

## Research Article

# Prediction of Repeat Customers on E-Commerce Platform Based on Blockchain

Huibing Zhang  and Junchao Dong 

Guangxi Key Laboratory of Trusted Software, Guilin University of Electronic Technology, Guilin 541004, China

Correspondence should be addressed to Huibing Zhang; zhanghuibing@guet.edu.cn

Received 25 May 2020; Revised 4 August 2020; Accepted 16 August 2020; Published 28 August 2020

Academic Editor: Hongju Cheng

Copyright © 2020 Huibing Zhang and Junchao Dong. 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.

In recent years, blockchain has substantially enhanced the credibility of e-commerce platforms for users. The prediction accuracy of the repeat purchase behaviour of e-commerce users directly affects the impact of precision marketing by merchants. The existing ensemble learning models have low prediction accuracy when the purchase behaviour sample is unbalanced and the information dimension of feature engineering is single. To overcome this problem, an ensemble learning prediction model based on multisource information fusion is proposed. Tests on the Tmall dataset showed that the accuracy and AUC values of the model reached 91.28% and 70.53%, respectively.

## 1. Introduction

As one of the most disruptive and innovative technologies to emerge in the fourth industrial revolution, blockchain has significant potential for the industrial transformation of e-commerce platforms. Blockchain can ensure payment security, protect users' privacy, and enhance transaction transparency of e-commerce platforms [1]. Therefore, it can improve the credibility of a platform and expand its user base. More importantly, blockchain redistributes the benefits of a centralized e-commerce platform to every merchant or seller, which effectively increases the income and improves the enthusiasm of the merchants. To attract the attention of more buyers and the rate of purchases, many merchants tend to launch large-scale promotions through e-commerce platforms on specific dates. However, the new buyers that are attracted during the promotion period are mostly one-time buyers, who do not provide long-term returns for the merchants. Therefore, repeat buyer prediction after a promotion can help an e-commerce platform carry out effective marketing to build access to long-term customers. Using historical behavioural data about buyers obtained during promotions, such as data about browsing and adding to cart and wishlist, can help e-commerce platforms understand user preferences

and shopping habits in order to achieve repeat buyer prediction.

The existing prediction models can be divided into two categories: individual models and ensemble models. Individual models, such as logistic regression (LR), support vector machine (SVM), recurrent neural network (RNN), and multilayer perception (MLP), are widely applied in repeat buyer prediction [2, 3]. As the scale of e-commerce buyers increases, the size of historical behavioural data increases rapidly. Repeat buyer individual prediction models based only on impact factors cannot achieve desired results. In recent years, many researchers have proposed repeat buyer ensemble learning prediction models, such as random forest (RF), gradient boosting decision tree (GBDT), and XGBoost, that integrate multiple individual prediction models. Results demonstrate that ensemble learning prediction models are superior to individual prediction models in terms of accuracy and robustness [4, 5]. However, the existing GDBT-based ensemble learning models perform poorly on processing the buyers' behaviour sequence data. Neural networks, represented by RNN and long short-term memory (LSTM) models, have shown notable success in sequence data modeling, which makes it possible to build a buyers' behaviour sequence prediction model based on a RNN [6]. The

individual models based on GDBT and neural networks are quite different. After fusion, they can effectively process various types of historical behaviour data about buyers. At the same time, new requirements have emerged for the fusion method of individual models in ensemble learning. To meet the new requirements, we propose a repeat buyer ensemble learning prediction model that combines the DeepCatboost, DeepGBM, and double attention BiGRU (DABiGRU) individual models using the vote-stacking method. The proposed model is better at modeling the discrete purchase records and historical behaviour sequence features, thus improving the accuracy of repeat buyer prediction.

The main contributions of this article are summarized as follows:

- (1) This paper analyzed the impact factors of repeat purchase behaviour and constructed corresponding features. The DeepCatboost and DABiGRU individual models are proposed, and the DeepGBM individual model is introduced to predict the repeat buyers
- (2) By differentiating the training data of the individual models, adding a base learning layer, and applying a majority voting decision-making mechanism to improve the test set, this paper presents an improved vote-stacking fusion approach
- (3) Experiments were conducted on the widely used the Tmall dataset. Compared with the individual and other reference models, the vote-stacking ensemble learning model achieved better results in predicting repeat buyers

The rest of this paper is organized as follows. Section 2 discusses existing studies on repeat buyer prediction methods and ensemble learning. Section 3 describes the sample balancing and feature construction and proposes the DeepCatboost and DABiGRU individual models, as well as an ensemble learning prediction model based on the vote-stacking fusion approach. The proposed method is evaluated on a real dataset in Section 4. Section 5 summarizes the paper.

## 2. Related Works

The individual prediction models predict the repeat buyers by constructing impact features and feeding them into a single machine learning or deep learning algorithm. Using of a large amount of historical behaviour data of buyers' browsing, clicking, and purchasing on e-commerce platforms, Liu and Li established a prediction model based on an SVM, which predicted the repeat purchase behaviour of buyers after promotions. They demonstrated the feasibility of using historical behaviour data to establish a prediction model and to identify the repeat buyers after promotions [7]. Tang et al. proposed a purchasing behaviour prediction framework based on an SVM model that optimized parameters and used the firefly algorithm [8]. They achieved better results than the traditional SVM model. Sakar et al. compared visitors' purchase intention prediction models based on RF, SVM, and

MLP and found that the accuracy and F1 score of the MLP model were significantly higher than those of the RF and SVM models [9]. The performance of traditional machine learning models depends heavily on feature engineering. However, the construction of features requires huge human resources, which makes traditional machine learning models not suitable for repeated buyer prediction tasks with massive user purchase behaviour data. Therefore, many repeat buyer prediction models based on deep learning algorithms have been proposed recently. For the task of customer purchase prediction in multichannel online promotions, Ling et al. used a fully connected long short-term memory network to model the interaction between customers and promotion channels, the nonlinear sequence correlation, and the cumulative effect between customers' browsing behaviour [10]. The above-mentioned neural network based models substantially reduce the dependence on feature engineering. However, there is still a gap between the ability of a neural network to handle the dense numerical features of the historical purchase behaviour data of buyers and that of the ensemble learning model based on the GDBT.

With the development of ensemble learning techniques, many researchers have introduced ensemble prediction models by combining different individual prediction models to improve the accuracy and robustness of prediction results. Common ensemble learning prediction models are mainly XGBoost [4] and RF [11, 12]. Li and Shao proposed a purchase behaviour prediction model that combined LSTM and RF [13]. The model not only extracts features from the user behaviour data and commodity attributes but also dynamically extracts features relating to the user behaviour data as a successive associated sequence by using LSTM on sequence data. Next, the model applies RF to predict user purchase behaviour. However, the dynamic feature extraction using LSTM is not accurate enough on user behaviour data, as it simply considers that each piece of historical behaviour data has equal impact on the purchase behaviour. In addition, it cannot represent the interrelation between different user behaviours (such as browsing, collecting, or purchasing). Kumar et al. adopted a hybrid method combining the artificial bee colony algorithm with machine learning techniques to predict repeat buyers. Their study found that the seller features and the buyer features were the main factors that affected the intention of repeat purchase, which inspired the feature construction of this paper [14].

The ensemble learning models can fuse a variety of prediction results from different individual models for collaborative decision-making, enabling a more accurate, stable, and robust final result. At present, the commonly used fusion strategies are the voting, averaging, and learning methods [15–18]. In recent years, many studies have shown that the performance of an ensemble model can be effectively improved by improving the strategy of fusing the individual models [19–21]. In reference [22], a stacking ensemble model using URL and HTML features was proposed to detect fishing pages in websites. This model effectively combined three individual models, GDBT, LightGBM, and XGBoost, so that the different individual models complemented each other, thus improving the accuracy of fishing web page detection.

A credit score ensemble prediction model based on a multi-stage adaptive classifier by using statistical learning and machine learning was proposed in [23]. According to the performance of the classifiers on the dataset, base classifiers were selected adaptively from the candidate classifiers set, and the parameters of the base classifiers were optimized using the Bayesian optimization algorithm. Then, the optimized base classifiers were ensemble with a multilayer stacking ensemble method to produce new features. Compared with the individual models, the ensemble models RF and AdaBoost showed better performance and data adaptability.

### 3. Repeat Buyer Prediction Model

**3.1. Model Framework.** Figure 1 demonstrates the overall framework for the proposed repeat buyer prediction model, which fuses information about buyers, sellers, and the interaction behaviour between them. Cleaning the historical behaviour data about buyers and eliminating erroneous data and missing data are implemented in the reprocessing stage, and the subtime under sampling is adopted to balance the sample [24]. The three-dimensional hidden features are constructed from the buyers, sellers, and the interaction between the buyers and sellers, respectively. The sample-balanced historical purchase behaviour data is inputted into the DABiGRU model. Next, the well-constructed features are fed to the DeepCatboost and DeepGBM individual models to predict the repeat buyers. The individual models are ensemble with the vote-stacking fusion approach, and the final prediction results are obtained.

#### 3.2. Sample Balancing and Feature Engineering

**3.2.1. Sample Balancing.** Usually, there are very few repeat buyers after a promotion, resulting in the problem that the repeat buyer sample and the one-time buyer sample are not balanced. To correct this problem, the sub-time under sampling method shown in Algorithm 1 is introduced to balance the samples. According to the time-sensitive characteristic of the purchase behaviour of buyers, the original sample of repeat buyers and one-time buyers are segmented according to the day. For each buyer in the original sample, the three nearest neighbour buyers are determined according to their Euclidean distance. If the buyer is a one-time buyer and more than two of its three nearest neighbour buyers are repeat buyers, delete it; if the sample is a repeat buyer and more than two of its three nearest neighbours are one-time buyers, remove the one-time buyers from the nearest neighbours. The rest are kept in the original buyer sample.

**3.2.2. Feature Engineering.** After analyzing the data before and after promotion, it is found that the historical interactive data of buyers is scattered in the information of buyers, sellers, and historical behaviour and so on. There are very few features in the original data that can be applied directly for repeat buyer prediction and the predictions are not ideal. There, statistical analysis and latent Dirichlet allocation (LDA), principal component analysis (PCA), and factorization machine (FM) machine learning methods were employed to build the repeat buyer features. Then, these features were

fed into the DeepCatboost and DeepGBM individual prediction models for selection and training. In addition to the basic features of buyers, sellers, and their interaction, three hidden features were constructed, and their specific meanings are as follows:

(1) *Topic Features.* Drawing on the method of LDA topic modelling of text, the seller is regarded as a document, and the IDs of all historical buyers are regarded as the words in the document [25]. Potential factors generated in the reset low-dimensional space are regarded as the topic features of sellers, and the topic features of the buyers can be obtained in the same way.

(2) *Similarity Features.* The similarity features include seller similarity and buyer similarity. The greater the number of cobuyers (similar buyers) between the sellers, the higher is the similarity of the two sellers. Before feeding into the training model, PCA is applied to the highly-sparse similarity matrix to improve the training efficiency [26].

(3) *Feedback Features.* The feedback features are defined as the inner product of the potential factors between the seller and the buyer, where the potential factors are obtained by the FM and the feedback matrix [27]. The meaning of a feedback feature is that different buyers who engage in purchasing behaviour with the same seller may have similar preferences, which denotes that if one buyer becomes a repeat buyer, another buyer is more likely to be a repeat buyer of the same seller.

**3.3. DeepCatboost Individual Model.** Catboost can automatically process the category features and use the relationship between features to enrich the original feature dimensions [28]. However, owing to the various representations of the historical behaviour data of buyers and the existence of missing data, there is a risk of overfitting in model training.

To improve the generalization ability of the Catboost model, the idea of extracting features layer-by-layer of representation learning is implemented to group and train the input data of the Catboost model, and the upper layer classification results are fed to the lower training set. The specific steps are as follows:

- (1) Multiple Catboost submodels are trained independently by randomly selecting a subset of features
- (2) The prediction results of the Catboost submodels are fused with the original features, and the output features are fed into the next layer for learning. The random noise represented by Formula (1) is introduced to mitigate the overfitting problem of the fusion process

$$F^* = F + \Theta(C_i), \quad (1)$$

where  $F^*$  denotes the historical buying feature after fusion,  $F$  stands for the original feature,  $C_i$  represents the prediction

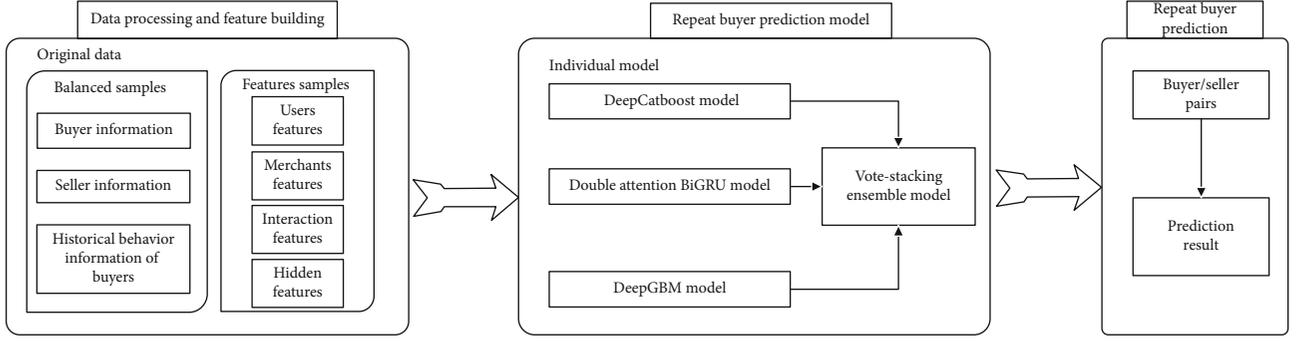
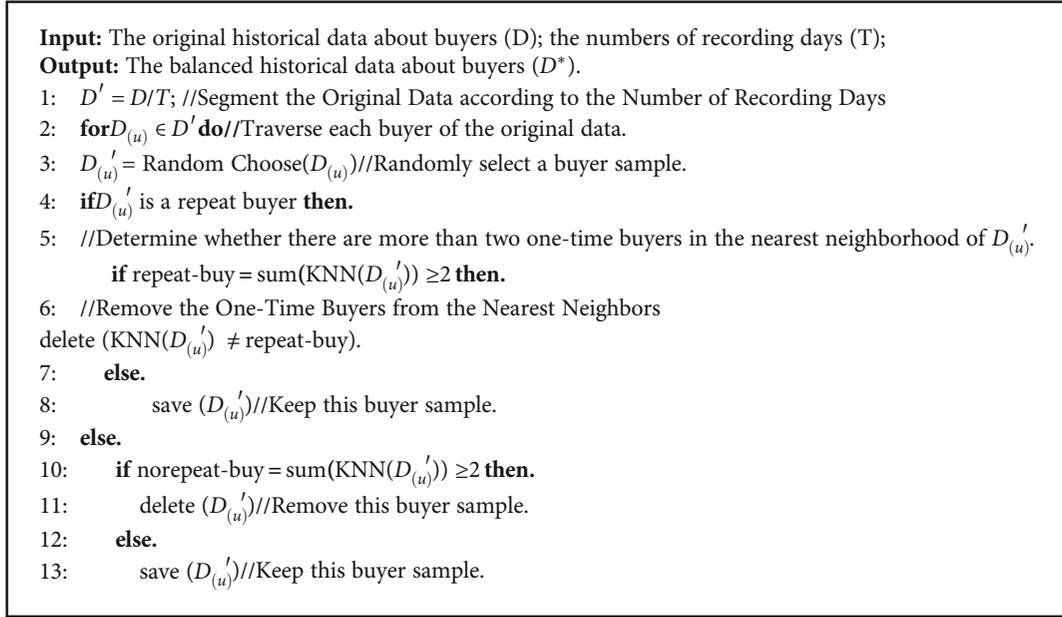


FIGURE 1: The overall framework of the proposed repeat buyer prediction model.



ALGORITHM 1: The sub-time under sampling sample balancing algorithm.

result of the Catboost submodel, and  $\Theta(\bullet)$  denotes the random noise

- (3) The features after fusion are learned in the second layer, and the prediction results of multiple Catboost submodels are fused by weighting to produce the final prediction results

The final DeepCatboost model is shown in Figure 2.

**3.4. DeepGBM Individual Model.** The DeepGBM individual model shown in Figure 3 combines the advantages of a neural network in dealing with large-scale sparse classification data and GBDT in dealing with dense numerical features. The model can produce strong classification and numerical features while maintaining the ability of efficient learning [29]. The model mainly includes two components: the GBDT2NN that focuses on the dense numerical features of the historical purchase behaviour of buyers, and the CatNN that deals with the sparse classification features of the buyer's age and sex.

To apply the DeepGBM individual model to predict repeat buyers, a GBDT model is trained using historical purchase behaviour data of buyers, and then, the DeepGBM is trained by the leaves of the GBDT tree from

$$\min_{\omega, \omega_0, \omega^T} \frac{1}{n} \sum_{i=1}^n L''(\omega^T H(\|_{t \in T} (L^{t,i})); \omega^T) + \omega_0, \sum_{t \in T} p^{t,i}, \quad (2)$$

where  $n$  stands for the number of units in the training sample,  $\|(\bullet)$  represents the connection operation, and  $G^{T,i} = (H(\|_{t \in T} (L^{t,i})); \omega^T)$  is a fully connected network that connects multiple one-hot indexes and converts them to a dense embedding vector of tree  $T$ .

The output of the DeepGBM model is expressed as

$$\hat{y}(x) = \sigma'(\omega_1 \times y_{\text{GBDT2NN}}(x) + \omega_2 \times y_{\text{Cat}}(x)), \quad (3)$$

where  $\omega_1$  and  $\omega_2$  are the training parameters of the GBDT2NN component of Formula (6) and the CatNN

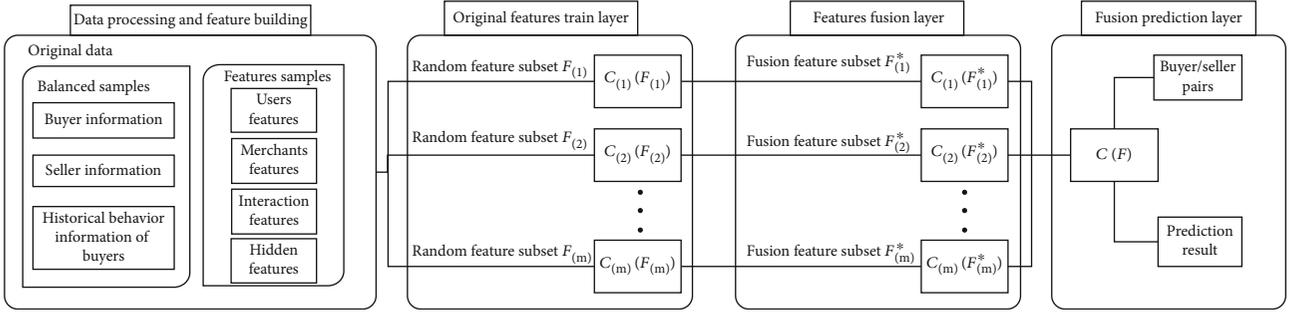


FIGURE 2: The DeepCatboost model.

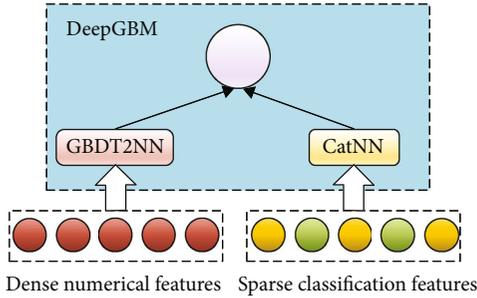


FIGURE 3: The DeepGBM framework.

component of Formula (7), respectively; and  $\sigma'$  represents the binary output transformation of whether the buyer will purchase again or not.

The model is trained using the following loss function:

$$L_{\text{DeepGBM}} = \alpha L''(\hat{y}(x), y) + \beta \sum_{j=1}^k L^{T_j}, \quad (4)$$

where  $y$  denotes a real repeat buyer,  $\hat{y}$  is the prediction result regarding a repeat buyer,  $L''$  denotes the cross-entropy loss function of the classification task whether a buyer will purchase again or not, and  $L^T$  represents the embedding loss of the tree group  $T$ . It can be inferred from Formula (5) that  $k$  is the number of tree groups and  $\alpha$  and  $\beta$  are the predefined super parameters that control the intensity of the end-to-end loss and the embedding loss, respectively.

$$L^T = \min_{\theta^T} \frac{1}{n} \sum_{i=1}^n L\left(N\left(x^i [I^T]; \theta^T\right), G^{T,i}\right), \quad (5)$$

where  $I^T$  represents the impact feature of buyer repeat purchase in the tree group  $T$ . Because of the large scale and complex structure of the historical behaviour data of buyers, the number of trees in the tree group  $T$  is large and each tree contains many features. To improve the ability of feature extraction, the top features are selected to represent all the features in the tree group according to the importance of each feature.

$$y_{\text{GBDT2NN}}(x) = \sum_{j=1}^k \omega^T \times N\left(x [I^T]; \theta^T\right) + \omega_0, \quad (6)$$

where  $k$  is the number of tree groups.

$$y_{\text{Cat}(x)} = y_{\text{FM}(x)} + y_{\text{Deep}(x)},$$

$$y_{\text{FM}(x)} = \omega_0 + \langle \omega, x \rangle + \sum_{i=1}^d \sum_{j=i+1}^d \langle E_{V_i}(x_i), E_{V_j}(x_j) \rangle x_i x_j, \quad (7)$$

where  $E_{V_i}(\bullet)$  is the embedding vector,  $d$  stands for the number of features,  $\omega_0$  and  $\omega$  are the linear part parameters, and  $\langle \bullet, \bullet \rangle$  represents the inner product.

$$y_{\text{Deep}(x)} = N\left(\left[E_{V_1}(x_1)^T, E_{V_2}(x_2)^T, \dots, E_{V_d}(x_d)^T\right]^T; \theta\right), \quad (8)$$

where  $N(x; \theta)$  denotes a multilayer neural network with input  $x$  and parameter  $\theta$ .

**3.5. DABiGRU Individual Model.** The DeepCatboost model is often inefficient in processing sparse historical behaviour data of buyers. Therefore, a DABiGRU model is proposed, which can make full use of the sparse and complex features that are automatically learned from massive data and can meet the basic requirement that individual models have substantial differences in ensemble learning.

The DABiGRU individual model is depicted in Figure 4. The model includes a feature embedding layer for encoding the original data, a bidirectional recurrent layer for modeling buyers' purchase behaviour, a double attention layer for fusing the bidirectional recurrent layer, and a classification layer for expressing the prediction results.

**3.5.1. Feature Embedding Layer.** To predict the repeat buyers, the original data of the three aspects of information from buyers, sellers, and interaction between buyers and sellers are automatically extracted by the word embedding method and the DABiGRU neural network. Firstly, the word embedding model is applied to code the information about the interaction behaviour between buyers and sellers, age and sex of the buyer, etc., and the coding length is obtained by experiments. After the word vector code is obtained, the feature submodel is used to train the interaction behaviour information between buyers and sellers to get the feature vectors.

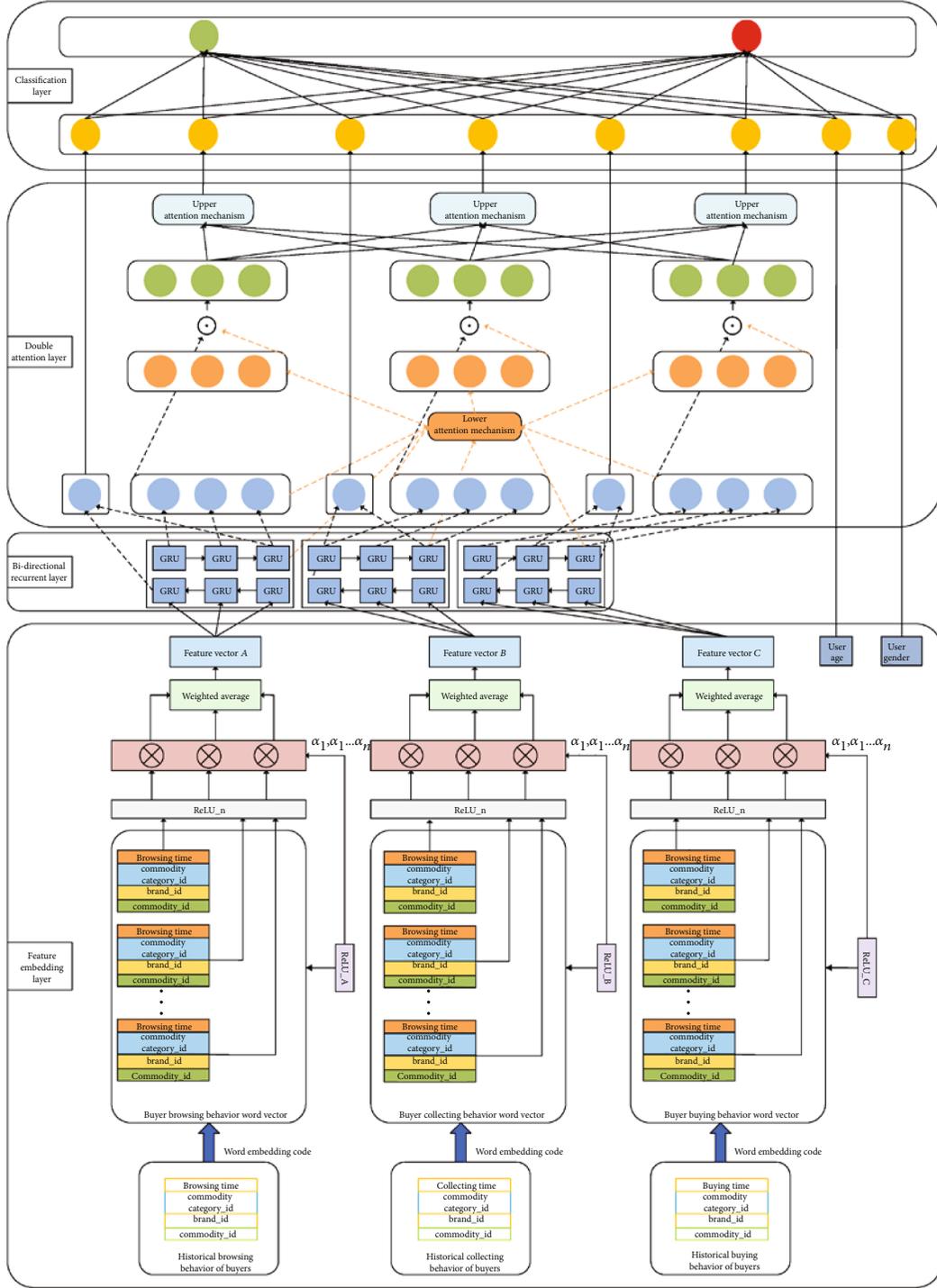


FIGURE 4: The DABiGRU individual prediction model.

As shown in the feature embedding layer of Figure 4, each record of buyers' behaviour is converted into word vectors by embedding and coding words containing the three variables commodity\_id, brand\_id, and commodity category\_id. Then, the word vectors of each buyer are transformed into an  $n$ -dimensional vector through the neural network ReLU<sub>n</sub>. The corresponding weight vector  $\alpha$  of the  $n$ -dimensional vector is obtained by ReLU<sub>M</sub>, and the final

eigenvector  $M = \{A : \text{browsing}, B : \text{purchase}, C : \text{collection}\}$  can be obtained by averaging all the  $n$ -dimensional vectors according to the weight vector  $\alpha$ .

**3.5.2. Bidirectional Recurrent Layer.** There exists a sequential relationship in the historical behaviour data of buyers. The BiGRU model is adopted to model the long-term dependency relationship [30]. Compared with the traditional

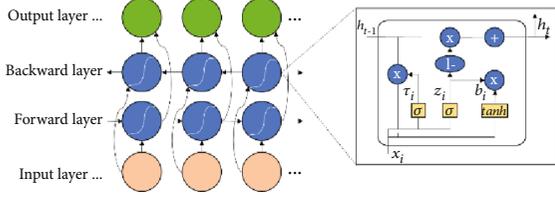


FIGURE 5: The architecture of the BiGRU model.

LSTM, this model is faster and avoids the vanishing gradient problem of the standard RNN, which is more suitable for predicting the purchasing behaviour of buyers.

The BiGRU model is a neural network that consists of a unidirectional and a heterogeneous GRU unit, as shown in Figure 5. The current hidden layer state  $h_t$  of the BiGRU model is codetermined by the current input  $x_t$ , the output  $\vec{h}_{t-1}$  of the forward hidden layer state, and the backward hidden layer state  $\overleftarrow{h}_{t-1}$ . As the BiGRU model can be regarded as two unidirectional GRU units, the hidden layer state of the BiGRU model at time  $t$  can be obtained by the weighted sum of the forward hidden layer state  $\vec{h}_{t-1}$  and the backward hidden layer state  $\overleftarrow{h}_{t-1}$ .

$$\begin{aligned}\vec{h}_t &= \text{GRU}(x_t, \vec{h}_{t-1}), \\ \overleftarrow{h}_t &= \text{GRU}(x_t, \overleftarrow{h}_{t-1}), \\ h_t &= \alpha_t \vec{h}_t + \beta_t \overleftarrow{h}_t + b_t,\end{aligned}\quad (9)$$

where the  $\text{GRU}(\bullet)$  function makes a nonlinear transformation of the inputted word vector of buyers' behaviour and converts it to the corresponding GRU hidden state,  $\alpha_t$  and  $\beta_t$  represent the weights of the forward and backward hidden states of the BiGRU at time  $t$ , respectively, and  $b_t$  is the bias term of time.

The historical purchase behaviour sequence of the buyers is expressed as  $h = \{h_1, h_2, \dots, h_n\} \in R^{n \times 2d}$ , where  $d$  indicates the dimension of the hidden state. The historical purchase behaviour sequence of the buyers includes three types: browsing, purchasing, and collecting. The corresponding hidden sequence can be obtained by inputting the corresponding coded three feature vectors into the bidirectional recurrence layer. The browsing behaviour sequence of the buyers is expressed as  $h_A = \{h_{a1}, h_{a2}, \dots, h_{an}\} \in R^{n \times 2d}$ ; the purchasing behaviour sequence of the buyers is expressed as  $h_B = \{h_{b1}, h_{b2}, \dots, h_{bn}\} \in R^{n \times 2d}$ ; and the collecting behaviour sequence of the buyers is expressed as  $h_C = \{h_{c1}, h_{c2}, \dots, h_{cn}\} \in R^{n \times 2d}$ .

**3.5.3. The Double Attention Layer.** In order to better integrate the three behaviour types of buyers' browsing, buying, and collecting, a double attention mechanism is proposed. The lower attention mechanism allocates enough attention to the key information in the behaviour sequence, and the upper attention mechanism mainly focuses on the relationship between the three behaviour sequences [31].

The lower attention mechanism achieves repeat buyer prediction by the self-attention mechanism to obtain those

behaviours that have a greater influence on repeat purchasing in each behaviour sequence. The self-attention mechanism, which usually does not require additional information, can automatically learn the weight distribution from the behaviour data of the buyers. Its formula is as follows:

$$\begin{aligned}v_t &= \tanh(W h_t + b), \\ a_t &= \frac{\exp(v_t A)}{\sum_{t=1}^T \exp(v_t A)},\end{aligned}\quad (10)$$

where  $a_t$  stands for the degree of importance of the  $t^{\text{th}}$  behaviour to the current behaviour sequence,  $v_t A$  represents the scoring system that is automatically learned from the behaviour data of buyers,  $W$  and  $A$  are the weight matrices, and  $b$  is the bias term.

The lower attention mechanism is introduced into the BiGRU model, and the input of the lower attention mechanism is the output vector of the BiGRU model. The calculation formulas of the lower attention mechanism are as follows:

$$\begin{aligned}\widetilde{h}_A &= h_A \odot a_a, \\ \widetilde{h}_B &= h_B \odot a_b, \\ \widetilde{h}_C &= h_C \odot a_c,\end{aligned}\quad (11)$$

where  $\widetilde{h}_A \in R^{n \times 2d}$ ,  $\widetilde{h}_B \in R^{n \times 2d}$ , and  $\widetilde{h}_C \in R^{n \times 2d}$  represent the lower attention mechanism output of the buyers' browsing, buying, and collecting behaviour sequences, respectively.

Different from the lower attention mechanism that focuses on the behaviour within the behaviour sequence of buyers, the upper attention mechanism pays attention to the three kinds of behaviour sequences of buyers, namely, browsing, buying, and collecting on repeat purchasing. For example, when a buyer purchases a product from a seller, the whole process may involve browsing, collecting, purchasing, etc. The upper attention model mines the interaction between different behaviour sequences to better model the buyer's purchasing behaviour. Imitating the self-attention mechanism method in Transformer [32, 33], as shown in Figure 6, the upper attention mechanism characterizes the interaction between the buyers' sequences by feeding two behaviour sequences and calculating the distances between the behaviours of the two behaviour sequences [34].

$$\begin{aligned}\text{SA}(h_A, h_B) &= A_A(h_B, h_A, h_A) \odot A_B(h_A, h_B, h_B), \\ A_A(h_B, h_A, h_A) &= \text{softmax}\left(\frac{h_B h_A^T}{\sqrt{2d}}\right) h_A, \\ A_B(h_A, h_B, h_B) &= \text{softmax}\left(\frac{h_A h_B^T}{\sqrt{2d}}\right) h_B,\end{aligned}\quad (12)$$

where  $A_A \in R^{n \times 2d}$  indicates the attention to the browsing behaviour sequence of the buyers,  $A_B \in R^{n \times 2d}$  indicates the attention to the purchasing behaviour sequence of

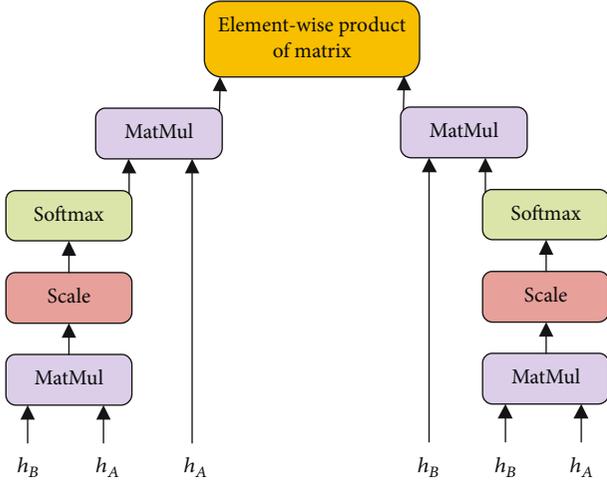


FIGURE 6: The upper attention mechanism.

the buyers,  $\odot$  stands for the element-wise product of matrix,  $SA(h_A, h_B) \in R^{n \times 2d}$  denotes the interattention to the browsing and purchasing behaviour sequences of the buyers,  $SA(h_A, h_C) \in R^{n \times 2d}$  is the interattention to the browsing and collecting behaviour sequences of the buyers, and  $SA(h_B, h_C) \in R^{n \times 2d}$  represents the interattention to the purchasing and collecting behaviour sequences of the buyers.

**3.5.4. Classification Layer.** The classification layer classifies the repeat buyers from the one-time buyers by the softmax function. The results of the bidirectional recurrent layer and the double attention layer are concatenated as the input of the classification layer [35].

$$\text{Input}_{\text{classify}} = [\text{output}_{\text{BiGRU}}, \text{output}_{\text{attention}}]. \quad (13)$$

The probability of each category  $P_c$  is calculated as

$$y = \omega_{\text{classify}} \text{input}_{\text{classify}} + b_{\text{classify}},$$

$$P_c = \frac{\exp(y_c)}{\sum_{j=1}^C \exp(y_j)}, \quad \text{for } c \text{ in } (1, 2, \dots, C), \quad (14)$$

where  $\omega_{\text{classify}}$  is the weight matrix of  $L \times C$  is the dimension of the input vector  $\text{input}_{\text{classify}}$  and  $C$  represents the number of categories of the repeat buyers and one-time buyers. After the prediction probability distribution is obtained, the cross-entropy loss function is introduced to calculate the difference between the real distribution and the predicted distribution. The parameters of the model are updated with back propagation [36, 37].

**3.6. Vote-Stacking Prediction Model.** The first layer of the different individual models of the traditional stacking models is implemented on the same training set, which makes the difference between the output values not significant, resulting in poor generalization performance. For this reason, the three-

layer vote-stacking model shown in Figure 7 is introduced. The first and second layers are both the base learning layer, and the third layer is the meta learning layer. The final prediction results are obtained by applying the majority voting decision mechanism on the outputs of the individual models and the learning layers [38].

The first base learning layer includes the three individual models of DeepCatboost, DeepGBM, and DABiGRU. The model uses nonidentical training data to further improve the difference between the output values and improve the prediction ability of the model. The historical purchase behaviour data of buyers is sensitive to time. It can be inferred from experience that the closer the predicted time is, the greater is the influence of the historical behaviour of buyers on the results. Therefore, the original data is divided into three groups according to the period, and the data in each group is randomly divided into three data clusters. One data cluster in each group is selected randomly without putting it back, and the three data clusters selected from the three groups are combined and fed into each individual model for training. The process of building the individual models  $\{y_i^1, y_i^2, y_i^3\}$  in the first base learning layer is as below.

The training set  $D$  consisting of three data clusters is inputted into the individual model  $f_k(u_i)$ ,  $k = 1, 2, 3$ ,  $i = 1, \dots, m$ , and the probability that each buyer  $u_i$  is a repeat buyer is predicted by five-fold cross-verification. The prediction results are expressed as  $y_i^k = f_k^{(j)}(u_i)$ ,  $j = 1, \dots, 5$ , where  $f_k^{(j)}$  stands for the classifier obtained from the individual model  $k$  on the  $j^{\text{th}}$  fold data subset; the value of  $y_i^k$  may be 0 (denoting one-time buyer) or 1 (denoting repeat buyers).

The test results  $\{\bar{f}_1(u), \bar{f}_2(u), \bar{f}_3(u)\}$  are obtained by inputting the test set into the trained three individual models, where  $\bar{f}_k(u)$  represents the average five-fold cross-validation results of the test samples  $u$  in each individual model.

The main difference between the second base learning layer and the first base learning layer is that the implicit relationship between the original features of the historical behaviour of buyers and the predicted probability of repeat buyers is retained in the second base learning layer. Compared with the base learning layer, the second learning layer includes an additional prediction result column for the five-fold cross-validation set and an additional test result column for the test set. In the first base learning layer, the prediction results  $\{y_i^1, y_i^2, y_i^3\}$  of the five-fold cross-validation set of the three individual models are added to the features  $x_i$  of the original training set as the features of the new training set  $x_i^* = \{x_i, y_i^1, y_i^2, y_i^3\}$ . The test set prediction result  $\{\bar{f}_1(u), \bar{f}_2(u), \bar{f}_3(u)\}$  is combined with the initial test set feature  $x$  as the new features  $x' = \{x, \bar{f}_1(u), \bar{f}_2(u), \bar{f}_3(u)\}$  of the test set.

The new training set is divided into five nonoverlapping subsets. One of the five subsets is chosen as the test set and the other four subsets as the training sets to train and test the individual models DeepCatboost, DeepGBM, and DABiGRU. The above process is repeated until all the five subsets are utilized as the test set for one time each. The prediction results are saved as  $B_{k,\text{train}} = (B_{k,1}, B_{k,2}, B_{k,3}, B_{k,4}, B_{k,5})^T$ . In the process of establishing the individual models, each model

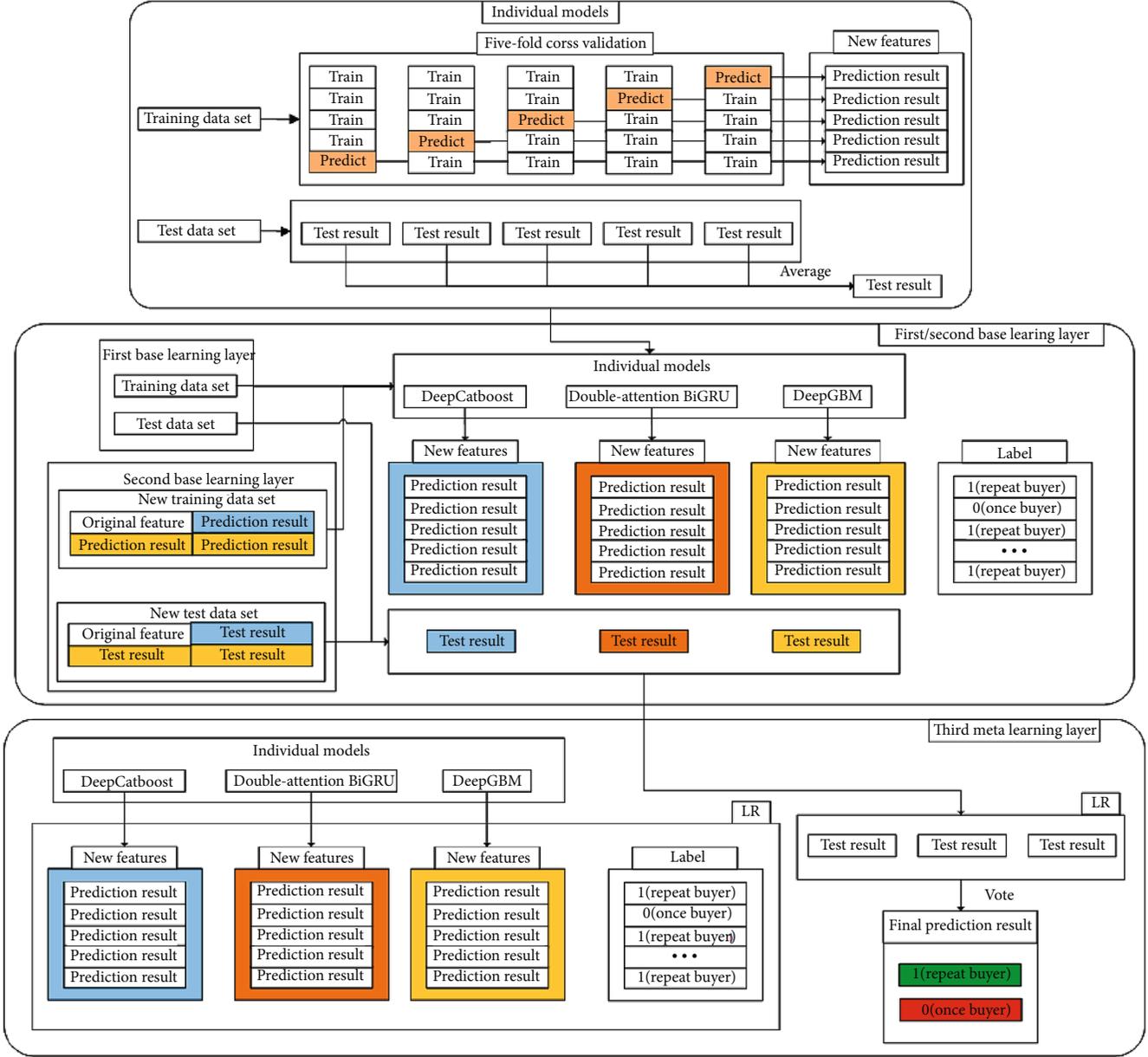


FIGURE 7: Ensemble learning prediction model based on the vote-stacking.

is tested five times on the test dataset, and the average value is taken as the test results  $B_{k,\text{test}} = (b_k)^T$ .

The third meta learning layer can effectively integrate the advantages of the three individual models and improve accuracy and stability. LR has a strong generalization ability and can reduce the overfitting risk of vote-stacking, so the meta learning layer is modeled by LR [39].

#### 4. Experiment and Analysis

**4.1. Experimental Data and Evaluation Indicators.** The experimental data contains approximately 260,000 anonymous buyers' shopping information six months before and after the "Double Eleven Shopping Festival" (<https://tianchi.aliyun.com/competition/entrance/231674>).

It is mainly composed of three tables of information relating to the buyers: the personal information table, the buyer behaviour log table, and the purchase behaviour table of the buyer. The ratio of the positive sample of repeat buyers to the negative sample of one-time buyers is about 1:10. The sample balancing process is carried out by the subtime under sampling method. Whether a buyer becomes a repeat buyer of a given merchant six months after the promotion is predicted by using the historical data information.

For the repeat buyer prediction problem, the sample category combination of the real category and the predicted category can be divided into four categories: true positive (TP), false positive (FP), true negative (TN), and false negative

(FN). After the digital confusion matrix is obtained, the accuracy and area under the curve (AUC) values can be calculated as

$$\text{Accuracy} = \frac{TP + TN}{P + N}, \quad (15)$$

where  $P$  is the total number of repeat buyers for positive samples and  $N$  is the total number of one-time buyers of negative samples.

$$\text{AUC} = \frac{\int_0^1 TPd_{FP}}{(TP + FN)(TN + FP)}. \quad (16)$$

## 4.2. Experimental Results and Analysis

**4.2.1. Word Vector Encoding Length Selection.** The encoding length of the word vector of the feature embedding layer in the DABiGRU individual model is obtained by experiments. The encoding length is positively related to the size of the coding objects; and the higher the total number of words, the longer is the encoding length. The ratio of the total number of words about commodities, brands, and commodity categories is obtained by statistics and is about 625:5:1. The experiments are carried out based on the five groups of word vector length combination shown in Table 1. As the total number of words about age and sex of the buyer information is small, the words vector length is set as 10.

Experimental results show that the AUC value improves with the increment of the word vector encoding length. The main reason is that the longer is the encoding length, the stronger is the representation ability of the code, and the overall representation ability of the DABiGRU individual model is improved accordingly. However, the longer is the word vector coding length, the more system resources are consumed by the DABiGRU individual model. Considering the limitations of the experimental environment, the word vector encoding length combination of [150, 80, 80, 10, 10] was selected for the following experiments.

**4.2.2. Impact Feature Analysis of Repeat Purchase by Buyers.** To analyze the impact features of the repeat purchases by buyers, Table 2 presents the top 10 impact features of the DeepCatboost individual prediction model.

The features listed in Table 2 show that the interaction feature has the strongest impact on repeat purchase, followed by the seller features and the buyer features. The result on the interaction between the buyer and seller demonstrates that feature (1) represents the degree of “preference” of a buyer for a seller. The actual situation is that buyers are often accustomed to buying goods from the sellers they have purchased from previously. Features (5) and (6) measure the length of time a buyer “stays” with a given seller. According to experience, the longer a buyer “stays” in a certain merchant, the more likely it is that they will buy. From the buyer’s point of view, it can be inferred from feature (1) that a buyer’s promotion intensity has the greatest impact on predicting repeat purchase. And from features (2) and (4), we can conclude that the buyer purchase conversion rate is also an important

TABLE 1: The AUC values of different combinations of word vector encoding length.

Word vector encoding length	Commodity	Brand	Commodity category	AUC value
1	64	32	32	0.6200
2	128	64	64	0.6316
3	150	80	80	0.6419
4	200	100	100	0.6543
5	256	128	128	0.6628

impact factor for repeat purchase prediction. From the seller’s point of view, the buyer’s age in feature (7) and gender in feature (10) also show a certain impact on repeat buyer prediction.

**4.2.3. Improved Individual Model and Vote-Stacking Ensemble Model Test.** Figure 8 demonstrates the difference in accuracy and AUC values before and after the improvement of the DeepCatboost and DABiGRU individual prediction models.

It can be observed from the experimental results shown in Figure 8 that the accuracy and AUC values of the improved DeepCatboost and DABiGRU individual models are superior to the original Catboost and LSTM models on both the training set and the test set, which indicates that the improved DeepCatboost individual model can effectively reduce the risk of overfitting and improve accuracy and robustness of predictions. The DABiGRU model outperforms the original LSTM neural network model by introducing the attention mechanism and improving the neural cell structure. This is because the added double attention mechanism can effectively mine and pay sufficient attention to the key information within and between the sequences of the historical purchase behaviour of buyers. Compared with the directional LSTM neural network model, the output state of the BiGRU neural network model is codetermined by several previous inputs and subsequential inputs, and the prediction result is more accurate. In addition, experimental results show that the accuracy and AUC values of the proposed model on the training set and the test set are not very different, and there are no overfitting or underfitting problems.

Table 3 shows the experimental results of the accuracy and AUC values of five ensemble learning prediction models. In order to evaluate the experimental results objectively and accurately, each model was executed 10 times on the same training and test sets, and the average and variance of the accuracy and AUC values were considered the final results.

It can be observed from Table 3 that the average values of the accuracy and AUC of the vote-stacking ensemble model are 91.28% and 70.50%, which are higher than those of the other four ensemble models for both the training set and the test set. This demonstrates the effectiveness of the improvement of the vote-stacking ensemble prediction model on selecting individual models and stacking fusing strategies. The main reasons are as follows: compared with the XGBoost and LightGBM ensemble learning models based on GBDT, the proposed ensemble model can not only

TABLE 2: The top 10 impact features of the DeepCatboost individual model.

Ranking of features	Feature detail
1	The ratio of shopping times during promotion to nonpromotion days
2	The ratio of purchasing times to browsing
3	The total historical shopping times of a buyer from some seller
4	The ratio of purchasing times to collecting
5	The time interval between the first purchase and the last purchase of a buyer from a seller
6	The interaction frequency between a buyer and a seller measured in months
7	The total number of goods purchased by different age groups of buyers from a seller
8	The total number of buyers browsing some seller
9	The total number of bands purchased by a buyer from a seller
10	The total number of goods purchased by different sexes of buyers from a seller

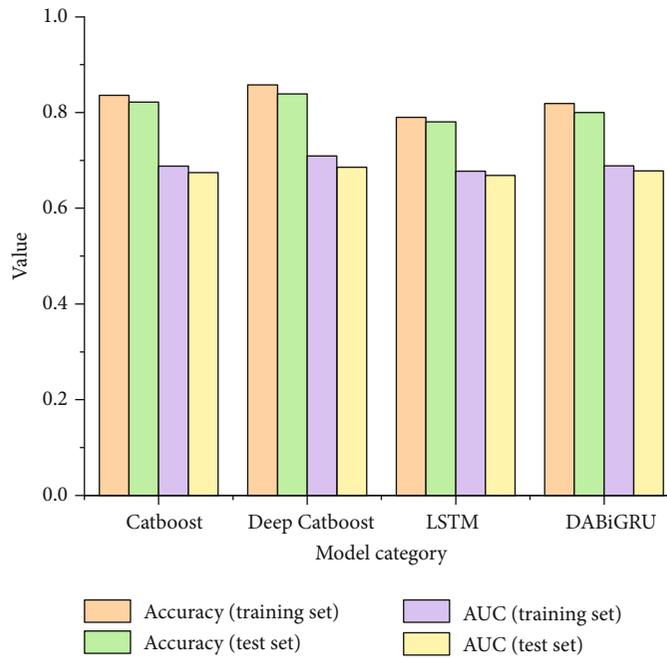


FIGURE 8: Comparison of the accuracy and AUC values of the individual models before and after improvement.

TABLE 3: The accuracy and AUC values of five types of ensemble prediction models.

Model	Training set				Test set			
	Accuracy		AUC		Accuracy		AUC	
	Mean	Variance	Mean	Variance	Mean	Variance	Mean	Variance
XGBoost	0.7762	0.0412	0.6752	0.0408	0.7751	0.0396	0.6744	0.0419
LightGBM	0.7930	0.0398	0.6878	0.0406	0.7908	0.0390	0.6875	0.0404
DeepGBM	0.8664	0.0361	0.7022	0.0356	0.8576	0.0371	0.6991	0.0366
Stacking	0.9080	0.0336	0.7027	0.0319	0.8943	0.0338	0.7023	0.0339
Vote-stacking	0.9128	0.0291	0.7053	0.0302	0.9014	0.0334	0.7050	0.0302

effectively deal with the dense numerical features in the historical behaviour data of buyers but also can process large-scale sparse classification features efficiently, which further enriches the feature types and improves the performance of the model. The nonidentical training data is employed in the vote-stacking model to increase the difference between

the output values. The features of the original data and the new features generated by the base learning layer are effectively combined to enrich the feature types of the model. The final test results are determined by the voting mechanism to effectively combine the test results of the base learning layer and the meta learning layer, which reduces the error

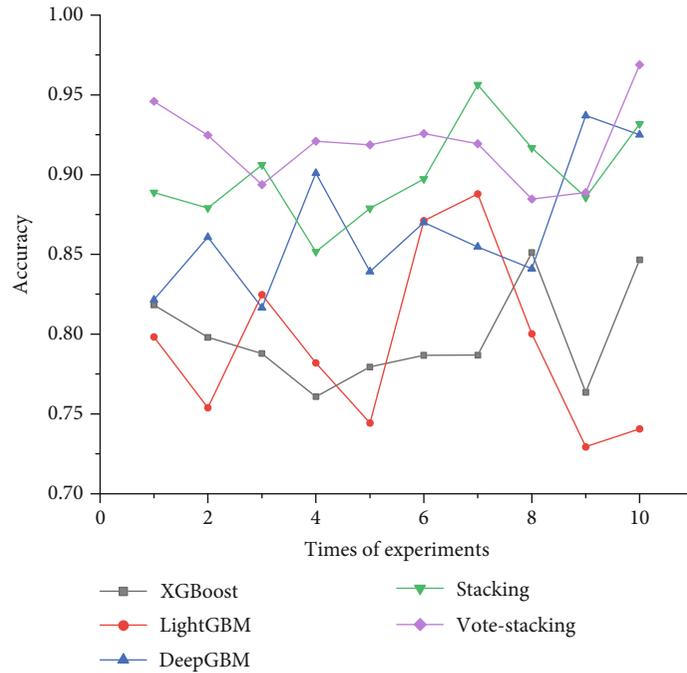


FIGURE 9: The accuracy curve of the five models.

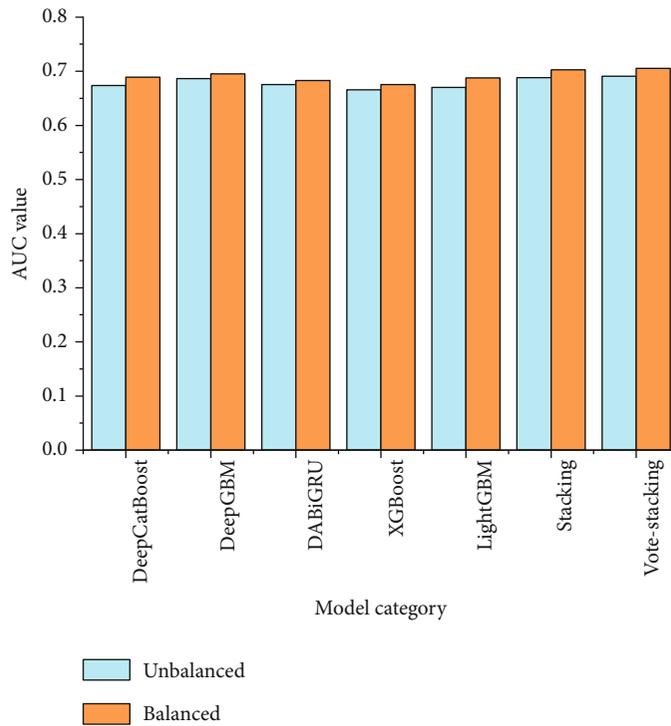


FIGURE 10: Comparison of the AUC values before and after sample balancing.

of the traditional stacking algorithm that relies only on the meta learning layer. In addition, the difference of accuracy and AUC values between the training set and test set is small, indicating that the model has better generalization ability.

In order to verify the stability of the vote-stacking model more intuitively, the accuracy values of the five models on 10 training sets are presented in a line graph as shown in Figure 9.

TABLE 4: Comparison of the accuracy and AUC values of six prediction models.

Model	Accuracy	AUC
SVM	0.6803	0.6245
RF	0.7045	0.6537
TMFBG	0.7028	0.6642
NN	0.7087	0.6951
Blending	0.8784	0.7023
Proposed model	0.9128	0.7053

It can be seen from Figure 9 that compared with the individual models, the vote-stacking and stacking models have less fluctuation in the interval with higher accuracy and their overall stability is better. Therefore, the vote-stacking fusion method is more suitable than the traditional stacking method for the repeat buyer prediction task where the dataset difference is significant. By combining the prediction results of the three individual models, the ensemble learning model can effectively avoid the selection of any individual model with a poor prediction result. By improving the performance of each individual model and collaborating between the individual models, an ensemble model can improve the overall accuracy and stability of the predictions.

It can be observed from Figure 10 that the AUC values of the seven models are improved by different extents after subtime under sampling sample balancing. Among them, the AUC values of the DeepCatboost, LightGBM, vote-stacking, and stacking models are enhanced more obviously. Experimental results show that, when the subtime under sampling sample balancing is used to process the historical behaviour data of buyers, even if the negative sample data is reduced leading to the loss of some information, the prediction accuracy of the model is improved.

In order to verify the effectiveness and functionality of the vote-stacking model, it is compared with the five other prediction models, namely, SVM [8], RF [11], TMFBG [4], neural network (NN) [10], and blending [5] with the same datasets and experimental settings. The results are shown in Table 4.

The proposed model outperforms the other five models in terms of the accuracy and AUC values. The reason is that the proposed model can effectively fuse the advantages of the DT and NN models on random and nonlinear training data and efficiently process high-dimensional nonlinear historical behaviour data of buyers. It also has better generalization ability than the reference models.

## 5. Conclusions

A repeat buyer prediction ensemble model based on the vote-stacking fusion method was introduced in this paper. First, the model applies the subtime under sampling sample balancing to process the unbalanced historical behaviour data of buyers. Then, the three individual models DeepCatboost, DeepGBM, and DABiGRU are constructed consecutively. Finally, the vote-stacking fusion method is used to fuse the prediction results of the three individual models and obtain

the final prediction results. The effectiveness of the proposed method has been evaluated on the Tmall real dataset. The experimental results showed that the proposed model is superior to the reference models. In addition, the DeepCatboost individual model was used to analyze the important impact features related to buyers' repeat purchase behaviour and to provide data support for sellers to carry out accurate marketing and increase buyers' loyalty.

In the future, we will continue to investigate the impact factors associated with buyers' repeat purchase behaviour, construct additional fine-tuned features, and design ensemble prediction models with a higher number of individual models that are also more different. In addition, we will attempt to derive better ensemble learning fusion strategies to further improve the performance of repeat buyer prediction.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

## Acknowledgments

This work was supported by the National Natural Science Foundation of China (grant numbers 61662013, U171123, U1811264, and 61662015); Guangxi Innovation-Driven Development Project (Science and Technology Major Project) (grant number AA17202024); and Graduate Student Innovation Program, Guilin University of Electronic Technology (grant number 2019YCXS045). We thank LetPub (<https://www.letpub.com>) for its linguistic assistance during the preparation of this manuscript.

## References

- [1] H. Cheng, Z. Su, N. Xiong, and Y. Xiao, "Energy-efficient node scheduling algorithms for wireless sensor networks using Markov Random Field model," *Information Sciences*, vol. 329, pp. 461–477, 2016.
- [2] A. H. Kazmi, G. Shroff, and P. Agarwal, "Generic Framework to Predict Repeat Behavior of Customers Using Their Transaction History," in *2016 IEEE/WIC/ACM International Conference on Web Intelligence (WI)*, pp. 449–452, Omaha, NE, USA, 2016.
- [3] H. S. Song, "Comparison of Performance between MLP and RNN Model to Predict Purchase Timing for Repurchase Product," *Journal of Information Technology Applications and Management*, vol. 24, no. 1, pp. 111–128, 2017.
- [4] G. Liu, T. T. Nguyen, G. Zhao et al., "Repeat Buyer Prediction for E-Commerce," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 155–164, San Francisco, CA, USA, 2016.
- [5] D. Xu, W. Yang, and L. Ma, "Repurchase Prediction Based on Ensemble Learning," in *2018 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing,*

- Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI)*, pp. 1317–1322, Guangzhou, China, 2018.
- [6] L. Guo, L. Hua, R. Jia, B. Zhao, X. Wang, and B. Cui, “Buying or Browsing?: Predicting Real-time Purchasing Intent using Attention-based Deep Network with Multiple Behavior,” in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp. 1984–1992, Anchorage, AK, USA, 2019.
  - [7] X. Liu and J. Li, “Using support vector machine for online purchase predication,” in *2016 International Conference on Logistics, Informatics and Service Sciences (LISS)*, pp. 1–6, Sydney, NSW, Australia, 2016.
  - [8] L. Tang, A. Wang, Z. Xu, and J. Li, “Online-purchasing behavior forecasting with a firefly algorithm-based SVM model considering shopping cart use,” *Eurasia Journal of Mathematics, Science and Technology Education*, vol. 13, no. 12, pp. 7967–7983, 2017.
  - [9] C. O. Sakar, S. O. Polat, M. Katircioglu, and Y. Kastro, “Real-time prediction of online shoppers’ purchasing intention using multilayer perceptron and LSTM recurrent neural networks,” *Neural Computing and Applications*, vol. 31, no. 10, pp. 6893–6908, 2019.
  - [10] C. Ling, T. Zhang, and Y. Chen, “Customer Purchase Intent Prediction Under Online Multi-Channel Promotion: A Feature-Combined Deep Learning Framework,” *IEEE Access*, vol. 7, pp. 112963–112976, 2019.
  - [11] H. Valecha, A. Varma, I. Khare, A. Sachdeva, and M. Goyal, “Prediction of Consumer Behaviour using Random Forest Algorithm,” in *2018 5th IEEE Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON)*, pp. 1–6, Gorakhpur, India, 2018.
  - [12] R. Joshi, R. Gupte, and P. Saravanan, “A Random Forest Approach for Predicting Online Buying Behavior of Indian Customers,” *Theoretical Economics Letters*, vol. 8, no. 3, pp. 448–475, 2018.
  - [13] X. LI and F. Shao, “Purchase Behaviour Forecasting Model Combining LSTM and Random Forest,” *Journal of Qingdao University (Engineering & Technology Edition)*, vol. 32, no. 2, pp. 17–20, 2018.
  - [14] A. Kumar, G. Kabra, E. K. Mussada, M. K. Dash, and P. S. Rana, “Combined artificial bee colony algorithm and machine learning techniques for prediction of online consumer repurchase intention,” *Neural Computing and Applications*, vol. 31, no. S2, pp. 877–890, 2019.
  - [15] H. Liu, Y. Xu, and C. Chen, “Improved pollution forecasting hybrid algorithms based on the ensemble method,” *Applied Mathematical Modelling*, vol. 73, pp. 473–486, 2019.
  - [16] W. Li, P. Ding, X. Zhang et al., “Ensemble learning methodologies to improve core power distribution abnormal detectability,” *Nuclear Engineering and Design*, vol. 351, pp. 160–166, 2019.
  - [17] Y. Wang, D. Wang, N. Geng, Y. Wang, Y. Yin, and Y. Jin, “Stacking-based ensemble learning of decision trees for interpretable prostate cancer detection,” *Applied Soft Computing*, vol. 77, pp. 188–204, 2019.
  - [18] S. Y. Yerima and S. Sezer, “DroidFusion: A Novel Multilevel Classifier Fusion Approach for Android Malware Detection,” *IEEE Transactions on Cybernetics*, vol. 49, no. 2, pp. 453–466, 2019.
  - [19] O. Sagi and L. Rokach, “Ensemble learning: A survey,” *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 8, no. 4, article e1249, 2018.
  - [20] B. Yadav, P. K. Gupta, N. Patidar, and S. K. Himanshu, “Ensemble modelling framework for groundwater level prediction in urban areas of India,” *Science of the Total Environment*, vol. 712, article 135539, 2020.
  - [21] H. Liu and C. Chen, “Prediction of outdoor PM<sub>2.5</sub> concentrations based on a three-stage hybrid neural network model,” *Atmospheric Pollution Research*, vol. 11, no. 3, pp. 469–481, 2020.
  - [22] Y. Li, Z. Yang, X. Chen, H. Yuan, and W. Liu, “A stacking model using URL and HTML features for phishing webpage detection,” *Future Generation Computer Systems*, vol. 94, pp. 27–39, 2019.
  - [23] S. Guo, H. He, and X. Huang, “A Multi-Stage Self-Adaptive Classifier Ensemble Model With Application in Credit Scoring,” *IEEE Access*, vol. 7, pp. 78549–78559, 2019.
  - [24] H. Cheng, D. Feng, X. Shi, and C. Chen, “Data quality analysis and cleaning strategy for wireless sensor networks,” *EURASIP Journal on Wireless Communications and Networking*, vol. 2018, article 61, 2018.
  - [25] H. Jelodar, Y. Wang, C. Yuan et al., “Latent Dirichlet allocation (LDA) and topic modeling: models, applications, a survey,” *Multimedia Tools and Applications*, vol. 78, no. 11, pp. 15169–15211, 2019.
  - [26] Y. Ait-Sahalia and D. Xiu, “Principal Component Analysis of High-Frequency Data,” *Journal of the American Statistical Association*, vol. 114, no. 525, pp. 287–303, 2019.
  - [27] B. Zhu, F. Ortega, J. Bobadilla, and A. Gutiérrez, “Assigning reliability values to recommendations using matrix factorization,” *Journal of Computational Science*, vol. 26, pp. 165–177, 2018.
  - [28] L. Prokhorenkova, G. Gusev, A. Vorobev, A. V. Dorogush, and A. Gulin, “CatBoost: unbiased boosting with categorical features,” in *Proceedings of the 32nd International Conference on Neural Information Processing Systems (NIPS’18)*, pp. 6638–6648, Red Hook, NY, USA, 2018.
  - [29] G. Ke, Z. Xu, J. Zhang, J. Bian, and T.-Y. Liu, “DeepGBM: A Deep Learning Framework Distilled by GBDT for Online Prediction Tasks,” in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp. 384–394, New York, NY, USA, 2019.
  - [30] T. Shen, T. Zhou, G. Long, J. Jiang, and C. Zhang, “Bi-directional block self-attention for fast and memory-efficient sequence modeling,” 2018, <https://arxiv.org/abs/1804.00857>.
  - [31] Y. Zhu, W. Zhang, Y. Chen, and H. Gao, “A novel approach to workload prediction using attention-based LSTM encoder-decoder network in cloud environment,” *EURASIP Journal on Wireless Communications and Networking*, vol. 2019, article 274, 2019.
  - [32] C. Park, D. Kim, and H. Yu, “An encoder–decoder switch network for purchase prediction,” *Knowledge-Based Systems*, vol. 185, article 104932, 2019.
  - [33] A. Vaswani, N. Shazeer, N. Parmar et al., “Attention is all you need,” in *Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017*, pp. 5998–6008, Long Beach, CA, USA, 2017.
  - [34] B. Pavlyshenko, “Using Stacking Approaches for Machine Learning Models,” in *2018 IEEE Second International Conference on Data Stream Mining & Processing (DSMP)*, pp. 255–258, Lviv, Ukraine, 2018.

- [35] H. Liang, J. Zou, K. Zuo, and M. J. Khan, "An improved genetic algorithm optimization fuzzy controller applied to the well-head back pressure control system," *Mechanical Systems and Signal Processing*, vol. 142, article 106708, 2020.
- [36] F. Hu and G. Wu, "Distributed Error Correction of EKF Algorithm in Multi-Sensor Fusion Localization Model," *IEEE Access*, vol. 8, pp. 93211–93218, 2020.
- [37] H. Cheng, N. Xiong, A. V. Vasilakos, L. Tianruo Yang, G. Chen, and X. Zhuang, "Nodes organization for channel assignment with topology preservation in multi-radio wireless mesh networks," *Ad Hoc Networks*, vol. 10, no. 5, pp. 760–773, 2012.
- [38] J. Hu, Y. Sun, G. Li, G. Jiang, and B. Tao, "Probability analysis for grasp planning facing the field of medical robotics," *Measurement*, vol. 141, pp. 227–234, 2019.
- [39] H. Liang, J. Zou, Z. Li, M. J. Khan, and Y. Lu, "Dynamic evaluation of drilling leakage risk based on fuzzy theory and PSO-SVR algorithm," *Future Generation Computer Systems*, vol. 95, pp. 454–466, 2019.