Journal of Network Intelligence ©2025 ISSN 2414-8105
Taiwan Ubiquitous Information Volume **10**, Number **4**, November **2025**

Research on Consumer Churn Prediction Based on Sentiment Analysis of Product Reviews

Qian Zhao*

School of Humanity and Foreign Languages
Xi'an University of Posts and Telecommunications
Xi'an 710121, P. R. China
m18161835690@163.com

Tian-Li Yu

National University of the Philippines
Manila 0900, Philippines
ve2056@163.com

*Corresponding author: Qian Zhao
Received September 5, 2024, revised March 9, 2025, accepted August 21, 2025.

Abstract. *Consumers share their experience of using products and services through online reviews, which contain their emotional tendencies. And customer churn is the process of customers' emotional changes towards products and services from positive to negative. Positive online comments correspond to low risk of user churn, while negative online comments correspond to high risk of user churn. Therefore, the risk of customer churn can be determined through sentiment analysis of users' online comments. Unlike the previous approach of studying the influencing factors of customer churn based on user consumption behavior data to construct a churn prediction index system, this study is based on online comment information and starts from the perspective of user emotions. Through word level sentiment analysis technology, the sentiment tendency of comments is quantified into two categories: positive and negative emotions. Then, BP neural network is used to model customer churn prediction for user churn prediction. The research results indicate that the sentiment polarity of comment content is the main indicator for identifying customer churn, and the method proposed in this paper has high accuracy and stability. This study provides a new perspective and method for researching customer churn.*
Keywords: customer churn; online comments; BP neural network

1. **Introduction.** With the rapid development of Internet technology, online shopping and service platforms have been widely used in the current digital and information era. Consumers are increasingly sharing their experiences using products and services through online reviews, which not only influence the purchasing decisions of other potential consumers, but also provide important references for businesses to understand customer needs and improve product and service quality. However, customer churn remains one of the main challenges that businesses face in the fierce market competition. Customer churn usually refers to consumers no longer choosing a certain brand or service provider, but turning to competitors, which can have a serious impact on the company's revenue and market share. Therefore, how to effectively prevent and predict customer churn has become an urgent issue for companies to solve [1,2]. Traditional research on customer churn prediction is mostly based on user consumption behavior data [3]. By analyzing behavioral

indicators such as purchase frequency, consumption amount, and usage frequency, prediction models are constructed [4]. However, with the increasing amount of unstructured data generated by users, especially online comments, user sentiment data has gradually become an important source for studying customer churn. Online comments often contain users' emotional tendencies, that is, their satisfaction or dissatisfaction with the product or service. Research has shown a significant correlation between changes in user emotions and customer churn: positive emotional tendencies typically correspond to lower churn risks, while negative emotional tendencies may indicate higher churn risks. Therefore, by analyzing the emotional tendencies in online comments, new ideas and methods can be provided for predicting customer churn. Sentiment analysis techniques, especially word level sentiment analysis, can deeply explore and quantitatively analyze user comments, converting their emotional expressions into specific indicators of emotional tendencies. By identifying positive or negative emotions contained in comments, it is possible to effectively assess changes in users' attitudes towards products or services, thereby providing a basis for predicting the risk of customer churn. However, the complexity and challenges of sentiment analysis technology in practical applications cannot be ignored. User comments usually contain rich language expressions and complex semantic structures. How to accurately capture and analyze these emotional information is the key to the research of sentiment analysis technology. This study uses online comments as the main data source, focusing on the perspective of user emotions. Through word level sentiment analysis techniques, the emotional tendencies of comments are quantified into two categories: positive and negative. Subsequently, a BP neural network was used to model customer churn prediction and construct a model that can accurately predict customer churn risk. In the process of model construction, the study not only considered the emotional polarity of comment content, but also integrated multiple online comment information such as comment star rating, likes, replies, etc., to analyze their impact on customer churn. Through empirical analysis, the research results indicate that emotional polarity in online comments is indeed an important indicator for identifying customer churn, especially based on variables such as membership level and comment star rating. Emotional polarity significantly affects the judgment of customer churn risk. A customer churn prediction method based on online comments is proposed from the perspective of user emotional expression through sentiment analysis technology. This method is different from traditional behavioral data analysis, starting directly from users' emotional changes, providing a new perspective and technical support for enterprise customer management and churn prevention. At the same time, this study also laid the foundation for further development and application of customer churn prediction models based on sentiment analysis, which has strong practical guidance significance.

1.1. **Related work.** With the rise of artificial intelligence and the further arrival of the big data era, many researchers have used various methods to construct different customer churn prediction models and have achieved some results. Around 95% of big data is made up of unstructured data, which can be sourced from a variety of locations both within and outside the company. There is a growing body of research focused on unstructured data across different fields. Customer churn prediction is crucial in various industries, and the biggest challenge is to distinguish which customers have a tendency to churn among many. The primary approaches for studying customer churn at present include logistic regression, decision trees, random forests, and neural networks. Bahnsen et al. [5] proposed a sample-based cost sensitivity decision tree model based on cost sensitivity research. This model has achieved good results in fraud detection, credit scoring, and market sales datasets. Kasiran et al. [6] predicted the churn of mobile phone users based on reinforcement learning algorithms and recurrent neural networks. Kuanchin et al. [7] demonstrated the effectiveness of the "length, proximity, frequency, monetary value, and profit" (LRFMP) customer value model in predicting customer churn for a logistics company. Sharma et al. [8] explored the behavior of users in online social platforms and some social functions of new users who have not yet created links or small networks in their research, in order to predict links and consumption preferences, nd introduced novel direct and latent models to represent link prediction and user consumption preferences on online social networks. The research team led by NhaLe et al. [9] identified reliability, product quality, and customer understanding as the key factors influencing the quality of port logistics services in Vietnam, with reliability and responsiveness also playing significant roles. This suggests that customers prioritize a port's professionalism in addressing their concerns, particularly its commitment to transportation safety, efficient customs clearance, and timely delivery or receipt. Additionally, customer satisfaction is positively correlated with overall service quality, highlighting the importance of strategically focusing on market share growth in the port logistics sector. All assumptions have been validated; the following conclusion can be drawn: high-quality service quality paves the way to customer satisfaction. Therefore, whether for port logistics customers or ordinary logistics customers, high-quality service quality is an important characteristic that customers need logistics companies to embody the most. Verbeke et al. [10] investigated the use of social network data to predict customer churn. In response, they developed a different modeling approach that uses relational learning algorithms to incorporate social network effects into customer churn prediction scenarios, efficiently handling large-scale networks, time-sensitive class labels, and imbalanced class distributions. According to the above research findings, current studies on consumer churn prediction mainly focus on e-commerce platforms and the operator industry. The evaluation indicators of existing customer churn prediction methods are not comprehensive, and the churn prediction system is not sound. On the basis of existing research, further combining sentiment analysis of product reviews to achieve customer churn prediction is the main problem that needs to be addressed in this article. Customers' online messages and other information can be used as the main features of customer churn prediction, combined with relevant prediction methods, to achieve consumer churn prediction.

1.2. **Contribution.** The above related research work provides a space for in-depth exploration in this study. The primary contributions of this article include the following:
(1) Innovatively applying sentiment analysis to customer churn prediction: Unlike traditional customer churn prediction methods based on user consumption behavior data, this study, for the first time, starts from the perspective of user sentiment and uses emotional information in online comments to predict customer churn risk. By introducing word level sentiment analysis techniques, the emotional tendencies of comment content are quantified into two categories: positive and negative, providing a new perspective and method for studying customer churn.
(2) Constructing a customer churn prediction model based on BP neural network: This study uses BP neural network to model the sentiment analysis results of user online comments, and incorporates multiple online comment information such as membership level, comment star rating, likes, replies, etc. into the model. Compared to traditional statistical models, BP neural networks have stronger nonlinear fitting capabilities and can more accurately capture the complex relationship between user emotions and customer churn, thereby improving the accuracy of customer churn prediction.
(3) Empirical analysis confirms the importance of sentiment polarity in comments: Through empirical analysis, this study found that sentiment polarity in online comments plays a

significant role in identifying customer churn. The research results indicate that the method proposed in this paper has high stability and accuracy.
(4) Enriched the research system of customer churn prediction: This study not only expands the research framework of customer churn prediction from an emotional perspective, but also provides new methods and tools for enterprises to identify potential high-risk customers in advance through emotional analysis of online comments, and take effective intervention measures to reduce customer churn rates in practical operations.

2. **Theoretical analysis.**

2.1. **Collection and preprocessing of comment text.**

2.1.1. *Comment text preprocessing.* It was found that JD's comments have a large amount of text and rich content, covering various aspects such as products and services. The review includes several parts: overall product evaluation, name, content, time, star rating, username, location, source, and popular tags. Therefore, some telecom mobile phone card products of JD MVNO were selected, and their comment information was captured, including comment content, membership level, star rating, likes, replies and other fields, with a time span of 2022 10 - 2023. 04, a total of over 10000 items. Although the corpus of CNKI Dictionary is relatively complete, some emerging network terms and non standardized terms have not been included. It is necessary to add some non standardized terms related to telecommunications vocabulary, with high frequency of comment text words, on the basis of CNKI Emotion Dictionary to meet the basic segmentation requirements. The added vocabulary includes telecommunications vocabulary such as Pioneer Card, Mighty Card, Number Segment, Activation, Real Name Authentication, etc. To improve the accuracy of word segmentation, ICTCLAS is used to segment the comment text after ignoring stop words.

2.1.2. *Text feature selection and sentiment classification.* After segmentation, not every word can be studied as a feature, which results in a high computational load and low research value for low-frequency words. Therefore, text feature selection is used to reduce dimensionality. Feature selection requires identifying representative words with strong relevance to the research content, and removing words with poor relevance. To improve efficiency, sentiment lexicons can be used to label sentiment words and classify overall emotions. The CNKI Dictionary has divided emotional words into two categories: positive and negative. Positive emotions include liking, satisfaction, and appreciation, while negative emotions include regret, dissatisfaction, disappointment, and high value. Simply match the representative emotional words in the comment text with the CNKI Dictionary as benchmark words. Emotional feature words can reflect the level of customer satisfaction with purchasing and using the product, and feature selection can also facilitate subsequent emotional classification processing after dimensionality reduction. Emotions can be classified into negative and positive, and a Boolean algorithm is used to construct an emotion classification model. Each user comment is represented as a document dataset $D = \{d_1, d_2, d_3, \ldots, d_n\}$, And the extracted feature items are represented as a set $V = \{t_1, t_2, \ldots, t_p\}$. Dataset D is used as the document, and Dataset V is used as the query, each consisting of a set of words. There are only two possibilities for querying whether the dataset is in the document: whether it appears or not (represented by 1 or 0).

2.2. **BP neural network.**

2.2.1. *Basic principles of BP neural network.* At the end of the 20th century, Rumelhart et al. designed an algorithm called backpropagation, also known as BP algorithm. BP stands for Back Propagation, which means backpropagation [11]. The signal of a neural network propagates forward, while the error propagates backward, which is its main characteristic that distinguishes it from other algorithms. The error is compared with the expected value through backpropagation, and the network is continuously adjusted when the sum of squares of the error is found to be large. The adjustment method is generally to change the threshold and weight values of the network. BP neural network (BPNN) forms mapping relationships between inputs and outputs during the network creation process, so these mapping relationships cannot be explained before the network is created. However, the internal structure of the network has already formed a certain computational program, so data can be input into the network and this mapping relationship can be run to predict other input data [12, 13]. The BP neural network has strong generality and is innovative in using mathematical methods to derive and implement data matrix analysis and prediction based on computer science theory [14]. Compared to other mathematical analysis methods, the BP neural network has a strong ability for nonlinear mapping, making it suitable for analyzing and predicting data sets with nonlinear relationships. Additionally, neural networks can continuously adjust and optimize the learned content, making BP neural networks effective for problems with complex internal mechanisms and uncertain factors [15]. BP neural networks also demonstrate strong fault tolerance; even when some data points in the training sample contain significant errors, the training results remain largely unaffected. This is due to the BP neural network's self-regulation capability, which independently adjusts the weights and thresholds of the learning algorithm. As a result, local damage to the BP neural network has minimal impact on overall training outcomes, allowing it to adapt to various simulation environments [16, 17].

2.2.2. *Basic structure of algorithm.* The BP neural network demonstrates robust fault tolerance, ensuring that the training results are minimally impacted even when some data points in the training sample contain substantial errors, and hidden layer. The neurons within each level are independent of each other, but they are connected to each other. The connection method is the edge neurons at the beginning and end of each level. These neurons constantly change their connection method through weights and thresholds, resulting in changes in the network's error [18, 19].

The complexity of the hidden layer affects the neural network's accuracy. Typically, as the number of hidden layers and neurons increases, the network's accuracy improves. To achieve more accurate predictions, more hidden layers and neurons need to be set. However, having too many of the two can also bring certain troubles. As the number of the two increases, the computational load of the neural network will grow exponentially. Not only will the network run slowly, but it is also prone to overfitting, resulting in a decrease in the accuracy of the neural network instead of an increase. Therefore, in general research, the hidden layers are generally set to 1–2 layers [20].

The number of neurons in the hidden layer also impacts the neural network's accuracy. While empirical formulas are often used to determine this number, it is not strictly confined to the calculated results and can vary slightly within a small range. The formula is given below:

$$n_j = +a \tag{1}$$

where n_j represents the number of neurons in the hidden layer; n_i and n_k rrepresent the number of indicators in the input and output layers, respectively; a is an integer within the range of $[1, 10]$, which can be chosen based on the specific circumstances.

In BP neural networks, the three-layer perception structure is the most widely used. Kolmogorov's theorem has demonstrated that a three-layer BP neural network structure can accurately map the relationship between nonlinear inputs and outputs with any desired level of precision, which can meet most analytical research purposes [21]. Its structural diagram is shown in Figure 1.

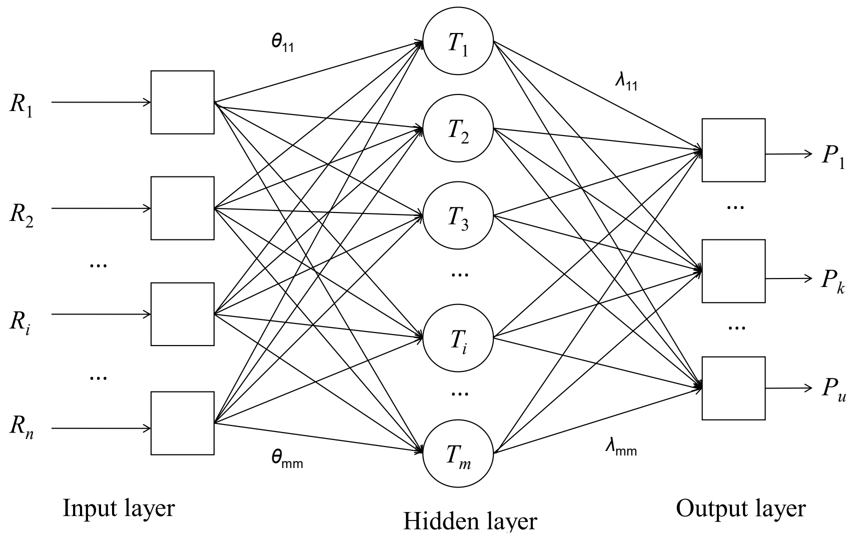

Figure 1. Structure diagram of three-layer neural network model

For the three-layer BP neural network, there is a certain transmission relationship and calculation method between the neurons in each layer. By following this relationship and method [22], the forward transmission of neural signals is completed. The specific conversion formula is as follows:

$$y_i = f\left(\sum w_{ij}\, x_i - \theta_j\right) \tag{2}$$

where w_{ij} is the coefficient conversion value between the input value and the output value, and θ_j represents the threshold from the input layer signal to the j-th node of the hidden layer.

$$y_i = f(\mathrm{net}_i) \tag{3}$$

where net_i represents the neural network structure formed by the i-th input node.

$$\mathrm{net}_i = \sum w_{ij}\, x_i - \theta_j \tag{4}$$

The calculated output of the output node is:

$$o_k = f\left(\sum w_{ij}\, y_i - \lambda_k\right) \tag{5}$$

where λ_k is the threshold from the hidden layer signal to the k-th output node of the output layer.

$$o_k = f(\mathrm{net}_i) \tag{6}$$

$$\mathrm{net}_i = \sum w_{ij}\, y_i - \lambda_k \tag{7}$$

The function f in the above two equations is usually referred to as a transformation function, which generally has the following form:

$$f(x) = \frac{1}{1+e^{-x}} \tag{8}$$

$$f(x) = \frac{1-e^{-x}}{1+e^{-x}} \tag{9}$$

To reduce computational complexity and improve model fitting efficiency, the output layer can use linear functions:

$$f(x) = kx \tag{10}$$

When the network's output signal o differs from the expected output d, an output error E is generated. The formula for this error is as follows:

$$E = \frac{1}{2}(d - o)^2 \tag{11}$$

2.2.3. *Basic functions.* **(1) Train function.** The train function is the fundamental function for training neural networks and is generally used to train relatively simple neural networks. Its function is called in the following way:

$$[\text{trainedNet, tr}] = \text{train}(\text{net}, X, T, X_i, A_i, EW) \tag{12}$$

If trainedNet is used to represent a pre trained neural network, then tr records the entire training process of the network, net represents the neural network to be trained, X and T represent the input values and target expected values of the matrix, X_i and A_i represent the conditions and output conditions required for matrix initialization, and EW represents the error weight. In neural networks, the train function can automatically match the error weights to the target matrix T. In use, the weights of any dimension of T can be defined according to the actual needs of the research problem, in order to distinguish the importance of the target values. The default value of error weight EW is 1. If $EW = 1$, it means that all target values have the same importance.

(2) Sim function. The SIM function is used to simulate neural networks, which can use pre trained neural networks to simulate and calculate a certain input matrix, output the predicted data of the input matrix, and achieve subsequent prediction accuracy calculation. Its function form is expressed as:

$$y = \text{sim}(\text{net}, x) \tag{13}$$

If y represents the predicted data generated by the neural network through computation, then in the sim function, x represents the input data of the target object that needs to be predicted, and net represents the neural network that has been trained before.

(3) Mapminmax function. The mapminmax function is used to standardize the input data matrix, that is, to standardize each row of data in the matrix to $[-1, 1]$, so that all input indicators are converted into a data matrix with consistent dimensions, and the data is unified into analyzable standards. Its function form is:

$$[Y, PS] = \text{mapminmax}(X, Y_{\min}, Y_{\max}) \tag{14}$$

where Y is the standardized data matrix, X is the original matrix, and $Y_{\min}$ and $Y_{\max}$ are the minimum and maximum values of each row of the standardized Y matrix, respectively. They are the control parameters for storing the standardized values. The general formula for this function is:

$$Y = (Y_{\min} - Y_{\max}) \cdot \frac{X - X_{\min}}{X_{\max} - X_{\min}} + Y_{\min} \tag{15}$$

In practical applications, this function has two other expressions:

$$Y = \text{mapminmax}(\text{'apply'}, X, PS) \tag{16}$$

$$X = \text{mapminmax}(\text{'reverse'}, Y, PS) \tag{17}$$

The former can obtain the standardized Y matrix under the given conditions of X matrix and control parameters, which is a function form of forward standardized data processing; The latter can restore the physical true values of the X matrix given the standardized Y

matrix and control parameter PS, that is, inverse normalization processing. Both forms will be used in the prediction model.

3. **Customer churn prediction model based on BP neural network.** Neural networks typically consist of input layers, output layers, and hidden layers. If the desired output is not reached, it indicates a difference between the actual output and the expected value. Then, the error path is returned, and the weights of each layer of neurons are continuously modified to minimize the error. Increasing the number of layers can reduce errors, but it will increase training time and complexity. Therefore, this article utilizes a three-layer BP neural network to build a consumer churn prediction model. This article opts to develop a 3-layer neural network model with j neurons in the input layer (where $j = 1, 2, \ldots, m$). For the design of the hidden layer, the following empirical formula is used to determine the number of hidden layer neurons:

$$l = \ +a \tag{18}$$

where I represents the number of neurons in the input layer, O signifies the number of neurons in the output layer, and a is a constant that falls within the range of $[1, 10]$.

Initialize network weights, randomly initialize w_{ih}, w_{ha}, b_h, b_a, with a range of values within the interval $[-1, 1]$, and define the error function as:

$$E = \frac{1}{2}\sum_{o=1}^{q}\left(d_o^{(k)} - y_o^{(k)}\right)^2 \tag{19}$$

Randomly select the output vector $x^{(k)}$ and the corresponding expected output $d^{(k)}$ for the k-th sample to compute the output $h_i^h(k)$ of each neuron in the hidden layer. Then, use the input and the Sigmoid activation function to determine the output $h_o^h(k)$ of the hidden layer neurons:

$$h_i^h(k) = \sum_i w_{ih}\, x_i^{(k)} - b_h \tag{20}$$

$$h_o^h(k) = f\left(h_i^h(k)\right) \tag{21}$$

$$y_i^o(k) = \sum_h w_{ho}\, h_o^h(k) - b_o \tag{22}$$

$$y_o^o(k) = y_i^o(k) \tag{23}$$

Calculate the partial derivative of the error function for each neuron in the output layer using the actual output $y_o^o(k)$ and the expected output vector $d^{(k)}$:

$$\delta_o(k) = \left(d_o(k) - y_o^o(k)\right) y_o^o(k)\left(1 - y_o^o(k)\right) \tag{24}$$

Calculate the partial derivative of the error function for each neuron in the hidden layer using the connection weights $w_{ho}(k)$, $\delta_o(k)$, and $h_o^h(k)$ from the hidden layer to the output layer:

$$\delta_h(k) = \left[\sum_{o=1}^{q}\delta_o(k)\, w_{ho}\right] h_h^o(k)\left(1 - h_h^o(k)\right) \tag{25}$$

The output layer neuron bias $\delta_o(k)$ and hidden layer neuron output $h_o^h(k)$ are used to modify the weight $w_{ho}(k)$ and threshold $b_o(k)$:

$$w_{ho}^{N+1}(k) = w_{ho}^{N}(k) - \eta\, m_t \tag{26}$$

$$b_o^{N+1}(k) = b_o^{N}(k) - \eta\, m_t \tag{27}$$

where N is before adjustment, $N+1$ is after adjustment, η is learning efficiency, with a value range of $(0, 1)$, $m_t = \frac{m_t}{1-\beta^t}$, $V_t = \frac{V_t}{1-\beta^t}$ are the first time mean and second time variance of the gradient, respectively, and $\lambda = 10^{-8}$. Adjust the connection weights and thresholds by utilizing the partial derivatives of each neuron in hidden layer $\delta_h(k)$ along with the inputs from each neuron in input layer $x_i(k)$:

$$w_{ih}^{N+1}(k) = w_{ih}^{N}(k) - \eta\, m_t\, x_i(k) \tag{28}$$

$$b_h^{N+1}(k) = b_h^{N}(k) - \eta\, m_t \tag{29}$$

Finally, calculate the global error:

$$E = \frac{1}{2}\sum_{k=1}^{m}\sum_{o=1}^{q}\left(d_o^{(k)} - y_o^{(k)}\right)^2 \tag{30}$$

Determine whether the global error fluctuates within a small range between two training sessions, that is, whether the global error remains anchored within a small range. If this condition is not met, loop the training process. Otherwise, end the current training session. And output the expected results of the dataset together with the actual predicted values. If the actual error is greater than the error function value, it indicates that the customer is a potential churn customer. Based on this judgment condition, output the potential churn customer matrix:

$$Churn\begin{bmatrix} yo_1(1) & yo_1(2) & ... & yo_1(s) \\ yo_2(1) & yo_1(2) & ... & yo_1(s) \\ ... & ... & ... & ... \\ yo_o(1) & yo_o(2) & ... & yo_o(s) \end{bmatrix} \tag{31}$$

where $s = 1, 2, \ldots, s$, $s \leq k$ represents the predicted potential number of lost customers.

4. **Experimental results and analysis.**

4.1. **Data preparation.**

The data is sourced from 2022.10 - 2023.10 Online comments from JD operator mobile card users, with comment star rating, membership level, and number of likes as input layer variables, and sentiment classification Boolean values of comment content as output layer variables. Data preprocessing includes deleting samples with missing values and removing duplicate data. Divide the modeling data into two mutually exclusive datasets with the same distribution. The output variables include comment star rating, membership level, and number of likes. Comment star rating A1 and membership level A2 are divided into 5 levels, represented by values from 1 to 5. The number of likes A3 is represented by a Boolean value, with likes being 1 and likes being 0. The output variable is represented as an emotional Boolean value, with positive being 1 and negative being 0. Determine the optimal range for the number of hidden layer nodes as [2, 12], and use the trial-and-error method to identify the best number. Start by splitting the data into training, validation, and test sets with ratios of 70%, 15%, and 15%, respectively. Validation and testing set; Secondly, a BP neural network is constructed using newf, where the transfer functions for the hidden layer and output layer are purelin and tansig, respectively; Network training and value learning are trainlm and tranlm, respectively; The maximum training frequency, initial learning rate, and training target accuracy are set to 1000 and 0, respectively 1. 0.001. From the above process, the preprocessed data and basic neural network structure are obtained. After multiple network trainings, when the input layer has 4 neurons and the hidden layer contains 10 nodes, the model achieves the best performance with a shorter training time.

4.2. **Experimental analysis.** After the experiment, 8 sets of prediction results were obtained, and the analysis of the prediction results is shown in Table 1:

Table 1. The prediction results

Number	Actual value	Predicive value	Absolute error/%
1	1	0.9999	0.01
2	1	0.7780	7.02
3	0	0.0646	6.46
4	1	0.9995	0.05
5	0	0.0621	6.61
6	0	0.0170	1.07
7	0	0.0532	5.32
8	0	0.0178	1.78

To further validate the accuracy of the method proposed in this paper, two representative algorithms, MPBP and Gray BP, were selected for comparison. The classic BP neural network algorithm was used to compare the prediction accuracy in the same sample space through simulation experiments. MPBP introduced the concept of meta path and established supervised link prediction through meta nodes. Gray BP utilizes BP neural network to compensate for the shortcomings of grey prediction models in terms of time complexity and non-linear data types. This article uses two indicators, accuracy and mean square deviation, to compare the method proposed in this article with the classical BP neural network algorithm. In the following mean square deviation formula, represents the predicted value, while represents the actual value. The comparison of accuracy and MSE line chart results between the two methods is shown in Figure 2 and Figure 3:

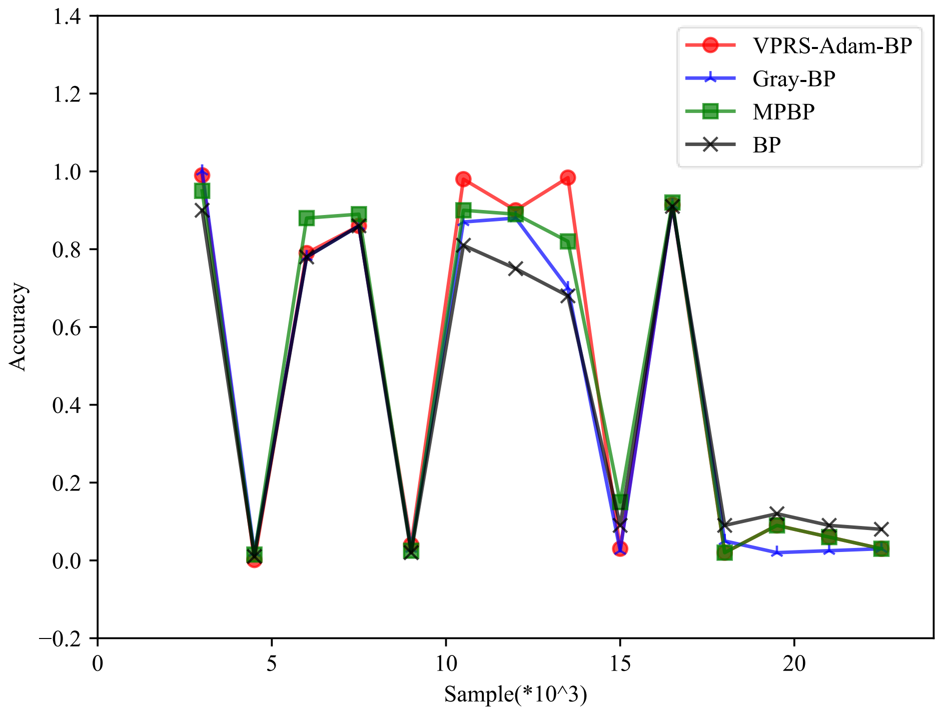

Figure 2. Line chart comparing prediction accuracy

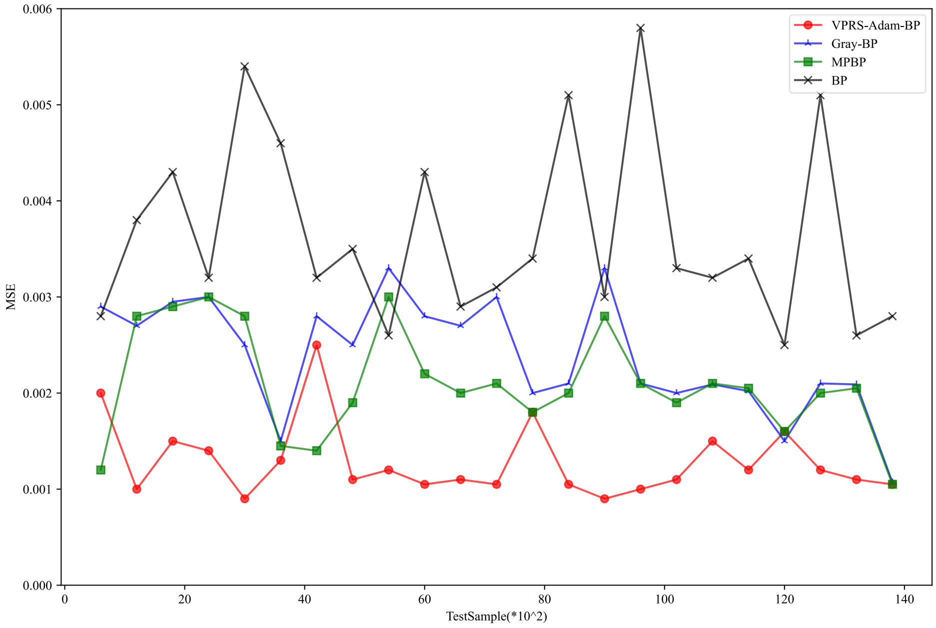

Figure 3. Comparison chart of test set mean square deviation results

Figure 2 and Figure 3 demonstrate that, in terms of overall performance for predicting large datasets, the method proposed in this paper outperforms traditional BP neural network approaches, particularly when the data volume exceeds 10,000. The accuracy of the VPRS Adam BP method is notably higher than that of the traditional BP neural network, with prediction accuracy nearing 1. At the same time, as the data volume increases, the accuracy of the VPRS Adam BP method still maintains relatively high quality prediction, which demonstrates the effectiveness of the algorithm. On the mean square error comparison chart of the test set, it can be seen that the mean square error of VPRS Adam BP is always at a low level, whether it is a small amount of data or a gradually increasing amount of data, and the line is relatively smooth, which indicates the robustness of the algorithm. In summary, this article verifies the performance of the VPRS Adam BP method in terms of accuracy and effectiveness through simulation experiments. At the same time, based on the prediction result table and simulation experiment graph, the method proposed in this article is more robust and stable in handling large amounts of dynamic data streams.

5. **Conclusion.** In this study, the risk of customer churn was predicted by analyzing the emotional tendencies of consumers in online reviews. Research differs from traditional methods of constructing churn prediction models based on user consumption behavior data, innovatively incorporating sentiment analysis of online comments as an important indicator for determining customer churn. The study uses word level sentiment analysis technology to quantify the emotional tendencies of comments into two categories: positive and negative emotions, and integrates the BP neural network to create a customer churn prediction model. The findings suggest that the proposed method demonstrates high stability and accuracy. This research offers a novel approach for businesses to anticipate customer churn.

REFERENCES

[1] B. Huang, M. T. Kechadi, and B. Buckley, "Customer churn prediction in telecommunications," *Expert Systems with Applications*, vol. 39, no. 1, pp. 1414-1425, 2012.

[2] P. Lalwani, M. K. Mishra, J. S. Chadha, and P. Sethi, "Customer churn prediction system: a machine learning approach," *Computing*, vol. 104, no. 2, pp. 271-294, 2022.
[3] C.-F. Tsai, and Y.-H. Lu, "Customer churn prediction by hybrid neural networks," *Expert Systems with Applications*, vol. 36, no. 10, pp. 12547-12553, 2009.
[4] T. Vafeiadis, K. I. Diamantaras, G. Sarigiannidis, and K. C. Chatzisavvas, "A comparison of machine learning techniques for customer churn prediction," *Simulation Modelling Practice and Theory*, vol. 55, pp. 1-9, 2015.
[5] A. C. Bahnsen, D. Aouada, and B. Ottersten, "Example-dependent cost-sensitive decision trees," *Expert Systems with Applications*, vol. 42, no. 19, pp. 6609-6619, 2015.
[6] Z. Kasiran, Z. Ibrahim, and M. S. M. Ribuan, "Customer churn prediction using recurrent neural network with reinforcement learning algorithm in mobile phone users," *International Journal of Intelligent Information Processing*, vol. 5, no. 1, 1, 2014.
[7] K. Chen, Y.-H. Hu, and Y.-C. Hsieh, "Predicting customer churn from valuable B2B customers in the logistics industry: a case study," *Information Systems and E-Business Management*, vol. 13, pp. 475-494, 2015.
[8] P. K. Sharma, S. Rathore, and J. H. Park, "Multilevel learning based modeling for link prediction and users' consumption preference in Online Social Networks," *Future Generation Computer Systems*, vol. 93, pp. 952-961, 2019.
[9] D. N. Le, H. T. Nguyen, and P. H. Truong, "Port logistics service quality and customer satisfaction: Empirical evidence from Vietnam," *The Asian Journal of Shipping and Logistics*, vol. 36, no. 2, pp. 89-103, 2020.
[10] W. Verbeke, D. Martens, and B. Baesens, "Social network analysis for customer churn prediction," *Applied Soft Computing*, vol. 14, pp. 431-446, 2014.
[11] S. Ding, C. Su, and J. Yu, "An optimizing BP neural network algorithm based on genetic algorithm," *Artificial Intelligence Review*, vol. 36, pp. 153-162, 2011.
[12] L. Zhang, F. Wang, T. Sun, and B. Xu, "A constrained optimization method based on BP neural network," *Neural Computing and Applications*, vol. 29, pp. 413-421, 2018.
[13] K. Cui, and X. Jing, "Research on prediction model of geotechnical parameters based on BP neural network," *Neural Computing and Applications*, vol. 31, pp. 8205-8215, 2019.
[14] S. Song, X. Xiong, X. Wu, and Z. Xue, "Modeling the SOFC by BP neural network algorithm," *International Journal of Hydrogen Energy*, vol. 46, no. 38, pp. 20065-20077, 2021.
[15] A. T. Goh, "Back-propagation neural networks for modeling complex systems," *Artificial Intelligence in Engineering*, vol. 9, no. 3, pp. 143-151, 1995.
[16] W. Jia, D. Zhao, T. Shen, S. Ding, Y. Zhao, and C. Hu, "An optimized classification algorithm by BP neural network based on PLS and HCA," *Applied Intelligence*, vol. 43, pp. 176-191, 2015.
[17] J.-X. Han, M.-Y. Ma, and K. Wang, "Product modeling design based on genetic algorithm and BP neural network," *Neural Computing and Applications*, vol. 33, pp. 4111-4117, 2021.
[18] Z. Xiao, S.-J. Ye, B. Zhong, and C.-X. Sun, "BP neural network with rough set for short term load forecasting," *Expert Systems with Applications*, vol. 36, no. 1, pp. 273-279, 2009.
[19] H. Yang, X. Li, W. Qiang, Y. Zhao, W. Zhang, and C. Tang, "A network traffic forecasting method based on SA optimized ARIMA–BP neural network," *Computer Networks*, vol. 193, 108102, 2021.
[20] Z. Wang, F. Wang, and S. Su, "Solar irradiance short-term prediction model based on BP neural network," *Energy Procedia*, vol. 12, pp. 488-494, 2011.
[21] Y. Ma, L. Li, Z. Yin, A. Chai, M. Li, and Z. Bi, "Research and application of network status prediction based on BP neural network for intelligent production line," *Procedia Computer Science*, vol. 183, pp. 189-196, 2021.
[22] F. Yu, and X. Xu, "A short-term load forecasting model of natural gas based on optimized genetic algorithm and improved BP neural network," *Applied Energy*, vol. 134, pp. 102-113, 2014.