
Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org
© 2025 SCPE. Volume 26, Issues 4, pp. 1831–1840, DOI 10.12694/scpe.v26i4.4641

CONSUMER PURCHASE BEHAVIOR PREDICTION ON E-COMMERCE PLATFORMS BASED ON MACHINE LEARNING FUSION ALGORITHM

YIBO HU*, RONG FU†, AND WENBO NIU‡

Abstract. To enhance the precision of predicting consumer purchasing behavior, the author conducts a study focused on forecasting buying patterns on e-commerce platforms through the use of machine learning fusion techniques. The research specifically integrates logistic regression and support vector machine algorithms to analyze shopping behavior data from Alibaba's e-commerce platform. The experiment revealed that, out of 1,445 test samples fed into the model, 571 were predicted to exhibit purchasing activity on the 32nd day, as indicated by a prediction outcome of "1." Compared with the samples with actual purchasing behavior on the 32nd day, their F1 score was 7.77%. The practical results show that the fused model is more accurate in prediction performance than a single algorithm model.

Key words: Machine learning, Fusion algorithm, purchasing behavior

1. Introduction. As mobile device costs and internet fees continue to decrease, a growing number of people are adopting smartphones, sparking a surge in online shopping activity. This trend has led to a significant rise in the number of individuals engaging in e-commerce. Data from the China E-commerce Research Center indicates that online retail sales in the first half of 2017 surpassed the total sales figures of 2016. With e-commerce expanding rapidly, consumers are becoming increasingly demanding, expecting higher standards from businesses in terms of brands, products, pricing, and services [1]. In addition, some giant enterprises have emerged in China's e-commerce market, such as Tmall, Taobao, and JD.com. In recent years, the government has successively introduced relevant policies to encourage young people to start businesses, and some capable entrepreneurs have also developed through this opportunity. These startups face challenges in breaking into the established e-commerce market. Recently, leading e-commerce platforms have responded by introducing innovations that focus on targeting specific consumer segments and catering to their unique needs, with platforms like Pinduoduo Shopping Mall being prime examples [2]. This platform is essentially a group buying platform, where consumers attract friends to bargain or initiate group buying to purchase goods at a low price. But at the same time, some problems have arisen, such as the inability to guarantee the quality of goods while lowering their prices, leading to consumer loss. In today's competitive e-commerce landscape, those who can accurately identify and cater to the specific needs of consumers, anticipate their preferences, and deliver superior services will be the ones to secure a strong position in the market [3].

In practical terms, with the fast-paced advancements in big data and cloud computing, effectively utilizing analytical and predictive techniques enables the prediction and understanding of consumer purchasing habits, intentions, and behavior patterns by examining both explicit and implicit feedback, such as purchase histories. This approach not only enhances the shopping experience for consumers but also allows businesses to transition from passively displaying products to actively recommending them, addressing specific consumer needs, attracting more potential buyers, and boosting conversion rates [4]. From a technological development standpoint, consumers generate vast amounts of online behavior data that is highly granular and multidimensional. Some e-commerce platforms, both domestic and international, offer anonymized data to researchers, enabling further exploration in machine learning and artificial intelligence, which holds considerable academic value. When consumers interact with e-commerce platforms and their intentions are clear, text analysis through search engines

*The School of Business, Xi'an International University, Xi'an, Shaanxi, 710077, China. (Corresponding author, YiboHu7@163.com)
†The School of Business, Xi'an International University, Xi'an, Shaanxi, 710077, China. (RongFu8@126.com)
‡The School of Business, Xi'an International University, Xi'an, Shaanxi, 710077, China. (WenboNiu9@163.com)

can effectively identify their needs. However, when consumer intentions are ambiguous and cannot be easily articulated, search engine analysis becomes less effective. At this point, using implicit feedback data generated by consumers to predict their next behavior and understand their purchasing preferences for a certain product becomes a better choice. However, due to concerns around data privacy and information protection, certain data may be encrypted and challenging to access. As a result, the challenge the author seeks to address is how to predict consumer purchasing behavior using consumer behavior data specifically by leveraging implicit feedback without needing to access detailed personal information [5].

2. Literature Review. Based on the review of previous related research, scholars' research on predicting consumer behavior in e-commerce can be divided into two categories. One is the study of recommendation systems, which mines the products that consumers are interested in by obtaining consumer behavior data from e-commerce platforms, thereby identifying the types of products that consumers are likely to be interested in and proactively presenting them with tailored product recommendations. Another approach involves utilizing various machine learning and data mining techniques to analyze consumer online behavior data and forecast purchasing patterns. For example, Pandi, V. et al. conducted sentiment analysis on Twitter to gain insights into customer purchasing behavior. E-commerce has grown significantly, especially for people who buy products on the Internet. The results prove that it is truly used to predict using the most accurate analysis model for checking customer conclusions/emotions. Analyze the accuracy of each machine learning algorithm and consider the most precise calculation as ideal [6]. Nakao, J. et al. colleagues created a production planning model that integrates customer preferences into its framework. This model generates insights into customer purchasing behavior using machine learning techniques. By clustering data based on purchase information from various customers, the model effectively segments the customer base. The findings indicate that, in contrast to models that do not incorporate customer purchase data, the proposed approach delivers improved customer satisfaction and increased profitability [7]. Prakash, R. et al. set out to explore the purchasing patterns of e-commerce customers to enhance service and product offerings. By analyzing customer data, they aim to identify product preferences based on prior transaction records. They employed a range of techniques, including logistic regression, K-nearest neighbors, support vector machines, random forests, recurrent neural networks, and long short-term memory networks. The performance of these methods was evaluated to predict which products customers are most likely to buy [8].

Based on this research, the author proposes a study on predicting consumer purchasing behavior on e-commerce platforms using machine learning fusion algorithms. By collecting and analyzing massive shopping behavior data on the Alibaba platform, these data are transformed into valuable information, providing valuable insights for e-commerce companies to optimize product recommendation systems, enhance user experience, and ultimately drive sales growth. Through this deep fusion method, the author not only demonstrates the effectiveness of machine learning fusion algorithms in the e-commerce field, but also provides new ideas and directions for future related research and practice.

3. Research Methods.

3.1. Prediction Algorithm Model.

3.1.1. Overall framework of the model. From a qualitative standpoint, consumer interest in products can be ranked as follows: click (pv) < like (fav) < add to cart < buy. Additionally, both the sequence and timing of consumer interactions with a product influence their final purchasing decision. Furthermore, when a consumer explores multiple products within the same category, their behavior towards subsequent products is influenced by their prior interactions [9]. The author will focus on performing statistical analysis of consumer behavior sequences over time. This involves using consumer ID, product category ID, and product ID as key variables to analyze behavior chronologically and establish a sequence of actions. It should be noted that the author ignored the interspersion and selection behavior of consumers between different products, and only started with the time point when the first action was taken on a certain product to conduct behavior statistics on the operation behavior of that product, and converted the behavior time attribute to a 24-hour clock.

The author applies two machine learning methods to analyze consumer behavior sequences and other features, starting with the analysis of consumer behavior sequences arranged in chronological order. Due to the different lengths of consumer behavior, before analyzing consumer behavior, the behavior sequence is first

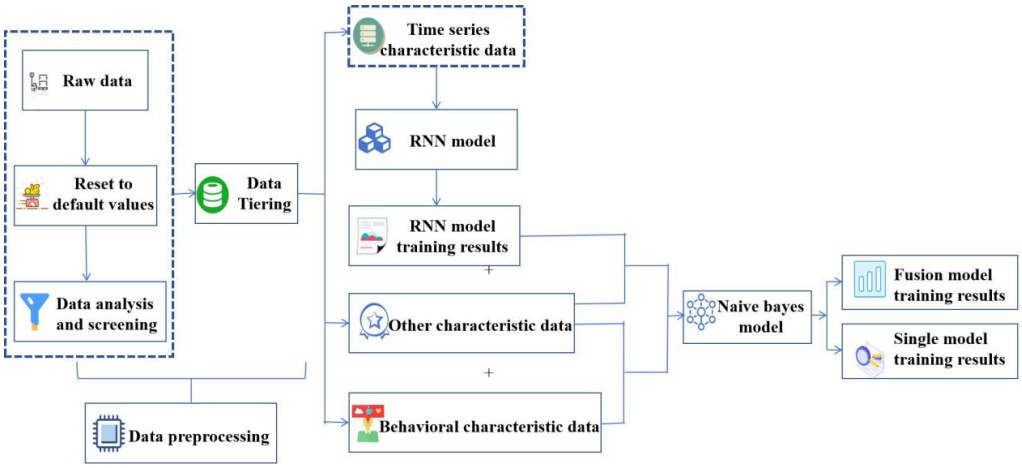

Fig. 3.1: Fusion Model Structure Diagram

stratified according to the number of behaviors. Assuming an analysis is conducted on a hierarchy with 5 behaviors, the first step is to extract the behavior sequence data from the dataset, which includes time point sequences associated with it and the true identification of whether the consumer has made a purchase, and process it into 5x24 dimensions. Behavior is divided according to the level of interest in the product, with integer values, and the time point when the behavior is empty is set to 0. Let the dataset be X and use X as the input for the RNN. RNN is an Nvs1 structure, and the input for each record in the dataset is an array of length n+1. The desired output is the probability value of whether the consumer will make a purchase, ranging from [-1,1]. The obtained probability values are added as new features to other feature sets as inputs for the Naive Bayes model [10]. The model is trained to determine the probability of whether consumers will purchase, and then the final decision result of the model is obtained, represented by 0 and 1. The fusion model structure is shown in Figure 3.1.

3.1.2. Characteristics of Online Shopping Behavior. Consumers' daily shopping behavior is generally captured through two types of feedback: explicit and implicit. Explicit feedback includes actions such as ratings, likes, and reviews, where consumers directly express their opinions about a product. In contrast, implicit feedback involves behavioral data generated during the shopping process, such as browsing, clicking, bookmarking, and adding items to the cart. Since users often do not actively rate products after purchase and may not always provide genuine feedback if they do, due to incentives like cashback or simply following requests for reviews, explicit feedback may not always accurately reflect true consumer sentiments and often suffers from data sparsity. Therefore, implicit feedback provides a more comprehensive view, addressing the limitations of explicit feedback, to some extent expressing users' true feelings, and due to the strong database support of e-commerce behavior, it can ensure data quality and accuracy. Based on the literature review above, the current research focus on conversion rates has shifted from explicit feedback data to implicit feedback data [11]. Meanwhile, due to the accumulation of massive purchasing data, research on consumer shopping behavior is no longer limited to analyzing a small amount of questionnaire data to identify influencing factors or customer segmentation and value recognition. Instead, it directly predicts consumer purchasing behavior in a more accurate way. Consequently, the author focused on user implicit feedback related to shopping behavior for the study. Utilizing a real dataset from the Alibaba e-commerce platform, the author applied big data analysis techniques to develop a model aimed at forecasting user purchasing behavior [12].

3.2. Model Algorithm.

3.2.1. Logistic Regression. Logistic regression (LR) is a probability based linear binary classification algorithm that assumes that the binary variable y of whether an event occurs follows a Bernoulli distribution, that is, the probability of event A occurring is equal to the probability of $\rho(\overline{A})$ not occurring, which is equal to

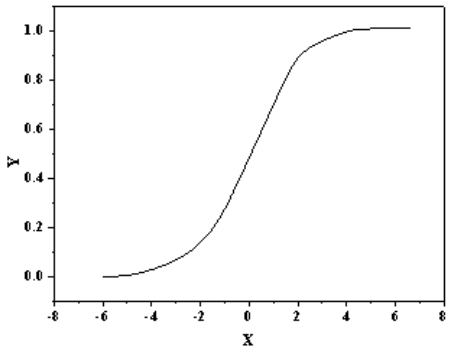

Fig. 3.2: Sigmoid function diagram

p(A). If y=1,0 is used to represent the occurrence and non occurrence of an event, the distribution function of the Bernoulli distribution is:

$$\rho(y|xl\theta) = \rho(A)^y[1-\rho(\overline{A})]^{1-y} \tag{3.1}$$

Logistic regression uses the Sigmoid function $h(\theta)$ to describe the probability of event A occurring (A):

$$\rho(A) = \rho(y=1|x;\theta) = h_\theta(x) \tag{3.2}$$

$$\rho(\overline{A}) = 1-\rho(A) = \rho(y=0|x;\theta) = 1-h_\theta(x) \tag{3.3}$$

Among them:

$$h(\theta) = 1/(1+e^{-(\theta^T x+b)}) \tag{3.4}$$

Logistic regression assumes that each sample in the dataset is independently and identically distributed. Using maximum likelihood estimation, the likelihood function for the parameter θ is given by:

$$L(\theta) = \rho(y|x;\theta) = \prod_{i=1}^{n}\rho(y_i|x;\theta) = \prod_{i=1}^{n}(h_\theta(x_i))^{y_i}[1-h_\theta(x^{(i)})]^{1-y_i} \tag{3.5}$$

By taking the logarithm of the likelihood function $L(\theta)$ and then differentiating it, we can determine the parameter θ that maximizes this function. This process leads to the formulation of the Sigmoid function, as illustrated in Figure 3.2. Data points with Sigmoid function values exceeding 0.5 are classified into one category, while those with values below 0.5 are classified into a different category [13,14].

In order to avoid overfitting, it is necessary to add regularization term $r(\theta) = ||w||^2$ and penalty parameter C to the optimization equation in practical applications. The original problem becomes:

$$max_{\theta,b} ln[\rho(y|x,\theta)] + Cr(\theta) \tag{3.6}$$

Among them, C is a hyperparameter that needs to be manually set.

3.2.2. Support Vector Machine. Support Vector Machine (SVM) is a classification technique grounded in the concept of maximizing the margin between classes. The fundamental approach involves identifying a hyperplane that divides the samples with the largest possible margin. This means that the distance between

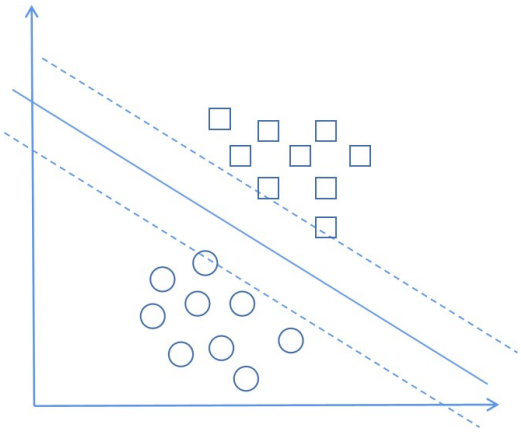

Fig. 3.3: Schematic diagram of support vector machine

the hyperplane and the nearest data points from either class is maximized. The optimal classification surface is defined by:

$$max_{w,b}\frac{2}{||w||}s.t.y_i[(w\cdot x_i)+b]-1\geqslant 0, i=1,2,\cdots,n \tag{3.7}$$

Among them, 2/|Iwl is twice the distance between the sample closest to the classification plane and the classification plane, with constraint s t. ensure correct classification for all samples. As shown in Figure 3.3, circles and crosses represent two types of samples, respectively. The circle and cross that fall on the dashed line are the samples closest to the solid line. If there is a solid line that maximizes the absolute distance between the circle and cross closest to it, then this solid line is the optimal classification surface.

In this context, 2/|Iwl represents twice the distance from the sample closest to the classification hyperplane to the hyperplane itself, while adhering to constraints that ensure all samples are classified correctly. As shown in Figure 3.3, circles and squares represent two types of samples, respectively. The circles and blocks that fall on the dashed line are the samples closest to the solid line. If there is a solid line that maximizes the absolute distance between the circles and blocks closest to it, then this solid line is the optimal classification surface [15].

Determining the optimal classification boundary can be reformulated as a convex quadratic programming problem, which involves optimizing a quadratic objective function subject to linear constraints.

$$min_{w,b}\frac{1}{||2||}||\omega||^2 s.t.y_i[(w\cdot x_i)+b]-1\geqslant 0, i=1,2,\cdots,n \tag{3.8}$$

The extent to which a sample deviates from the constraint conditions is captured by the relaxation variable ξ. Consequently, finding the optimal classification boundary is reformulated as:

$$min_{w,b,\xi}\frac{1}{||2||}||\omega||^2+C\sum_{i=1}^{n}\xi_i s.t.y_i[(w\cdot x_i)+b]\geqslant 1-\xi_i,\xi_i\geqslant, i=1,2,\cdots,n \tag{3.9}$$

Among them, C is used to "punish" those samples that violate the constraint conditions, and the larger the C, the greater the punishment. C, like δ, is also a hyperparameter [16].

3.2.3. Fusion Algorithm. The core of machine learning involves selecting the best hypothesis from a wide range of possible hypotheses through various algorithms. Each specific learning task often requires a different algorithm suited to its particular needs. In practice, it is frequently unclear which algorithm is most appropriate for a given problem, and no single algorithm is universally effective across all domains. Algorithm fusion addresses this by combining the results from multiple individual algorithms to create a new composite

model, which enhances the overall accuracy of the learning process. As computing and storage resources become more accessible, the use of ensemble methods, which integrate multiple algorithms, is gaining traction. The technique used for combining these algorithms is critical to improving the final model's performance. Common fusion methods include bagging, as seen in random forests, and boosting, used in algorithms like AdaBoost [17].

Researchers have shown that for a group of independent classification algorithms, when their accuracy in classifying a problem is greater than 0.5 (that is better than random guessing), the accuracy of using majority voting for classification will increase with the increase of the number of algorithms. Assuming di is the posterior probability of each sample classification result, and di is independent and identically distributed, E (di) is the expectation, Var (di) is the variance, if the weights of each base algorithm are set to wi=1/T (i=1,2,..., T), i.e., using the simple average method for algorithm fusion, the expected and average values of the fused algorithm are:

$$E(\overline{d_i}) = E(\sum_{i=1}^{T} \frac{1}{T} d_i) = \frac{1}{T} TE[d_i] = E(d_i) \tag{3.10}$$

$$Var(\overline{d_i}) = Var(\sum_{i=1}^{T} \frac{1}{T} d_i) = \frac{1}{T^2} TVar[d_i] = \frac{1}{T} Var(d_i) \tag{3.11}$$

From equation 3.12, it can be concluded that the expected value of the fusion algorithm remains unchanged compared to the original base algorithm, while the variance decreases as the number of base algorithms T increases [18]. As a result, a fusion algorithm typically offers higher classification accuracy compared to individual algorithms. However, in practice, increasing the number of fusion algorithms, denoted as T, does not always lead to better performance. It is essential to balance factors such as model complexity and computation time when selecting the optimal value for T. For the author's dataset, experiments indicated that using 2 to 3 fusion algorithms yields the best results. Generally, fusion algorithms tend to provide superior generalization compared to single algorithms, which can be understood through the following three intuitive perspectives:

1. From a data perspective, a single sample set might not provide enough information for a learning algorithm to identify the correct hypothesis. However, by combining multiple hypotheses that each achieve a degree of accuracy, the fusion algorithm can better approximate the true hypothesis within the hypothesis space.
2. From an algorithmic perspective, the ideal hypothesis for a given sample set might not exist within the hypothesis space of a single algorithm. By merging hypotheses from multiple algorithms, we effectively broaden the hypothesis space and get closer to the correct solution.
3. From a computational perspective, many algorithms perform localized searches within the hypothesis space, which can lead to missing the optimal solution and getting stuck in local optima. Fusion algorithms, by starting from various initial points, can collectively cover more of the hypothesis space and more accurately approach the optimal hypothesis.

The weighted average method is of particular significance for algorithm fusion, and it was popular in the 1950s. Perrone and Cooper officially used it for algorithm fusion in 1993. Algorithm fusion fundamentally involves assigning weights to each base algorithm based on learning from the sample data. In essence, different fusion techniques can be viewed as specific cases or variations of the weighted average method. For example, the simple average method is a particular instance of the weighted average approach where each algorithm is given an equal weight wi=1/T (i=1, 2..., T) being the total number of algorithms in the fusion. However, in practical scenarios, due to incomplete data and inherent noise, the weights determined from the sample data might not be entirely accurate and could potentially lead to overfitting. Research indicates that the weighted average method does not always outperform the simple average method in real-world learning tasks. For fusion algorithms to outperform the individual algorithms they comprise, it is essential that the constituent algorithms are both highly accurate and exhibit significant diversity.

Logistic regression and support vector machines are based on different principles: logistic regression uses probability-based classification, while support vector machines focus on maximizing margins between classes. Despite these differing approaches, combining them in a fusion algorithm is believed to reduce the variance

of the learning outcomes and enhance overall accuracy. Empirical evidence from the author supports this view. Additionally, the simple average method for fusion has been shown to provide better prediction accuracy compared to the weighted average method. Consequently, the author employs the SoftVoting technique from the Scikit-learn library, a popular Python-based machine learning platform, to integrate logistic regression and support vector machines. The SoftVoting method determines the final classification by averaging the probabilities predicted by each model for all possible categories and selecting the category with the highest average probability. To effectively use SoftVoting, it is essential to first obtain the probability estimates from each individual model for the categories of "purchased" and "not purchased" before performing the fusion. The prediction result of logistic regression is represented as a probability value using the Sigmoid function $h(\theta) = 1/(1 + e^{-}(\theta^T x + b))$, while the support vector machine directly provides a binary output of "1" or "0" for its predictions, it is necessary to transform these binary results into corresponding probability values to integrate with the SoftVoting method. The author uses PlattScaling method to probabilistically convert the prediction results of SVM.

4. Results Analysis.

4.1. Research on Predicting Purchase Behavior on E-commerce Websites. The empirical framework of the machine learning model is shown in Figure 4.1. Firstly, after conducting certain statistical and visual analysis on the raw data, it is preprocessed to remove duplicate and default values from the data. Before filtering, the deduplicated data is first used to calculate the behavior sequence of each consumer towards a certain product, with the product as the smallest unit. Through preliminary analysis of consumer operation sequences, records that do not conform to behavioral patterns and have too few operation times are screened out, and the extraction of other feature indicators is completed. A new dataset is formed by fusion and set as D1. Perform a single naive Bayes training on the dataset and save the training results for comparing model performance. Secondly, establish the necessary environment for RNN experiments and divide the dataset obtained in the previous step into two parts: consumer behavior sequence data d1 and other feature data d2; Divide the training and testing sets based on the size of the specific behavior sequence dataset, and input them into the RNN model. The training ultimately yields the optimal behavior preference score. Then, the behavior preference score obtained by RNN is added to the D1 dataset as one of the new feature items, and a new dataset is formed as D2. Split the dataset into a training set and a testing set with a 10:1 ratio. Train the model using the training set, where the outputs are designated as 0 and 1. After training, evaluate and compare the results by calculating and visualizing the prediction outcomes from both the single Naive Bayes model and the fusion model. By comparing the two, the advantages of the fusion model are summarized, and the conditions and scenarios for the model to be established are explained.

4.2. Introduction to Empirical Data and Predictive Objectives. The dataset is derived from actual user shopping behavior on Alibaba's mobile e-commerce platform and includes five key fields: user ID, product ID, product category ID, type of user interaction with the product (e.g., click, bookmark, add to cart, purchase), and the timestamp of the behavior. It encompasses data from 19,772 users, 422,858 products, and 1,054 product categories, collected over a period from November 18, 2014, to December 18, 2014—spanning 31 days. The goal is to predict each user's purchase behavior on December 19, 2014 (the 32nd day) for all products they interacted with during the previous 31 days. Each sample is uniquely identified by a combination of user ID, product ID, and behavior timestamp, resulting in a total of 2,084,859 records.

Given the class imbalance in the prediction problem—where the number of non-purchase samples significantly outweighs the purchase samples—traditional classification algorithms may struggle with such imbalanced data. Additionally, due to constraints on computation time and memory, processing all samples is impractical. To address these issues, the author uses sampling techniques. Since purchase samples are rare but crucial for predicting future purchases, all 730 purchase samples are retained. From the 87,383 non-purchase samples, 1,500 are randomly selected to create the training set. This adjustment helps balance the sample distribution, approximating a 1:2 ratio between the two categories. For evaluating the model's performance in this imbalanced classification scenario, the author opts for the F1 score, which is a weighted harmonic mean of precision

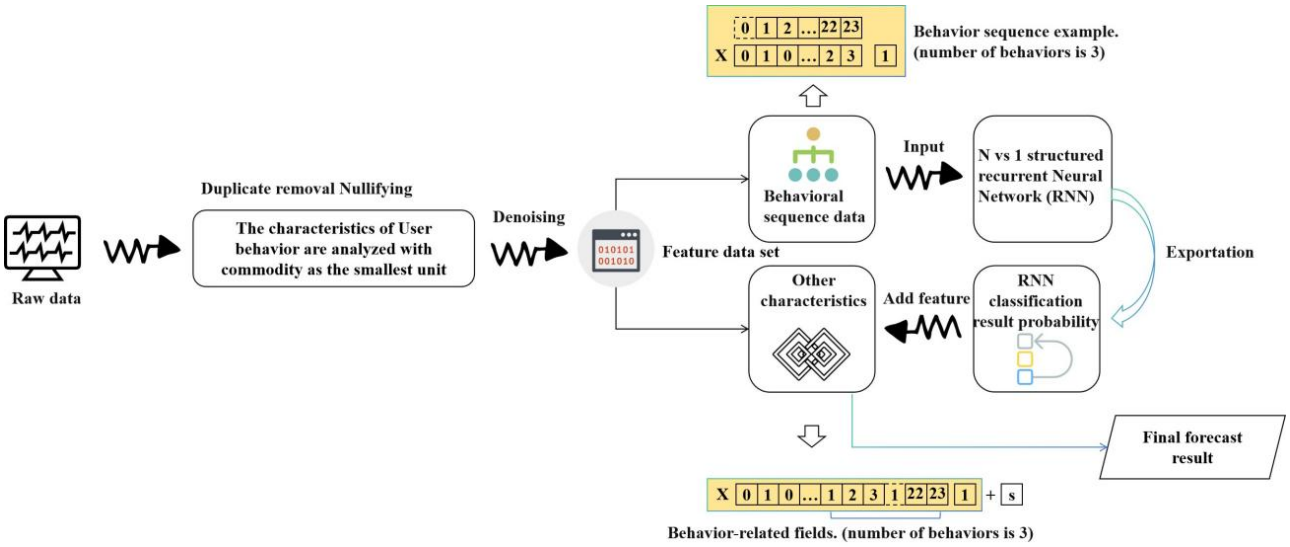

Fig. 4.1: Empirical framework of the model

and recall, rather than the traditional error rate. The formula for calculating the F1 score is:

$$F1\frac{2 \times P \times R}{P + R} \tag{4.1}$$

The formulas for computing precision P and recall R are:

$$P = TP/(TP + FP) \tag{4.2}$$

$$R = TP/(TP + FN) \tag{4.3}$$

In this context, TP refers to the count of samples where the model correctly predicts the occurrence of purchasing behavior. Meanwhile, FP denotes the number of samples where the model incorrectly predicts a purchase, and FN indicates the count of samples where the model fails to detect actual purchasing behavior.

4.3. Prediction results of online purchasing behavior based on logistic regression algorithm. A training set containing 2,230 samples—comprising 730 purchase instances and 1,500 non-purchase instances—was used with the logistic regression algorithm. The hyperparameter C was explored within the range of [2,240], with 10 values selected in an exponential scale within this range. For each C value, the algorithm was subjected to 3-fold cross-validation, resulting in 30 learning iterations to identify the optimal model. The test set was then evaluated with this model, yielding 656 samples predicted to exhibit purchasing behavior on the 32nd day. The F1 score of this prediction, when compared to the actual purchases made on the 32nd day, was found to be 7.73%.

4.4. Prediction results of online purchasing behavior based on support vector machine algorithm. The "soft margin" Support Vector Machine (SVM) algorithm with an RBF kernel was utilized to build the prediction model. The process began by training the algorithm with 2,230 samples from the training set. To find the optimal hyperparameters, namely C and the Gaussian kernel bandwidth δ, a layered 3-fold cross-validation approach was employed. The ranges for C and δ were set within [2,210], and 10 values for each were chosen in an exponential scale, resulting in a total of 100 parameter combinations (10 values of C $\times$ 10 values of δ). Each combination underwent 3-fold cross-validation, leading to 300 iterations to determine the best model. Finally, the model was tested with 1,445 samples from the test set, producing 577 samples with a prediction of "1," indicating the likelihood of purchasing behavior on the 32nd day. Compared with the sample that actually made a purchase on the 32nd day, its F1 score was 7.75%.

Table 4.1: Comparison of Prediction Results of Three Models

	Sample size for predicting purchase behavior	F1 score
Logistic regression	656	7.73
Support vector machine	577	7.75
fusion algorithm	571	7.77

4.5. Logistic regression support vector machine fusion algorithm for predicting online purchasing behavior. The steps for building a model using a single algorithm are the same. Initially, a training set of 2,230 samples is utilized to train a hybrid model that combines logistic regression and support vector machine algorithms, employing the Soft Voting method for integration. Then use layered 3-fold cross validation on the hybrid algorithm to obtain the optimal fusion algorithm. Due to limitations in computing resources, the author restricts the selection range of fusion algorithm parameters. Considering the predictive performance of both the logistic regression and support vector machine algorithms across various parameter settings, the author defined the hyperparameter ranges for the support vector machine within $[2^7, 2^{10}]$. Three specific values for C and δ were chosen within this range, arranged in exponential increments; Select the range of values for the logistic regression hyperparameter C within $[2^{15}2^{20}]$, and choose 5 values within this range in exponential order as the values for C; The 3 parameters (SVM-C, 8, LR — C) have a total of 3x3x5 = 45, and the algorithm carries out 3 times cross-verification for every value, making a total of 3x45 = 135 computations. Finally, we input 1445 samples into the model and get the predicted results. A sample of "1" predicted that buying behaviour would take place on Day 32, in a total of 571 samples. Compared to those who bought it on the 32nd day, their F1 score was 7.77 percent.

4.6. Comparison of Prediction Effects of Three Models. The F1 results of 3 models built with various algorithms are presented in Table 4.1, respectively. Comparison of F1 results shows that this method has the least amount of data for forecasting buying behaviour, but it is superior to that of other individual models. Though the F1 of this method is only 0.02 percent better than that of the single-model, it is not a negligible increase since the F1 is needed for the forecast.

5. Conclusion. The rapid development of technology has also driven the continuous improvement and optimization of machine learning algorithms, but machine learning algorithms have not yet been fully utilized in practical business applications. Machine learning algorithms are simple algorithms that can perform big data analysis. Many large-scale companies are fully aware of the importance of machine learning algorithms, and the most obvious application effect is the "Double Eleven" event held by Alibaba and Tmall every year. Machine learning algorithms can be applied in a variety of ways in the process of business development. It can be a single model algorithm or a combination of two single model algorithms to form a fusion algorithm. The author mainly studies which of the two algorithms, the single model algorithm and the fusion algorithm algorithm, can achieve more accurate budget results in the business application process. The proposed method is applied to forecast the buying behaviour of consumers under on-line shopping environment and it is found that the forecast precision of this method is better than that of single-model.

6. Acknowledgement. This work was supported by Shaanxi Provincial Social Science Foundation of China: Research on the Circulation Mechanism of "Agriculture Consumer Connection" of Shaanxi Characteristic Agricultural Products (No. 2022D052).

REFERENCES

[1] Feldman, J. B., Zhang, D. J., Liu, X., & Zhang, N. (2022). Customer choice models vs. machine learning: finding optimal product displays on alibaba. Oper. Res., 70, 309-328.

[2] Chen, S., Ngai, E. W. T., Xiao, F., & Xu, Z. (2024). From comparison to purchasing: effects of online behavior toward associated co-visited products on consumer purchase. Information & Management, 61(3).

[3] Fekete-Farkas, M. (2022). Startups and consumer purchase behavior: application of support vector machine algorithm. Big Data and Cognitive Computing, 6.

[4] Chen, C. W., & Li, M. E. (2024). Utilizing a Hybrid Approach to Identify the Importance of Factors That Influence Consumer Decision-Making Behavior in Purchasing Sustainable Products, 28(6), 2579-2595.
[5] Nguyen, K., Mai, T. N., Nguyen, H. A., & Nguyen, V. A. (2023). Retracted article: a computational model for predicting customer behaviors using transformer adapted with tabular features. International Journal of Computational Intelligence Systems, 16(1).
[6] Pandi, V., Nithiyanandam, P., & Sindhuja ManickavasagamIslabudeen Mohamed MeerashaRagaventhiran JaganathanMuthu Kumar Balasubramanian. (2022). A comprehensive analysis of consumer decisions on twitter dataset using machine learning algorithms. IAES International Journal of Artificial Intelligence, 11(3), 1085-1093.
[7] Nakao, J., & Nishi, T. (2022). A bilevel production planning using machine learning-based customer modeling. Journal of Advanced Mechanical Design, Systems, and Manufacturing.
[8] Prakash, R., Anoosh, R. K., & Sharon, S. (2022). Comparative study of machine learning algorithms for product recommendation based on user experience. ECS transactions(1), 107.
[9] Bangyal, W., Ashraf, A., Shakir, R.,& Rehaman, N. U. (2022). A review on consumer behavior towards online shopping using machine learning. International Journal of Emerging Multidisciplinaries: Computer Science & Artificial Intelligence, 77(6), 395-406.
[10] Parihar, V., & Yadav, S. (2022). Comparative analysis of different machine learning algorithms to predict online shoppers' behaviour. International Journal of Advanced Networking and Applications, 13(6), 5169-5182.
[11] Chaudhary, P., Kalra, V., & Sharma, S. (2022). A hybrid machine learning approach for customer segmentation using rfm analysis.
[12] Sun, J., Pang, S., Qiu, M., Li, J., Zhao, C., & Song, W., et al. (2022). Predictive Analysis of Power Purchase Behavior of Beijing Residents and Research on Active Equipment Repair and Operation, 19(4), 486-495.
[13] Habbab, F. Z., & Kampouridis, M. (2024). An in-depth investigation of five machine learning algorithms for optimizing mixed-asset portfolios including reits. Expert Systems with Application(Jan.), 235.
[14] Pawar, N., & Tijare, P. A. (2023). A review on phishing website detection using machine learning approach. International Journal of Scientific Research in Computer Science, Engineering and Information Technology, 24(2), 125-137.
[15] Rastogi, S., Agarwal, D., Jain, J., & Arjun, K. P. (2022). Demographic filtering for movie recommendation system using machine learning, 6(2), 157-168.
[16] Munde, A., & Kaur, J. (2024). Predictive modelling of customer sustainable jewelry purchases using machine learning algorithms. Procedia Computer Science, 235, 683-700.
[17] Priel, R., & Rokach, L. (2024). Machine learning-based stock picking using value investing and quality features. Neural Computing and Applications, 36(20), 11963-11986.
[18] Tran, D. T., & Huh, J. H. (2022). Building a model to exploit association rules and analyze purchasing behavior based on rough set theory. The Journal of Supercomputing, 78(8), 11051-11091.

Edited by: Bradha Madhavan
Special issue on: High-performance Computing Algorithms for Material Sciences
Received: Aug 17, 2024
Accepted: Feb 26, 2025