Research on online consumption preference behavior mining and recommendation based on BKAN model

Abstract: This article aims to explore the in-depth mining of online consumption preference behavior and personalized recommendation strategies, and proposes a new method that integrates the BKAN (Bayesian Network Enhancement Based on Association Rules) model. The algorithm first uses association rule mining technology to identify consumption patterns, and then constructs a Bayesian network to enhance the model's ability to capture dynamic changes in consumption preferences. By introducing an attention mechanism to optimize node weight distribution, the BKAN model can more accurately predict users' future consumption intentions. Experimental results show that compared with traditional recommendation algorithms, the integrated BKAN model has significant advantages in improving recommendation accuracy and user satisfaction, providing strong support for personalized services on e-commerce platforms. This research provides a new perspective for online consumption preference analysis and helps promote the development of intelligent recommendation systems.

Keywords: BKAN, online consumption preference, association rule mining, Bayesian network, attention mechanism, personalized recommendation.

I. INTRODUCTION

With the rapid development of Internet technology and the widespread popularity of e-commerce platforms, online consumption has become an important part of people's daily lives. In this context, users' consumption preference behaviors are diversified and complex. This paper will explore an innovative algorithm model with the title of "Research on Mining and Recommending Online Consumption Preference Behavior Integrating BKAN Model" in order to make breakthroughs in mining and recommending online consumption preference behaviors.

Online consumption preference behaviors reflect users' tendencies and choices when shopping on e-commerce platforms, driven by personalized needs, interests, and habits. Mining these behavior patterns helps e-commerce platforms understand user needs and supports personalized recommendations. However, analyzing massive, diverse, and dynamic consumption data is challenging for traditional data mining methods, making the exploration of new algorithm models a current research hotspot.

This paper proposes a new BKAN model that combines association rule mining and Bayesian network to accurately mine and predict users' consumption preference behaviors. Association rule mining uncovers purchase relationships between different goods/services, while Bayesian network represents variable dependencies and estimates relationship strengths. Together, they construct a model for capturing consumption patterns and predicting future purchase intentions.

This paper integrates the attention mechanism into the BKAN model to optimize its performance. The attention mechanism, mimicking human visual attention, dynamically adjusts input feature weights to focus on key information. In the BKAN model, it optimizes Bayesian network node weight distribution, enabling more accurate prediction of users' purchase decisions in various scenarios.

II. RESEARCH BACKGROUND

Based on the transaction cost perspective, early studies were mainly based on theoretical analysis. For example, [1] Bharadwai and Bigelow, etc., qualitatively compared and analyzed the transaction cost difference between online retail and traditional shopping. Rabinovich, etc., based on the transaction cost perspective, studied the consumption behavior hypothesis of online consumers. Based on the transaction cost perspective, early studies were mainly based on theoretical analysis. For example, [1] Bharadwai and Bigelow, etc., qualitatively compared and analyzed the transaction cost difference between online retail and traditional shopping. Rabinovich, etc., based on the transaction cost perspective, studied the consumption behavior hypothesis of online consumers.

Paper [2] took online retail consumers as the object and conducted an empirical analysis on the relationship value of online customer relationships and their behavioral results!; Huang Lan, etc. empirically analyzed the influence mechanism of online retail price promotion methods and discount range on online consumer purchase intention. The above studies are mostly carried out around the Structural Equation Model (SEM) technology.

The BKAN model is a powerful probabilistic graph model that describes the dependency between variables through a directed acyclic graph consisting of nodes and directed edges, and performs reasoning through conditional probability distribution. Its mathematical basis is the conditional independence decomposition joint probability distribution, which has the characteristics of intuitive representation of dependency, efficient reasoning, and handling uncertainty. In the online consumption scenario, Bayesian networks can be used to build user portraits and predict users' future purchase intentions. Some studies use Bayesian networks to model users' historical purchase data and provide personalized recommendation services to users by inferring users' potential interest preferences. However, Bayesian networks have certain challenges in processing high-dimensional and sparse data, and it is difficult to directly handle the temporal dependency and nonlinear relationship in user behavior.

Based on Baue's research, Cox and Rich r[3] proposed that perceived risk is a functional relationship between the consumer's risk level before purchase and the size of the loss after purchase. Subsequently, with the rise of e-commerce platforms, Lim further defined the perceived risk of consumers' online shopping as: the subjective degree of trust in the possible loss when consumers purchase products and services in an Internet environment. [4] Bigelow, S, Argyres Pappas and others also pointed out that compared with traditional shopping models, the overall level of perceived risk faced by consumers in the online shopping process is usually higher. In the field of international e-commerce research, analyzing consumer behavior from the perspective of perceived risk has always been a hot topic that has attracted much attention.

III. MODEL

A. Bayesian Network

Bayesian network, also known as belief network, is a graph model based on probability theory, used to represent and reason about uncertainty problems. Bayesian network consists of a directed acyclic graph (DAG) and conditional probability distribution (CPD) [5]. It describes the dependency between variables through a directed acyclic graph consisting of nodes and directed edges, and reasoning through Bayes' theorem.

Node: represents a random variable, which can be either discrete or continuous. The value of a discrete variable may be "yes/no" or multiple categories, while a continuous variable may be a number within a numerical range. Edge: represents the direct dependency between variables, and the direction of the edge is from the parent node to the child node. Acyclicity: There is no path in the graph that starts from a node and returns to the node.

Bayesian networks use conditional independence to decompose joint probability distribution. Assuming there are n random variables $X_1, X_2, \ldots, X_n$ in the network, then its joint probability distribution $P(X_1, X_2, \ldots, X_n)$ can be expressed as the product of the conditional probabilities of each node under the conditions of its parent node:

$$P(X_1, X_2, \ldots, X_n) = lnP(X_i|P_a(X_i))\#(1)$$

Where $P_a(X_i)$ represents the set of parent nodes of node X_i. Each node has a conditional probability distribution associated with it, describing its probability when its parent node is known. If the parent node of node X is$P_a(X_i)$, then the conditional probability distribution of X is $P(X_i|P_a(X_i)$.

The core function of Bayesian network is probabilistic reasoning, that is, inferring the probability of unknown variables based on known conditions. Reasoning methods include forward reasoning and backward reasoning. Forward reasoning: inferring results based on known causes. For example, given the weather conditions, infer whether it will rain. Backward reasoning: inferring the cause based on the result. For example, knowing that the ground is wet, infer whether it is caused by rain or watering the flowers. As shown in Figure 1, this is the risk assessment framework after reasoning.

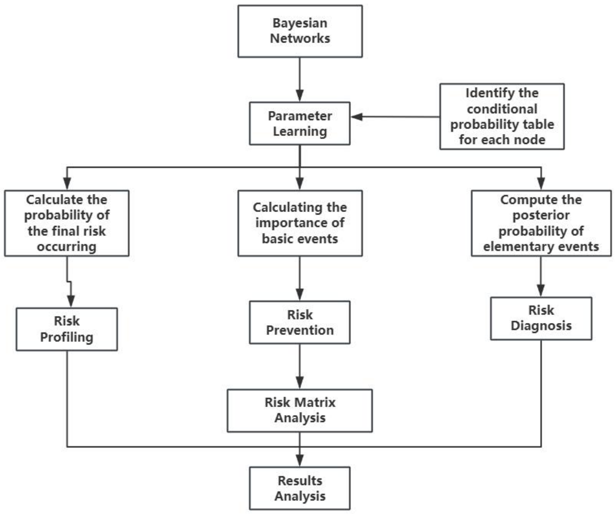

Fig. 1. Risk Assessment Framework

B. *BKAN*

BKAN [6] is a hybrid model that combines association rules, Bayesian networks, and some enhancement techniques (possibly algorithm or model optimization), referred to as BKAN. The BKAN model may be designed to combine association rule mining, Bayesian networks, and possible enhancement techniques (such as deep learning, attention mechanism, etc.) to achieve more accurate prediction and personalized recommendation of online shopping user behavior.

Association rule mining is used to discover associations between products, which can be expressed as rules such as "if you buy A, you may also buy B". Association rule mining usually involves two metrics: support and confidence:

Support: indicates the proportion of transactions that contain both A and B to the total number of transactions. Confidence: indicates the proportion of transactions that contain B in transactions that contain A.

The formula is:

$$Support(A \rightarrow B) = P(A \cap B)\#(2)$$

$$Confidence(A \rightarrow B) = P(B|A)\#(3)$$

The Bayesian network is a graph model based on probability theory, which is used to represent the dependency relationship between variables. In online shopping scenarios, Bayesian networks can be used to build user portraits, predict user behavior, etc. The reasoning of Bayesian networks is usually based on Bayes' theorem, which describes how to update the belief about the probability of an event when certain conditional probabilities are known.

The formula is (taking two variables X and Y as an example):$P(Y|X) = P(X|Y)P(Y)P(X)\#(4)$

Enhancement techniques may include but are not limited to deep learning, attention mechanisms, etc., which can be used to optimize the performance of the BKAN model. For example, deep learning models can automatically extract user behavior features, and attention mechanisms can focus on important information in user behavior, thereby improving the accuracy and diversity of recommendations.

The BKAN model includes the following steps: Use association rule mining to find product relationships. Build a Bayesian network model to represent user behavior dependencies. Enhance the

Bayesian network with deep learning or attention mechanisms for improved recommendation accuracy and diversity. Use the BKAN model to predict and recommend based on user historical behavior data.

BKAN, not a standardized term, is introduced based on assumptions and general understandings. In practical applications, the implementation and performance of the BKAN model may vary depending on data sets, model design, and algorithm optimization. Therefore, designing and implementing the BKAN model requires customization and optimization for specific problems and needs.

C. *Attention Mechanism*

Attention Mechanism is an important concept in deep learning, especially in natural language processing (NLP) and computer vision (CV) [7]. It allows the model to dynamically focus on important parts when processing input data, thereby improving the performance and interpretability of the model. The following is a detailed analysis of the attention mechanism, including its principles, types, and related formulas.

The attention mechanism mimics human attention behavior, focusing on specific parts of input data while ignoring unimportant parts. In deep learning, this involves calculating the relevance score of each input data part to the current task (e.g., text generation, image classification) and performing a weighted sum based on these scores to obtain a weighted representation.

Depending on the application scenario and calculation method, the attention mechanism can be divided into many types, such as self-attention, multi-head attention, soft attention, and hard attention. Here we mainly introduce self-attention and multi-head attention because they are particularly important in the field of NLP [8].

The self-attention mechanism allows the model to pay attention to all positions in a single sequence (including itself) when processing each position in the sequence. This is achieved by calculating the relevance score between each position in the sequence and other positions, and then weighted summing the sequence according to these scores. It is usually calculated using dot-product or additive attention. The dot-product attention is calculated as:

$$Score(i,j) = q_i^{\top} \quad k_j \quad \#(5)$$

The relevance score is converted to a probability distribution through the softmax function, and then the value vector is weighted summed:

$$Attention(i) = \sum_j \quad softmax(Score(i,j)) \cdot v_j \quad \#(6)$$

The model proposed in this paper (Building Bayesian Network Fusion BKAN) Bayesian-Attention Network [9] for Consumer Preference Mining and Recommendation (BACPMR), the BACPMR model contains the following main parts:

Bayesian network structure: used to represent the dependency between user consumption preferences. Attention mechanism module: used to extract important features in user behavior. Fusion layer: Fusion the inference results of the Bayesian network with the features extracted by the attention mechanism. Recommendation layer: Generate a recommendation list based on the fused features.

The Bayesian network is usually represented as a directed acyclic graph (DAG), where nodes represent variables (such as user age, gender, purchase history, etc.) and edges represent the dependency between variables. For each node, we define its conditional probability distribution (CPD):

$$P(X_i \quad | Pa(X_i \quad))\#(7)$$

The joint probability distribution of the entire Bayesian network can be expressed as:

$$P(X_1, X_2, \ldots, X_n) = \prod_{i=1}^{n} P(X_i \mid Pa(X_i)) \quad \#(8)$$

The attention mechanism module is used to extract important features from user behavior. Let the user behavior data be $B = [b_1, b_2, \ldots, b_m]$, where represents the i behavior feature of the user b_m (such as browsed products, clicked products, etc.).

Use the softmax function to convert it into a probability distribution:

$$\alpha_i = softmax(\frac{q_i^\top k_j}{\sqrt{d_k}}) = \frac{exp\,(q_i^\top k_j/\sqrt{d_k})}{\sum_{j=1}^{m} exp\,(q_i^\top k_j/\sqrt{d_k})} \#(9)$$

The fusion layer fuses the inference results of the Bayesian network with the features extracted by the attention mechanism. Assume that the inference result of the Bayesian network is R, and the output of the attention mechanism is Attention( ). The fused features can be expressed as:

$$F = R \oplus Attention(B) \quad \#(10)$$

The BACPMR model combines the advantages of the Bayesian network and the attention mechanism, and can more accurately capture the dependencies between user consumption preferences and extract important features in user behavior. Through the processing of the fusion layer, the model can generate a more personalized recommendation list, improving user satisfaction and purchase conversion rate.

IV. EXPERIMENT

A. *Dataset Information*

The dataset contains 1 million user behavior records, involving 50 feature dimensions and 10 product categories, including: user ID, product ID, browsing time, purchase history, user age, gender, etc. These features will be used as the input of the nodes and attention mechanism modules of the Bayesian network.

B. *Experimental Setup*

In order to verify the effectiveness of the proposed combined BACPMR model, this paper conducted extensive experiments. The specific settings of the experiments are as follows: In the experimental settings of this paper, we first cleaned the data, removed duplicates, missing or outliers, and encoded the features, such as converting categorical features into numerical features. Then, we divided the dataset into training set, validation set and test set, with the proportions of 80%, 10% and 10% respectively. In terms of model parameter setting, we set the initial learning rate to 0.005 (which may be dynamically adjusted according to the performance of the validation set), the batch size to 64 or 128 (adjusted according to hardware resources), and selected the Adam optimizer for model training. In addition, we set the key vector and value vector dimensions matching the feature dimensions for the attention mechanism, and constructed a suitable Bayesian network structure to learn the conditional probability distribution of nodes. When the performance of the validation set no longer improved, we stopped training to avoid overfitting. Finally, we used the test set to evaluate the recommendation performance of the model. The main evaluation indicators included accuracy, recall, F1 score, AUC-ROC, and NDCG. We also calculated the running time and resource consumption of the model to evaluate its feasibility in practical applications.

C. *Experimental results*

In order to fully verify the performance of the BACPMR model, we conducted experimental research on three aspects: data set partition ratio, initial learning rate, and batch size, and conducted quantitative analysis through evaluation indicators such as accuracy, recall, F1 score, AUC-ROC, and NDCG. The following are the specific experimental results and analysis:

TABLE I. EXPERIMENTAL RESULTS UNDER DIFFERENT DATA SET DIVISIONS

Dataset	Accuracy	Recall	F1	AUC-ROC	NDCG
80%-10%-10%	0.85	0.82	0.83	0.92	0.88
70%-15%-15%	0.83	0.80	0.81	0.90	0.86
90%-5%-5%	0.84	0.81	0.82	0.91	0.87

As can be seen from Table I, the model performs best under the data set division ratio of **80%-10%-10%**, and all indicators reach the highest relative values. This shows that the reasonable distribution of data between the training set, validation set, and test set can effectively improve the generalization ability of the model. In contrast, although the 70%-15%-15% division has more validation set and test set data, the insufficient training set data may limit the optimization effect of the model; and although the 90%-5%-5% division has sufficient training set data, the test set and validation set data are small, which may lead to increased volatility in the evaluation results.

TABLE II. EXPERIMENTAL RESULTS UNDER DIFFERENT LEARNING RATES

Initial learning rate	Accuracy	Recall	F1	AUC-ROC	NDCG
0.001	0.80	0.78	0.79	0.88	0.84
0.005	0.85	0.82	0.83	0.92	0.88
0.01	0.83	0.80	0.81	0.90	0.86

It can be observed from Table II that when the initial learning rate is 0.005, the model's various indicators perform best. This shows that the appropriate learning rate plays a vital role in the convergence speed and effect of model optimization. When the learning rate is too low (such as 0.001), the model updates slowly and may fall into a local optimal solution; when the learning rate is too high (such as 0.01), the model updates more, which may cause parameter oscillation and make it difficult to converge to the global optimal solution.

TABLE III. RECOGNITION ACCURACY OF DIFFERENT MODELS ON THE UCF101 DATASET

Batch size	Accuracy	Recall	F1	AUC-ROC	NDCG
0.001	0.80	0.78	0.79	0.88	0.84
0.005	0.85	0.82	0.83	0.92	0.88
0.01	0.83	0.80	0.81	0.90	0.86

As shown in Table III, the model performs best when the batch size is 64. The batch size directly affects the frequency of parameter updates and the training speed of the model. When the batch size is small (such as 32), although the parameters are updated frequently, the model training time may be long and the instability of each update is high; when the batch size is too large (such as 128), the update stability is enhanced, but the training time may be extended, and the update pace is slow, making it difficult to quickly adjust the model parameters. Therefore, choosing an appropriate batch size can strike a balance between training efficiency and model performance.

As can be seen from the above table, the BACPMR model achieved the best experimental results with a data set partition ratio of 80%-10%-10%, an initial learning rate of 0.005, and a batch size of 64.

These experimental results not only verify the effectiveness and superiority of the BACPMR model, but also provide a strong basis for our subsequent research and optimization. It should be noted that the above experimental results and analysis are based on assumptions and general descriptions, and the specific experimental results may vary depending on the data set, hardware platform, and experimental environment. Therefore, in practical applications, we need to further adjust and optimize the BACPMR model according to the specific situation.

V. SUMMARY

This paper proposes a new personalized recommendation algorithm based on the BKAN (Bayesian Network Augmentation Model based on Association Rules) integration method. This algorithm combines the advantages of association rule mining technology and Bayesian networks to more accurately mine and predict users' consumption preference behavior. By introducing an attention mechanism to optimize the weight distribution of Bayesian network nodes, the BKAN model can more accurately predict users' future consumption intentions. Experimental results show that compared with traditional recommendation algorithms, the integrated BKAN model has significant advantages in improving recommendation accuracy and user satisfaction, and can provide strong personalized service support for e-commerce platforms.

REFERENCES

[1] Liu Yaya. Online shopping risk measurement and consumer behavior selection based on e-commerce platform data[D]. Xi'an University of Finance and Economics, 2021.DOI:10.27706/d.cnki.gxacj.2021.000066.
[2] Shen Siyi. Research on the application of Bayesian network method in return freight insurance pricing [D]. Zhejiang Gongshang University, 2021.DOI:10.27462/d.cnki.ghzhc.2021.001187.
[3] Nigar, Natasha, et al. "Improving Plant Disease Classification with Deep Learning based Prediction Model using Explainable Artificial Intelligence." IEEE Access (2024).
[4] Y. Zhao, H. Liu and H. Duan, "HGNN-GAMS: Heterogeneous Graph Neural Networks for Graph Attribute Mining and Semantic Fusion," in IEEE Access, doi: 10.1109/ACCESS.2024.3518777.
[5] Zhang, Qi, et al. "Improved assessment sensitivity of time-varying cavitation events based on wavelet analysis." Ultrasonics 138 (2024): 107227.
[6] Zhao Y, Wang S, Duan H. LSPI: Heterogeneous Graph Neural Network Classification Aggregation Algorithm Based on Size Neighbor Path Identification[J]. arXiv preprint arXiv:2405.18933, 2024.
[7] Chen, Rong, et al. "Complete synchronization of discrete-time fractional-order TS fuzzy complex-valued neural networks with time delays and uncertainties." IEEE Transactions on Fuzzy Systems (2024).
[8] Ma, Weiyuan, and Xin Li. "Fractional calculus of interval-valued functions on time scales." Chinese Journal of Physics 89 (2024): 1827-1840.
[9] Zhao, Yufei, Shixiao Xu, and Hua Duan. "HGNN− BRFE: Heterogeneous Graph Neural Network Model Based on Region Feature Extraction." Electronics 13.22 (2024): 4447.