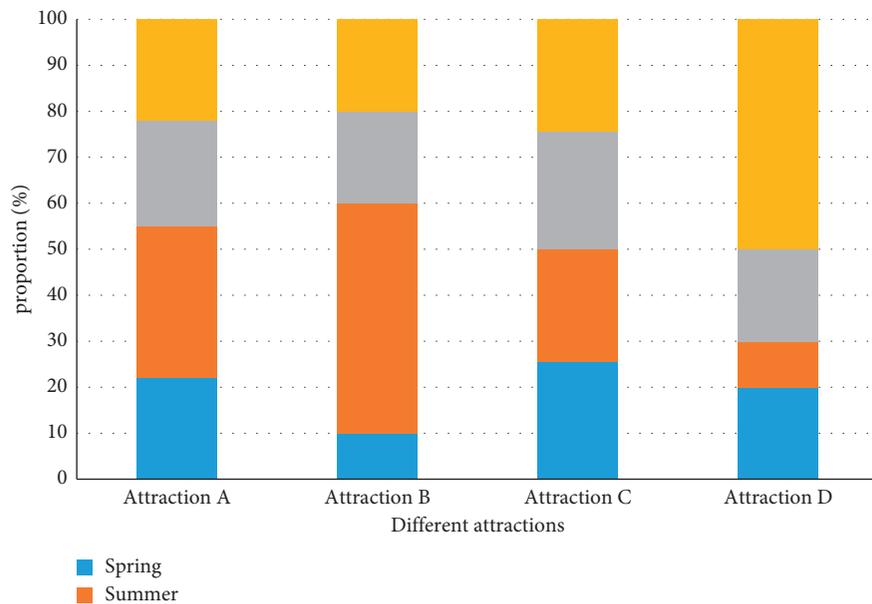
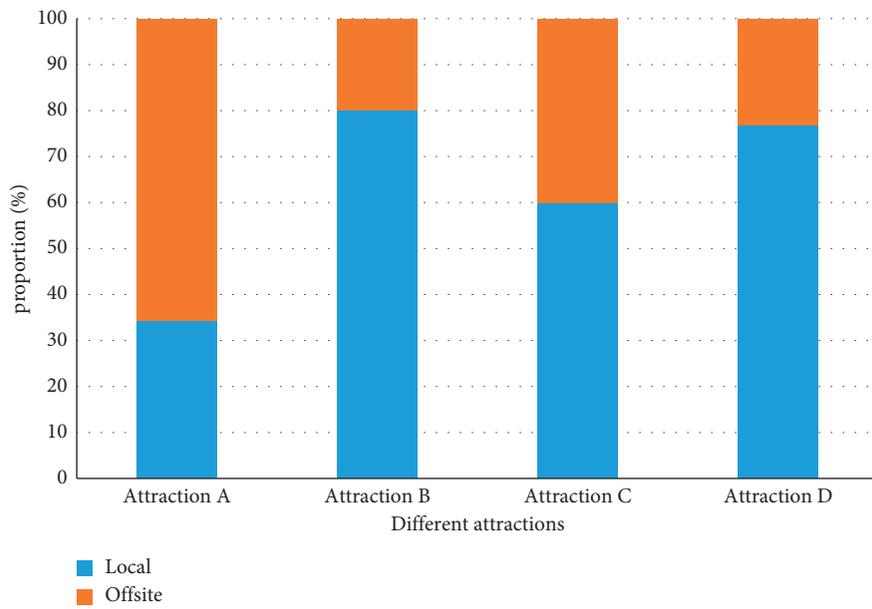


FIGURE 8: Order sources for different attractions and data graphs for different seasons.

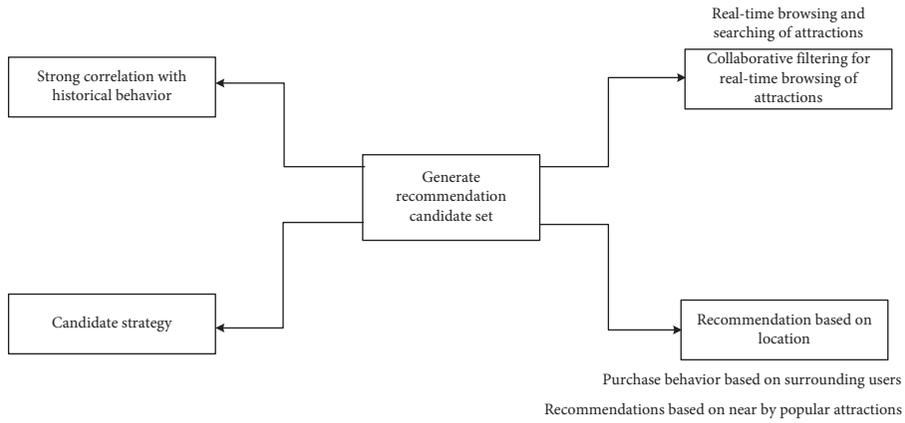


FIGURE 9: Generate recommendation candidate set recall strategy.

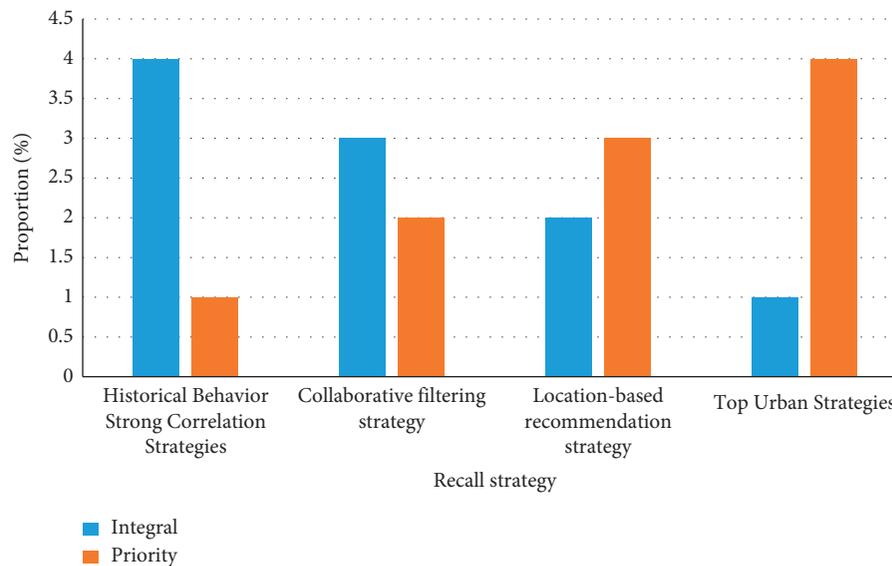
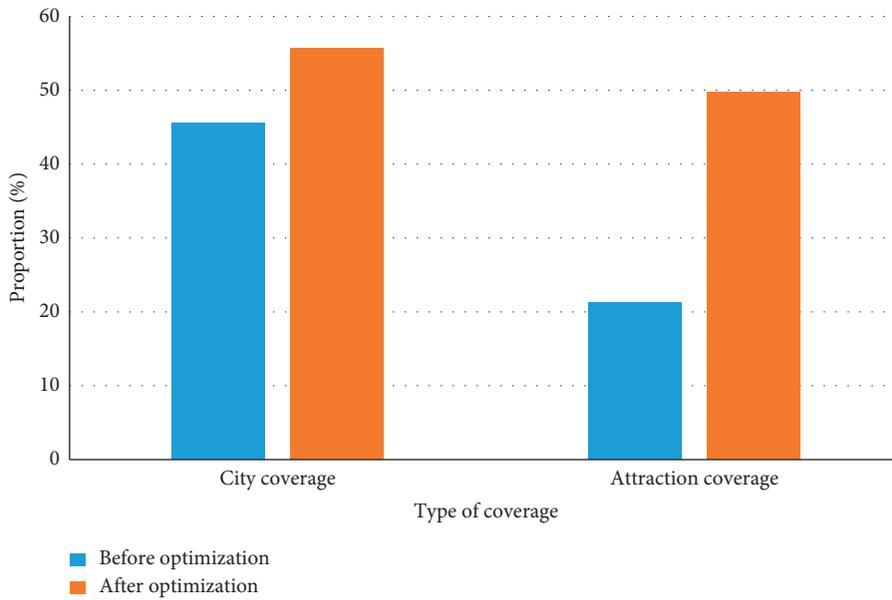


FIGURE 10: Schematic diagram of coverage before and after optimization.

user's behavior. If the user does not have any historical behavior, then this situation must rely on the recommendation based on location and the recommendation of popular cities, which are aimed at a popular recommendation method and can only indicate the interest of most people. Therefore, the results of the four recall strategies can be given a point, and the result set obtained by the strong correlation strategy of historical behavior is given 4 points. The result set obtained by the collaborative filtering strategy is given 3 points, the result set obtained by the geographic location-based recommendation strategy is given 2 points, and the result set of the last popular city strategy is given 1 point [24].

5. Discussion

The travel recommendation system needs to record the user's historical behavior. It includes the tourist attractions that users purchased, browsed, and collected, comments and ratings on purchased attractions, comments on some guides, and interactions with authors who wrote guides, and filter, analyze, and mine the information to form the user's preference model. Then, in the recommendation engine, the user's preference model is used as input to obtain the recommended item set. Finally, the recommended item set is displayed to the user, and the user's feedback results are recorded, whether they are satisfied with the recommended results, and what needs to be improved. The system analyzes and optimizes according to the feedback results and adjusts a new user preference model. Therefore, the functions of a complete travel recommendation system can be divided into the following modules: user behavior collection, collection information preprocessing, collection information mining, establishment of user preference model, item recommendation, UI display, and user feedback analysis.

The tourism information recommendation system based on data mining is to build a user-oriented personalized tourism information recommendation platform based on the existing tourism information system, combining data mining knowledge and recommendation system algorithm. First, through a large amount of tourism information, the key factors affecting the choice of tourist attractions by tourists are excavated. Through the analysis of the key factors, the difficulties in realizing the tourism recommendation system are obtained. Then, aiming at the difficulties, a combined recommendation strategy is proposed, and the collaborative filtering algorithm is optimized, which can solve the coverage problem to a certain extent. Finally, some functions of the tourism information recommendation system are realized.

In different recommendation systems, the best recommendation algorithm should be selected according to specific scenarios. After decades of development, personalized recommendation algorithms have produced a variety of recommendation algorithms. These algorithms have their own unique application scenarios, advantages and disadvantages. The three most common recommendation algorithms are described below.

For content-based recommendation, the algorithm first needs to extract the feature points of the items, then calculate the similarity between the items, and then give the final recommendation through calculation. The advantage is that the recommended results are intuitive, easy to explain to users, and do not require a certain domain of knowledge. The disadvantage is that it is very time-consuming to extract features and calculate similarity, and usually requires a lot of offline calculations, but the recommended results are not ideal.

For collaborative filtering recommendation, this algorithm is the current mainstream recommendation technology, and it does not need to extract the features of the items nor the association rules of the items, but only needs to calculate the similarity between users or items. The collaborative filtering algorithm is suitable for personalized recommendation and can handle complex unstructured objects, but it also has inevitable shortcomings, such as poor recommendation quality at the beginning of the system and new user problems.

6. Conclusion

The tourism information recommendation system based on data mining is to build a user-oriented personalized tourism information recommendation platform based on the existing tourism information system, combining data mining knowledge and recommendation system algorithm. First, through a large amount of tourism information, the key factors affecting the choice of tourist attractions by tourists are excavated. Through the analysis of the key factors, the difficulties in realizing the tourism recommendation system are obtained. The paper proposes that different parameter distributions are reflected in different precision performances of neural networks and obtains a comparison of different smoothing filters. In this paper, a preliminary prediction research is carried out. Given the limited data sources and academic level, the research will inevitably have omissions. In the current situation analysis stage, the analysis is not thorough enough, only showing the changes of relevant indicators, and lack of internal judgment analysis; in the theoretical research stage, the grasp of the theory is not deep enough. The improved algorithm in this paper only solves the coverage problem to a certain extent, but the cold start problem of the recommendation system is still not well solved. When the user's historical record does not exist, it is converted to recommend popular attractions in the target city to the user. Although the recommendation can be achieved, it cannot solve the problem of user personalization.

Data Availability

No data were used to support this study.

Conflicts of Interest

The author declares that there are no conflicts of interest regarding the publication of this article.

References

- [1] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, and C. I. Sánchez, “A survey on deep learning in medical image analysis,” *Medical Image Analysis*, vol. 42, no. 9, pp. 60–88, 2017.
- [2] D. S. Kermany, M. Goldbaum, W. Cai, and C. C. S. Valentim, “Identifying medical diagnoses and treatable diseases by image-based deep learning,” *Cell*, vol. 172, no. 5, pp. 1122–1131, 2018.
- [3] H. He, C.-K. Wen, S. Jin, and G. Y. Li, “Deep learning-based channel estimation for beamspace mmWave massive MIMO systems,” *IEEE Wireless Communications Letters*, vol. 7, no. 5, pp. 852–855, 2018.
- [4] W. E and B. Yu, “The Deep Ritz method: a deep learning-based numerical algorithm for solving variational problems,” *Communications in Mathematics and Statistics*, vol. 6, no. 1, pp. 1–12, 2018.
- [5] T. Young, D. Hazarika, S. Poria, and E. Cambria, “Recent trends in deep learning based natural language processing [review article],” *IEEE Computational Intelligence Magazine*, vol. 13, no. 3, pp. 55–75, 2018.
- [6] X. X. Zhu, D. Tuia, L. Mou, and G.-S. L. F. F. Xia, “Deep learning in remote sensing: a comprehensive review and list of resources,” *IEEE Geoscience and Remote Sensing Magazine*, vol. 5, no. 4, pp. 8–36, 2017.
- [7] S. Xie, Z. Yu, and Z. Lv, “Multi-disease prediction based on deep learning: a survey,” *Computer Modeling in Engineering and Sciences*, vol. 128, no. 2, pp. 489–522, 2021.
- [8] J. Lee, “Integration of digital twin and deep learning in cyber-physical systems: towards,” *Smart Manufacturing*, vol. 38, no. 8, pp. 901–910, 2020.
- [9] Y. Li, Y. Zuo, H. Song, and Z. Lv, “Deep Learning in Security of Internet of Things,” *IEEE Internet of Things Journal*, 1 page, 2021.
- [10] X. Sun, P. Wu, and S. C. Hoi, “Face detection using deep learning: an improved faster RCNN approach,” *Neurocomputing*, vol. 299, pp. 42–50, 2018.
- [11] H. Li, K. Ota, and M. Dong, “Learning IoT in edge: deep learning for the Internet of things with edge computing,” *IEEE Network*, vol. 32, no. 1, pp. 96–101, 2018.
- [12] a admin, I. e el-henawy, K. Ahmed, fnm au, I. e. el-henawy, and K. Ahmed, “An improved equilibrium optimizer algorithm for tackling global optimization problems,” *Fusion: Practice and Applications*, vol. 3, no. 1, pp. 01–28, 2021.
- [13] a. admin, A. M. Abdulazeez, fnm au, and A. M. Abdulazeez, “Electrocardiogram classification based on deep convolutional neural networks: a review,” *Fusion: Practice and Applications*, vol. 3, no. 1, pp. 43–53, 2021.
- [14] J. Han, D. Zhang, G. Cheng, and N. D. Liu, “Advanced deep-learning techniques for salient and category-specific object detection: a survey,” *IEEE Signal Processing Magazine*, vol. 35, no. 1, pp. 84–100, 2018.
- [15] M. Elsharkawy, N. Ahmed, and A. Masri, “A novel image encryption with deep learning model for secure content based image retrieval,” *Journal of Cybersecurity and Information Management*, vol. 0, no. 2, pp. 54–64, 2019.
- [16] M. Hammoudeh, S. M. Aljaberi, fnm au, and S. M. Aljaberi, “Modeling of deep learning based intrusion detection system in Internet of things environment,” *Journal of Cybersecurity and Information Management*, vol. 8, no. 1, pp. 17–25, 2021.
- [17] A. S. Heinsfeld, A. R. Franco, A. Buchweitz, and F. Meneguzzi, “Identification of autism spectrum disorder using deep learning and the ABIDE dataset,” *NeuroImage: Clinic*, vol. 17, no. C, pp. 16–23, 2018.
- [18] N. Sünderhauf, O. Brock, W. Scheirer, and R. D. J. B. P. W. M. P. Hadsell, “The limits and potentials of deep learning for robotics,” *The International Journal of Robotics Research*, vol. 37, no. 4-5, pp. 405–420, 2018.
- [19] Q.-s. Zhang and S.-c. Zhu, “Visual interpretability for deep learning: a survey,” *Frontiers of Information Technology & Electronic Engineering*, vol. 19, no. 1, pp. 27–39, 2018.
- [20] S. S. Han, M. S. Kim, W. Lim, and G. H. I. S. E. Park, “Classification of the clinical images for benign and malignant cutaneous tumors using a deep learning algorithm,” *Journal of Investigative Dermatology*, vol. 138, no. 7, pp. 1529–1538, 2018.
- [21] M. Araya-Polo, J. Jennings, A. Adler, and T. Dahlke, “Deep-learning tomography,” *The Leading Edge*, vol. 37, no. 1, pp. 58–66, 2018.
- [22] K. Grm, V. Štruc, A. Artiges, M. Caron, H. K. Ekenel, and H. K. Ekenel, “Strengths and weaknesses of deep learning models for face recognition against image degradations,” *IET Biometrics*, vol. 7, no. 1, pp. 81–89, 2018.
- [23] R. Kemker, C. Salvaggio, and C. Kanan, “Algorithms for semantic segmentation of multispectral remote sensing imagery using deep learning,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 145, pp. 60–77, 2018.
- [24] R. Ranjan, S. Sankaranarayanan, A. Bansal, and N. Bodla, “Deep learning for understanding faces: machines may Be just as good, or better, than humans,” *IEEE Signal Processing Magazine*, vol. 35, no. 1, pp. 66–83, 2018.