

Coupled Optimization Algorithm of Intelligent Recommendation Model Based on Computer Natural Language Processing

Dianqing Bao^{1,a*} and Wen Su^{2,b}

¹*School of Mathematics and Information Engineering, Lianyungang Normal University, Lianyungang 222000, Jiangsu, China*

²*School of Information Engineering, Lianyungang Technical College, Lianyungang 222000, Jiangsu, China*

To address the issues of fragmented semantic representation of items and the difficulty of capturing dynamic user interests in traditional recommendation models, this paper proposes a coupling optimization algorithm for intelligent recommendation models based on natural language processing (NLP). A heterogeneous multi-source, information-deep, semantic-unified representation module is constructed, integrating knowledge graph embedding and a Bi-LSTM (Bidirectional Long Short-Term Memory) model to combine structured and unstructured item information. A time-aware user dynamic interest coupling modeling module is designed, using GRU (Gated Recurrent Unit) sequence encoding to capture behavioral temporal dependencies and combining an interest-semantic coupling attention mechanism to dynamically track interest evolution. A multi-objective collaborative adversarial coupling optimization module is introduced to achieve multi-objective collaborative optimization. Experiments show that the proposed model achieves 0.937 on Recall@20 and 0.959 on NDCG@20 (Normalized Discounted Cumulative Gain); the intra-list similarity is as low as 0.25, and the coverage reaches 0.934.

Keywords: computer natural language processing, intelligent recommendation models, coupling optimization, dynamic interest coupling, deep semantic unification

1. INTRODUCTION

The big data era has led to the generation of enormous amounts of information, resulting in unprecedented levels of information overload and choice dilemmas for the end-user. Intelligent recommendation systems are now the primary tools for information filtering and have been used in e-commerce, content platforms, social networking services, and many other applications in order to improve user experience, increase platform loyalty, and create value [1–2]. Today's ever-evolving complexity of various applications,

combined with the sophistication of user demand for customization in the form of application delivery, means that traditional collaborative filtering-based recommendations or basic content-based recommendation systems are encountering several difficulties such as poor semantic comprehension, difficulty processing extremely sparse click logs, increasing difficulty in discerning the current state of the user's evolving preferences, and inadequate ease-of-interpretability when conducting an analysis and/or generating additional recommendations utilizing these technologies. This has resulted in the generation of poor quality recommendations due to the amount/type of processed data available to the recommendation systems as well the limited technological capabilities of such systems [3–4]. Consequently, an important area of

^aEmail: lygzerg@163.com, ^bsuwenlyg@163.com,*Corresponding Author.

mutual interest to scientists and industry experts alike, is how to develop advanced technologies utilizing natural language processing (NLP) in conjunction with recommendation models that will more effectively create an integrated approach for generating advanced recommendations.

Several researchers have conducted studies on intelligent recommendation models [5–6]. Wang proposed a method for improving the accuracy of tourism destination recommendation systems by considering user preferences when calculating similarity metrics [7]. Song & He examined a tourism intelligent recommendation system developed with artificial intelligence and the IoT that allows visitors to choose from predefined itineraries based on personalized suggestions and travel insights [8]. They also analyzed the complexity of the Apriori algorithm and enhanced its efficiency by using an improved iteration on the weighted association rules. Widayanti et. al developed a new recommendation paradigm that is more comprehensive in its personalized approach, by combining collaborative filtering and content-based filtering in a synergistic way to create a personalized recommendation system [9]. Parthasarathy & Sathiyar also developed a hybrid recommendation system by extracting features and making the best recommendations based on user preferences using the firefly algorithm in conjunction with a crow search algorithm to arrive at the optimal rating [10]. However, previous researches often loosely combine or simply splice text processing modules with recommendation models, failing to achieve deep semantic alignment and interaction.

To address these challenges, researchers have explored tighter coupling mechanisms and optimization strategies [11–12]. Zhou et al. constructed an ad-relevance graph and trained graph embedding vectors based on a self-attention deep neural network to enhance the ad's ability to capture user interests [13]. They also learned the intrinsic correlation of user behavior through the self-attention mechanism and explored user interest preferences in historical behavior. Yu et al. proposed an algorithm based on a multi-feature attention mechanism and deep and convolutional neural networks to achieve personalized and accurate movie recommendations [14]. This algorithm adds user and movie attribute information, learns their features separately, and embeds users' static and dynamic preferences for items into the interaction. Zhang et al. proposed a neural time-aware recommendation network with temporal context, jointly modeling static and dynamic user preferences through a feature interaction network [15]. Cui et al. proposed a deep collaborative recommendation algorithm based on an attention mechanism, using this feature to assign weights to user item representations and capture information hidden in implicit feedback [16]. However, these methods focus mostly on single coupling and lack a unified framework. Therefore, this paper proposes a coupled optimization algorithm that combines the advantages of NLP and deep learning to solve problems such as insufficient semantic understanding, inadequate model optimization, and difficulty in coordinating the various elements of existing research.

This paper aims to address the problems of insufficient semantic understanding, difficulty in capturing user interests, and limited recommendation results in intelligent

recommendation systems under big data environments. It proposes a coupling optimization algorithm for intelligent recommendation models based on NLP to improve recommendation accuracy, diversity, and cold-start capability. First, a heterogeneous multi-source information deep semantic unified representation module is constructed. This module extracts structured semantics of items using knowledge graph embedding, combines BERT (Bidirectional Encoder Representations from Transformers) and Bi-LSTM deep encoding of review text, and fuses the two semantics through a cross-modal attention mechanism to form a unified item representation. Second, a time-aware user dynamic interest coupling modeling module is designed. This module uses GRU to model user historical behavior sequences and to capture the temporal evolution of interests. It introduces an interest-semantic coupling attention mechanism to dynamically focus on the user's current interests, and models the higher-order evolution of interests through a graph attention network to generate dynamic interest vectors. Finally, a multi-objective collaborative adversarial coupling optimization module is proposed, introducing an adversarial training mechanism to mitigate popular item bias and improve recommendation diversity and long-term coverage. Experiments verify the significant advantages of this algorithm in key metrics.

2. CONSTRUCTION OF RECOMMENDATION MODEL BASED ON DEEP SEMANTIC UNDERSTANDING AND DYNAMIC INTEREST COUPLING

2.1 Unified Deep Semantic Representation of Heterogeneous Multi-Source Information

When developing intelligent recommendation systems, a thorough understanding of items and user intent is key to improving recommendation performance [17–18]. However, it is challenging to create a cohesive item representation with contextual information using traditional methods that merely splice or treat structured and unstructured data separately, leading to semantic fragmentation and shallowness. This paper proposes a coupling optimization algorithm for intelligent recommendation models based on NLP, the core of which is the construction of a unified deep semantic representation module for heterogeneous multi-source information. This module integrates the explicit structured semantics of the item knowledge graph with the implicit unstructured semantics of user comment text, and uses representation learning and cross-modal alignment techniques to create a collaboratively enhanced deep semantic space, providing a solid semantic foundation for subsequent dynamic interest modeling and multi-objective optimization. Essentially, the structured attribute information of items constitutes a knowledge graph with items as entities and attribute relationships as edges. To capture this explicit semantic relationship, this paper applies knowledge-graph-embedding technology, mapping entities and relationships to different vector spaces

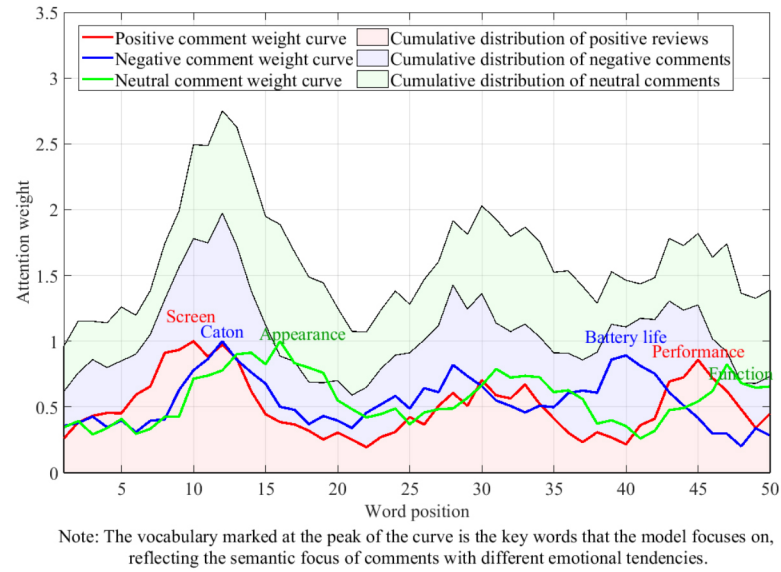


Figure 1 Weight curve distribution based on Bi-LSTM and attention mechanism.

to better model diverse relationship patterns. The set of item entities is represented by ε , and the set of relationships is R . For a triple (h, r, t) in a knowledge graph, where $h, t \in \varepsilon$ represent the head entity and tail entity respectively, and $r \in R$ represents a relation, knowledge graph embedding technology defines a projection matrix M_r for each relation r , projecting the entity vectors from the entity space onto the specific space of relation r . Its scoring function is defined as:

$$f_r(h, t) = \|M_r h + r - M_r t\|_2^2 \quad (1)$$

By minimizing a loss function constructed based on correct and negative sampling triples, a low-dimensional dense vector representation $e \in R^d$ of project entities and their attributes can be learned. This process encodes discrete label information such as project category and brand, and captures the complex semantic association network between projects connected by multiple relational paths, forming a deep understanding of the project's structured attributes [19–20]. Meanwhile, unstructured text related to projects, such as user reviews and official descriptions, contains nuanced, dynamic, and emotionally rich semantics, directly reflecting users' subjective perceptions of project functions and experiences, and is key to understanding the project's potential characteristics and user concerns. To extract deep semantics from unstructured text, the contextual representation capabilities of the BERT pre-trained model are used to encode the review text. Given a text sequence S , after BERT encoding, each word w_i can obtain a context-related vector representation $x_i \in R^h$, containing rich syntactic and preliminary semantic information. To obtain a global semantic representation of the document and capture long-distance dependencies and the distribution of important information, a Bi-LSTM is used for semantic aggregation. It consists of forward and backward LSTMs, processing the sequence from two directions respectively.

$$\vec{h}_i = LSTM_{forward}(x_i, \vec{h}_{i-1}) \quad (2)$$

$$\overleftarrow{h}_i = LSTM_{backward}(x_i, \overleftarrow{h}_{i+1}) \quad (3)$$

For each position i , the forward hidden state $\vec{h}_i$ and the backward hidden state $\overleftarrow{h}_i$ are concatenated to obtain a comprehensive representation $h_i = [\vec{h}_i; \overleftarrow{h}_i]$ containing complete contextual information.

To demonstrate the differences in the contribution of different parts of the text to the document's semantics, an attention mechanism is included to aggregate the hidden states, resulting in a document-level semantic vector. First, the importance score for each position is calculated, and then a weighted sum is obtained to produce the document vector. The Bi-LSTM attention mechanism constructed in this paper can dynamically focus on key sentiment words during semantic aggregation. With the help of attention weight distribution, the model can identify high-contribution words such as “battery life” and “performance” and assign them high weights, demonstrating the user's focus dimension, as shown in Figure 1.

Figure 1 presents positive, negative, and neutral comments. The peaks of the curves correspond to the keywords on which the model focuses, such as “screen” for positive comments, “lag” for negative comments, and “appearance” for neutral comments. Keyword annotation not only verifies that the attention mechanism can capture the emotional semantic focus, but also indicates the model's accurate grasp of user intent. The stacked area plot visually demonstrates the differences in the model's focus under different emotional tendencies through different regions.

Currently, the embedded vectors of structured project knowledge and the deep semantic vectors of unstructured text have been obtained. However, since these originate from different information sources and modeling methods, direct concatenation or simple weighting makes it difficult to

achieve complementarity and semantic depth alignment [21–23]. Therefore, a cross-modal attention network is designed to achieve dynamic alignment and deep fusion of the two modalities. This network uses the text semantic vector d as the query and the set of embedded vectors $\{e_1, e_2, \dots, e_m\}$ of a group of entities related to the project (such as the project and its key attribute entities) as keys and values. By calculating the similarity between the query and each key, the attention distribution is obtained. This distribution reveals the strength of the association between the semantic focus of the text and different entity concepts in the knowledge graph. The attention weight β_j is calculated as follows:

$$\beta_j = \frac{\exp(\text{score}(d, e_j))}{\sum_{k=1}^m \exp(\text{score}(d, e_k))} \quad (4)$$

$\text{score}(d, e_j)$ can employ scaled dot product attention or other similarity functions. Based on the calculated attention weights, the knowledge entity vectors are weighted and summed to obtain a knowledge context vector highly relevant to the current text semantics. Finally, the text semantic vector and the knowledge context vector are integrated through a fusion-gating mechanism to generate the project's final, unified deep semantic vector.

$$g = \sigma(W_g[d; c] + b_g) \quad (5)$$

$$v = g \odot d + (1 - g) \odot c \quad (6)$$

σ is the sigmoid function, W_g and b_g are learnable parameters, $\odot$ represents element-wise multiplication, and g is a gating vector that controls the flow of information. This approach combines the representation of the precise structured semantics of the knowledge graph with the nuanced contextual and emotional semantics of the comment text. Furthermore, it utilizes an attention mechanism to achieve dynamic alignment and adaptive weighting of the two semantic modalities, constructing a deep, rich, and consistent project semantic space.

2.2 Time-Aware User Dynamic Interest Coupling Modeling

After constructing a deep, unified semantic space, another key aspect of optimizing the coupling of intelligent recommendation models is the accurate capture of the dynamic changes in user interests. User interests have typically been recorded as a static vector through classical user interest modeling strategies, which combine individual interaction items and derive an average of all items in the pool [24–25]. These methods ignore both the changing nature of user interests over time and the connection between the user's interaction behavior(s) with them, resulting in recommendations that do not accurately match the user's current preferences. In response to this limitation, this paper proposes a new time-sensitive and dynamic user interest coupling modeling module, based on deep semantic understandings of user dynamic interests. The modeling module consists of three components: sequent behavior encoding, interest-semantic coupling attention, and a cascaded design of interest

evolution graph networks. The proposed modeling module will dynamically model the evolution of user interest, provide rich semantic coupling between user interests, incorporate higher-level patterns of user interest(s), and produce a time-updated user interest vector that most accurately reflects the current state of the user's preferences — resulting in an upgrade from traditional to dynamic user interest models.

The coding of the behavior chain pattern is the first step in the process of converting user-based historic data, and providing a basic behavioral framework based on time and order. Utilizing the input user 'u', the user history will be organized based on interaction timestamps. Subsequently, an ordered activity list is produced that has the following structure: $S_u = (i_1, i_2, \dots, i_T)$ (where T is the sequence length and i_t is the interaction item at time t). At the same time, to create deep semantic representations by constructively integrating both the structured knowledge and semantics of a particular comment-based/text description comment, from the single item user-based timeline sequence of activities we create deep semantic vectors for each activity in S_u , using a unified deep semantic representation module to perform an alignment for each of these single-user based timeline activities into a deep semantic representation. Finally, to encode the temporal dynamics and contextual dependencies within that user based behavioral pattern representation, the overall activity-based deep semantic vectors will then be input for processing by the activity sequence encoder built using the GRU-based unit. GRUs are lightweight recurrent neural networks capable of learning temporal dependencies and mitigating gradient vanishing [26–27]. For time step t , the GRU calculates the hidden state h_t of the current time step based on the current input v_t and the hidden state $h_{t-1} \in R^d$ of the previous time step. Its update process can be abstractly represented as:

$$z_t = \sigma(W_z v_t + U_z h_{t-1} + b_t) \quad (7)$$

$$r_t = \sigma(W_r v_t + U_r h_{t-1} + b_t) \quad (8)$$

$$\tilde{h}_t = \tanh(W_h v_t + U_h(r_t \odot h_{t-1}) + b_t) \quad (9)$$

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \quad (10)$$

z_t and r_t represent the update gate and reset gate, respectively. The final hidden state sequence can be seen as a preliminary temporal representation of user behavior at each time step, containing both the semantic information of the current interaction item and incorporating historical behavior patterns. In the temporally aware user dynamic interest coupling modeling module, GRU is the core of sequence behavior encoding, relying on the update gate and reset gate to dynamically model the evolution of user interests [28–29]. To intuitively present the mechanism by which GRU captures the dynamic changes in user interests, three consecutive time steps of interaction behavior of a typical user were selected, and the numerical change curves of the update gate, reset gate, and candidate hidden states were plotted. The corresponding interaction item category (such as mobile phone, computer, tablet) is labeled below with each time step in the figure, as shown in Figure 2.

Figure 2 depicts the internal mechanism of the GRU in the time-series-aware user dynamic interest coupling

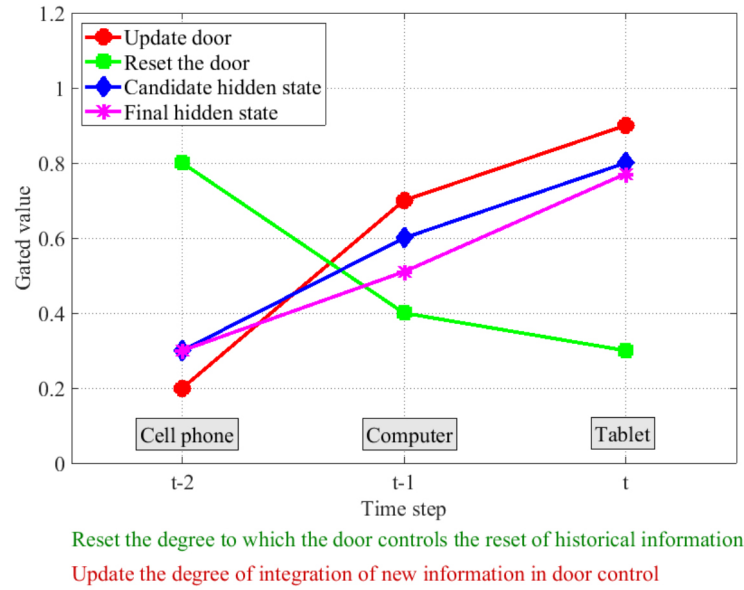


Figure 2 Numerical variation curves of gating mechanism and interest evolution in GRU-based temporal behavior coding.

modeling module. The typical user's interaction behavior at three consecutive time steps (t-2, t-1, and t) involves three items: "phone", "computer", and "tablet", respectively. By constructing numerical sequences of update gates, reset gates, candidate states, and final hidden states, the process of GRU controlling information flow is demonstrated: at time t-2, the update gate is 0.2 (low) and the reset gate is 0.8 (high), indicating that the model retains less historical information and focuses on the current "phone" feature; at time t-1, the update gate increases to 0.7 and the reset gate decreases to 0.4, balancing the fusion of historical (phone) and new information (computer); at time t, the update gate is high and the reset gate is low, the model retains a large amount of historical information, and the user's interest shifts to "tablet".

The preliminary time-series representation above is based only on the user's own interaction sequence and is not strongly coupled with the global semantic knowledge of the item space. To more proactively integrate the deep semantics of items into interest modeling, an interest-semantic coupling attention layer is designed. This layer treats the unified deep semantic vector set of items generated by the previous module as a global "semantic memory unit". When modeling a user's instantaneous interest at a given moment, it not only uses the current behavioral representation but also retrieves relevant semantic information from semantic memory through an attention mechanism [30–31]. The hidden state h_t at the current time step is used as the query, and an attention score is calculated between it and the item vector v_j in the semantic memory unit. This score reflects the user's attention to the semantics of different items, expressed as:

$$a_{t,j} = \frac{\exp(h_t^T W_{att} v_j)}{\sum_{k=1}^{|I|} \exp(h_t^T W_{att} v_k)} \quad (11)$$

W_{att} is a learnable bilinear transformation matrix. Based on the calculated attention weights, the entire semantic memory unit is weighted and summed to obtain the global

semantic context vector $m_t = \sum_{j=1}^{|I|} a_{t,j} v_j$, which is closely related to the current time step t. This context vector can be understood as the dynamically focused interest from the global item semantic space, targeting the user's current behavior. Subsequently, the preliminary behavioral representation is fused with the global semantic context to form the user's instantaneous interest vector p_t at time step t.

$$p_t = \tanh(W_f [h_t; m_t] + b_f) \quad (12)$$

With this combined form of attention, the user's instantaneous interests p_t are derived from a much stronger relationship with their own history of interactions and the overall semantic space of the project. This allows us to be able to accurately capture the user's subtle interest trends at that moment in time.

Although the instantaneous interest vector sequence $\{p_1, p_2, \dots, p_T\}$ can indicate dynamic changes in interest, it is only an independent representation of interest at a given point in time, without considering the complex high-order evolutionary patterns between interest points. User interest evolution is non-linear, exhibiting jumps, repetitions, and fusion. To capture high-order patterns, an interest evolution graph network is introduced, treating instantaneous interest vectors at different time steps as nodes in the graph, and connecting these nodes with directed edges in chronological order. To capture long-distance dependencies, additional edges are built based on interest similarity or time intervals, forming a fully connected temporal interest graph. On this graph, nodes update their representations by aggregating neighbor information through a graph neural network, capturing high-order transition patterns. A graph attention network is used as the information aggregation backbone because it can assign different weights to neighbor nodes of different importance. The updated representation a'_t of node a_t is obtained by aggregating its neighbor information:

$$a'_t = \sigma \left(\sum_{s \in N_t} \beta_{t,s} W_g a_s \right) \quad (13)$$

After information transmission and aggregation through an L-layer graph network, the final time step node representation a'_T can be considered as the final interest representation that fully considers the evolution of user interests and higher-order patterns. This is used as a preference vector representing the user's current dynamic interests.

2.3 Multi-Objective Collaborative Adversarial Coupling Optimization

Although some traditional recommendation models rely largely on one main outcome measure and perform well under those conditions [32–33], there is considerable bias associated with using just one metric, particularly because of how heavily the accuracy is relied on when popular items are recommended, creating the “Matthew Effect”. Because of the “Matthew Effect”, fewer long-term items are recommended to users than there are recommended to those who tend to be successful and there are fewer recommendations for users who are unsuccessful, ultimately reducing the satisfaction level of users in the long run and impacting the overall health of the platform's ecosystem [34–35]. Creating a multi-objective collaborative adversarial coupling optimization module for use within the methods developed in this paper allows for recommendations based on both semantic user interest and dynamic user interest through an end-to-end modeling approach across multiple objectives such as accuracy and diversity. This is accomplished through multiple task output/heads designed for parallel execution for recommending accuracy as well as for recommending diversity, the introduction of adversarial training, and the development of coupled loss functions for achieving collaboration across multiple objectives.

The architecture of the multi-task prediction head serves as the starting point to provide the capabilities for the multi-objective collaborative efforts in this module. Typical single-task recommendation models utilize only one prediction layer, which produces one probability of a user interacting with an item. While succinct and efficient, single-task models have restrictions on their ability to adequately leverage the data supervision signals available to them, as well as their ability to sufficiently control the multi-dimensional quality of recommendations produced by the model. To solve this design issue, a parallel multi-task prediction head design is developed in this research. Based on sharing the underlying semantic representation and dynamic interest vector, it derives multiple parallel neural network branches, each optimizing a specific recommendation task. Three parallel prediction heads are set up: the first is the main task prediction head that minimizes the prediction error of click or purchase behavior, which is the core indicator for measuring recommendation accuracy. It takes the user's dynamic interest vector p_u and the deep semantic vector v_i of the candidate item i as input, and calculates the interaction probability through a multilayer perceptron (MLP):

$$\hat{y}_{ui} = \sigma (MLP_{main} ([p_u; v_i])) \quad (14)$$

The second is the rating prediction auxiliary task head that generates predictions for item ratings while simultaneously refining the representation learning of user preferences that are created through user's explicit ratings on items. The input for this task head takes p_u and v_i and provides an output of predicted item rating $\hat{r}_{ui}$ using the other MLP structure to fit the regression problem. This type of supervised task may offer additional supervision signals that will provide further context for the way it builds and uses the understanding of user preferences to provide better generalizable and interpretable results. Another auxiliary head for the multi-task model is the diversity prediction head. It produces the probable distribution of diverse item categories a user may want based on their interests at the time. The output of this head is the probability distribution of different item categories (represented by $\hat{c}_u$), which serves to help the recommendation engine consider both the user's current interests along with other potentially interesting categories when generating recommendations, thereby producing more diverse recommendations as well. The synergy of these three tasks during training has been demonstrated using the convergence curves of the multi-task loss function for three separate tasks in terms of training epoch number. During the training process, Figure 3 illustrates this synergetic optimization effect.

In order to illustrate training stability, Figure 3 displays the decline curves of the losses for the primary task, rating prediction, and diversity prediction over 100 training rounds, creating confidence interval bands. The loss for the main task decreased from 2.91 to 0.58, optimizing click/purchase prediction accuracy; the loss for rating prediction decreased from 2.07 to 0.38, introducing explicit rating signals to refine user preferences; and the loss for diversity prediction decreased from 1.53 to 0.45, demonstrating significant effectiveness in exploring category diversity.

The sole reliance on implicit constraints for auxiliary tasks cannot adequately alleviate the popular item bias problem. Therefore, an adversarial regularization training mechanism is introduced. The core of this mechanism is to have the adversarial discriminator play a game with the main recommendation model (generator), forcing the recommendation results generated by the main model to “deceive” the discriminator, making it unable to distinguish whether the recommended items come from the distribution of popular items based on the user's historical interactions or from the distribution of non-popular but relevant items [36–37]. First, a popular item set I_{pop} is defined based on interaction frequency. The adversarial discriminator D is a binary classifier. Its input is either a vector of recommended items generated by the main model or a vector of items sampled directly from the popular item set. It outputs a probability value indicating whether the input item comes from the popular set. The optimization objective of discriminator D is to maximize its discriminative power, and its loss function uses the standard binary cross-entropy.

$$L_{adv}^D = -E_{i \sim I_{pop}} [\log D(v_i)] - E_{i \sim G(u)} [\log (1 - D(v_i))] \quad (15)$$

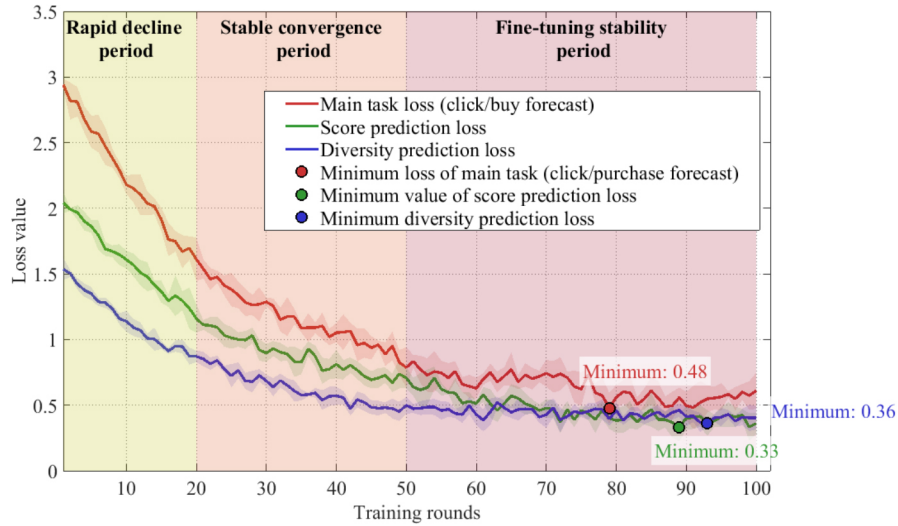


Figure 3 Convergence curve of multi-task learning loss function.

$G_{(u)}$ represents the distribution of recommended items generated by the main recommendation model for user u . When optimizing the accuracy of the main task, the main recommendation model G also needs to “deceive” the discriminator D , causing D to classify its generated recommended items as unpopular items. Hence, an adversarial regularization loss is introduced into the main model G , in order to minimize the probability of generated items being identified as popular items, i.e., maximizing the probability of them being identified as unpopular items.

$$L_{adv}^G = -E_{i \sim G_{(u)}} [\log(1 - D(v_i))] \quad (16)$$

The primary model uses a minimax approach to attempt to make an accurate recommendation to the user while focusing on less frequently requested long-tail items to confuse the discriminator. By using this method and the optimization objective, it encourages the model to develop findings of lesser-known, but relevant long-term items, thus enhancing the diversity and novelty of the recommendations while still achieving accuracy. The coupling loss function, which is a weighted sum of the main task loss, the auxiliary task loss, and the adversarial regularization loss, is used to integrate the various optimization objectives into a single training framework. A major task loss with the cross_entropy loss function as its primary function and an auxiliary task with a rating prediction loss using a mean_squared error is used for predicting clicks and purchases. The diversity prediction is used as an auxiliary task to compare the predicted click and purchase categories to the click and purchase categories that were interacted with; the cross_entropy for this auxiliary task is also used as the loss function. The adversarial regularization loss is added to the combination of total losses to create a total loss function defined as follows:

$$L_{total} = \lambda_1 L_{main} + \lambda_2 L_{rating} + \lambda_3 L_{diversity} + \lambda_4 L_{adv}^G \quad (17)$$

λ_1 , λ_2 , λ_3 , and λ_4 are hyperparameters used to balance the relative importance of different loss terms in the overall objective.

3. EXPERIMENT AND RESULTS ANALYSIS

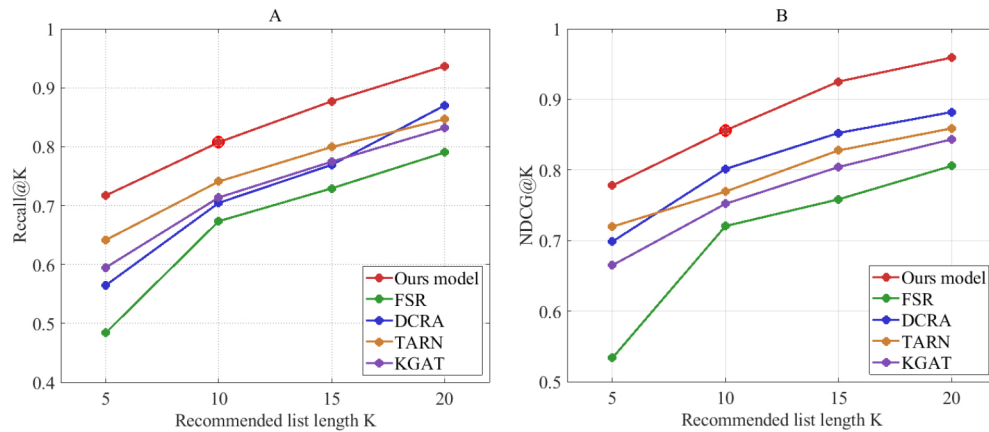
3.1 Experimental Setup and Dataset

The research presented in this paper is performed entirely on a common hardware/software platform to achieve the same type of results for all experiment types as well as provide identical experiment conditions. The hardware comprised an Intel Xeon Gold 6248R processor (3.0GHz, 24 cores), 128GB system RAM, and an NVIDIA Tesla V100 32GB GPU (Graphics Processing Unit) to improve both speed of training and inference of deep learning models. The software used a single Ubuntu 20.04 installation with the PyTorch 1.10.0 library running under Python 3.8.10 for all model training. All models were optimized using the Adam optimizer with an initial learning rate of 0.001, batch size of 256, and a 64-dimensional embedding vector. The hyperparameters used to create the baseline model were identical to those reported in the original paper and fine-tuned during validation only.

To ensure a thorough evaluation of model performance, four publicly available recommender system datasets were selected. The first of these datasets is the Amazon Electronics dataset, which consists of real user reviews (containing category ratings, reviews, and item metadata) for electronic products from the Amazon Electronics site; this dataset has rich text and a long tail distribution. The second dataset is MovieLens-1M (Million), a classical movie recommendation dataset consisting of 1 million user rating records, and is widely used to go with algorithms based on collaborative filtering and sequence-based recommendation models. The third dataset is Yelp, a business review site with restaurant and other merchant review and interaction records; the data is sparse but filled with rich text. Gowalla, a location-based social network check-in dataset, records user check-in behavior and timestamps, making it suitable for evaluating the interest evolution modeling capabilities of time-series-aware recommendation models. These four datasets have a range of different characteristics, appropriate for comprehensively

Table 1 Statistical information of experimental dataset.

Indicator	Amazon Electronics	MovieLens-1M	Yelp	Gowalla
Number of users	192403	6040	1968703	107092
Number of items	63001	3706	209393	1280969
Number of interactions	1689188	1000209	8021122	6442890
Sparsity	99.86%	95.53%	99.98%	99.99%
Comment text	Yes	No	Yes	No
Timestamp	Yes	Yes	Yes	Yes
Field	Electronic commerce	Movie	Local life	Social sign-in

**Figure 4** Accuracy comparison of different models under different recommendation list lengths. Figure 4A: Recall@K. Figure 4B: NDCG@K.

testing model generalization and robustness. Statistical information is shown in Table 1.

As shown in Table 1, Amazon Electronics has 192,403 users, 63,001 items, and 1,689,188 interactions, with a sparsity of 99.86%. It contains rich comment text, making it suitable for testing the model's textual semantic understanding and long-tail recommendation performance. MovieLens - 1M is smaller, with 6,040 users and 3,706 items, and a sparsity of 95.53%. As a classic benchmark, it can verify the model's accuracy under dense data. Yelp is larger, with 1,968,703 users and a sparsity of 99.98%, containing comment text, and can test the model's robustness in large-scale sparse scenarios. Gowalla has 1,280,969 items, a sparsity of 99.99%, and includes timestamps, making it suitable for evaluating the performance of time-series-aware models in interest evolution modeling.

3.2 Results Analysis

3.2.1 Comparative Analysis of Recommendation Accuracy

To comprehensively evaluate the performance of the proposed model under different recommendation list lengths, it is compared with FSR (Fuzzy Signature-based Recommendation), DCRA (Deep Collaborative Recommendation with Attention), TARN (Neural Time-Aware Recommendation Network), and KGAT (Knowledge Graph Attention Network). Two primary metrics are used to measure the overall accuracy of the results were Recall@K and Normalized

Decline Cumulative Gain (NDCG@K). Recall assesses the total number of items that have been found by the model that are actually good for the user as compared to how many total items are contained in the recommendations. NDCG @K measures user satisfaction by determining the project's ranking in the list. The performance of each model at four truncation lengths ($K = 5, 10, 15$, and 20) is compared. The results are shown in Figure 4.

Figure 4 shows the performance of different models under different recommendation list lengths. In Figure 4A, for the Recall@K metric, when $K=5$, the proposed model achieves a recall of 0.717, which is higher than FSR (0.484), DCRA (0.565), TARN (0.642), and KGAT (0.595). When K increases to 20, the recall of the proposed model rises to 0.937, while KGAT only achieves 0.832. In Figure 4B, for the NDCG@K metric, when $K = 5$, the proposed model's NDCG@5 is 0.778, which is higher than other models. When $K = 20$, the proposed model's NDCG@20 reaches 0.959, while the FSR is only 0.806. This advantage stems from the model design: a heterogeneous multi-source information deep semantic unified representation module, which uses knowledge graph embedding and BERT+Bi-LSTM comment text encoding to construct a unified space integrating structured and unstructured semantics, accurately understanding the relationship between item attributes and user perception; a time-series-aware user dynamic interest coupling modeling module, which captures behavioral temporal dependencies through GRU sequence encoding, retrieves global semantic memory through interest-semantic coupling attention, and mines higher-order interest evolution through graph attention networks to accurately track the

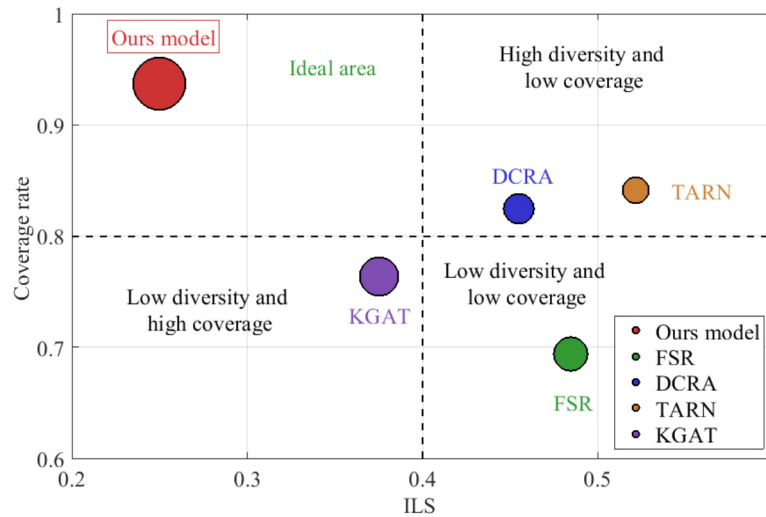


Figure 5 Bubble chart of recommendation diversity and coverage for each model.

evolution path of user interests; and a multi-objective collaborative optimization module, which improves diversity and long-term coverage capabilities through adversarial training while ensuring accuracy.

3.2.2 Evaluation of Recommendation Diversity and Coverage

While accuracy is crucial in recommender model evaluation, simply pursuing it can easily lead to recommendation convergence, forming a “Matthew effect”, where popular items are repeatedly recommended while long-tail items struggle to gain exposure. An experiment was set up to evaluate the proposed model’s advantages through diversity by comparing two key indicators: intra-list similarity (ILS) and coverage. ILS computes the average similarity among all items on an individual recommendation list; therefore, lower ILS numbers correlate to greater diversity among recommendations. Coverage indicates the percentage of all possible unique items that can be recommended. The greater the percentage of coverage, the better the recommendation system will be at discovering long-tail items. An ideal multi-objective optimization model should balance low ILS and high coverage. This paper compares the proposed method with other methods. The results are shown in Figure 5.

As shown in Figure 5, the ILS metric of the proposed model is only 0.25, far lower than FSR (0.484), DCRA (0.455), TARN (0.522), and KGAT (0.375), indicating that the recommendation results vary greatly in terms of item attributes and categories, effectively avoiding homogenization. In terms of coverage, the proposed model reaches 0.934, significantly higher than FSR (0.694), DCRA (0.825), TARN (0.841), and KGAT (0.764), indicating that it can uncover more long-tail items and provide exposure for niche products. The proposed model achieves the highest coverage and optimal diversity, located in the ideal region in the upper left corner of the figure (ILS < 0.4 and Coverage > 0.8). This multi-dimensional collaborative optimization benefits from the model design: the heterogeneous multi-source informa-

tion deep semantic unified representation module integrates knowledge graphs and comment text semantics, helping the model understand the intrinsic relationship between long-tail and popular items, avoiding the neglect of niche items; the time-series-aware dynamic interest coupling modeling module retrieves diverse items through interest-semantic coupling attention, broadening the recommendation scope; the multi-objective collaborative adversarial coupling optimization module combats popularity bias through adversarial training, achieving dual optimization of ILS and coverage. The strong coupling mechanism enables the proposed model to break through the traditional diversity bottleneck.

3.2.3 Robustness Verification in Cold Start Scenarios

In practical applications of recommendation systems, the cold start problem is a key bottleneck restricting performance improvement. New users lack historical interaction data, and new items lack exposure feedback information, which are its main manifestations. Traditional collaborative filtering methods experience a sharp performance drop in cold start scenarios. To verify the adaptability of the model in cold start scenarios, the proposed method is compared with other methods. Cold start users with no more than 3 interactions and cold start items with no more than 3 exposures are selected, and the NDCG@10 performance distribution of different models in the two scenarios is evaluated. The results are shown in Figure 6.

As shown in Figure 6, the performance and stability of different models vary significantly in the cold start scenario. In Figure 6A, the average NDCG@10 of the proposed model reaches 0.95, far exceeding DCRA (0.824), TARN (0.811), KGAT (0.821), and FSR (0.713). Furthermore, the proposed model has a concentrated performance distribution, small standard deviation, and no outliers; the baseline models all show abnormally low values, indicating that some models are completely ineffective for some cold start users. In the cold start project scenario (Figure 6B), the average value of the proposed model reaches 0.968, higher than other

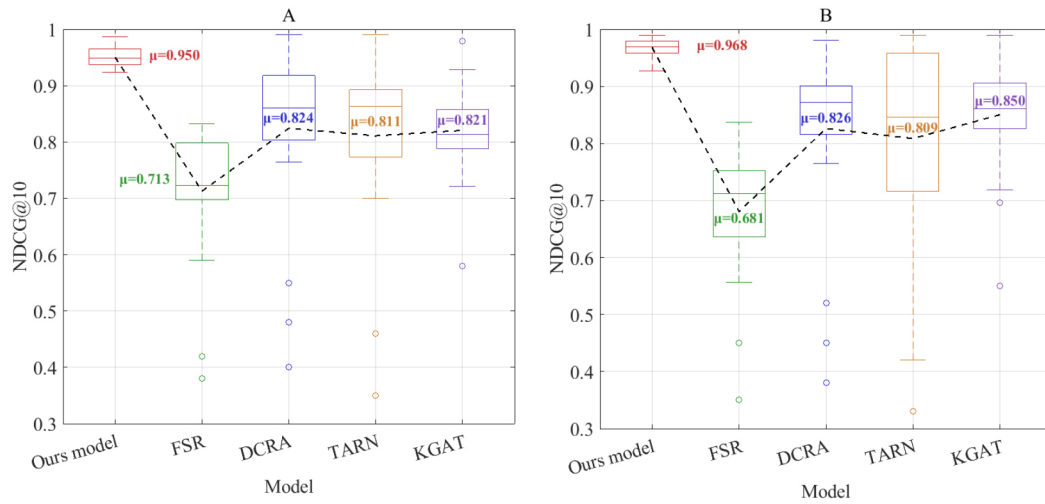


Figure 6 NDCG@10 Performance distribution of each model in cold start scenario. Figure 6A: Cold start user scenario. Figure 6B: Cold start project scenario.

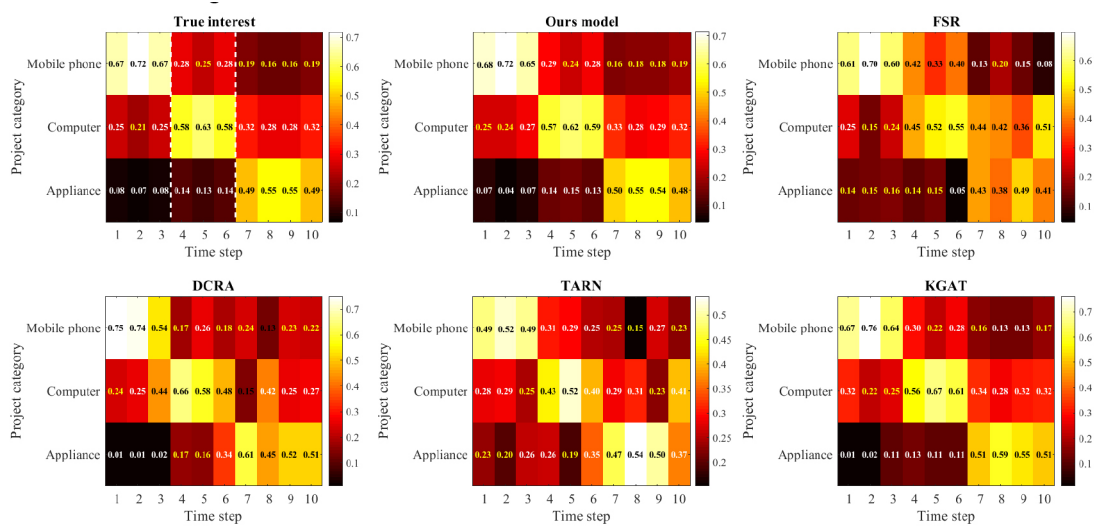


Figure 7 Comparison of heatmaps predicting user interest evolution trajectory.

models, demonstrating excellent stability and no outliers; the baseline models, however, fluctuate wildly, all showing abnormally low values. The advantage of the proposed model in the cold start scenario stems from its design: a heterogeneous multi-source information deep semantic unified representation module, which uses knowledge graph embedding and comment text encoding to construct semantic vectors for cold start projects independent of the number of interactions, helping the model understand the project's meaning; and an adversarial optimization module to prevent the model from favoring popular projects when data is scarce.

3.2.4 User Interest Evolution Capture Ability Test

To verify intuitively the advantages of being able to capture a user's change of interest over time, an interest evolution experiment was conducted. This experiment simulated a user's typical interest trajectory for ten consecutive time steps. Each time step was then compared to the actual distribution of a user's interest within each model versus

the model's predicted distribution of that time step. The horizontal axis of the data represents each time step, while the vertical axis represents the item category (i.e., mobile phone, computer, or home appliance). The color intensity of the cells in the data represents the predicted interest intensity of the user in a particular item category. By conducting this experiment and comparing the actual and predicted distributions for each time step, it evaluated the model's performance at capturing the overall trend of user interest evolution. The results of this experiment are shown in Figure 7.

As shown in Figure 7, the proposed model demonstrates a significant advantage in capturing the dynamic evolution of user interests. The actual interest distribution reveals that user interests evolve through three stages: $t = 1-3$ focusing on mobile phones (peak value 0.72), $t = 4-6$ shifting to computers (peak value 0.63), and $t = 7-10$ evolving to home appliances (peak value 0.55). The distribution predicted by the proposed model closely matches the actual distribution, with peak values of 0.72 for mobile phones, 0.62

Table 2 Comparison of module contribution measurements in ablation experiments.

Model variant	NDCG@20	ILS?	Coverage rate?	Cold start user scenario NDCG@10
Complete model	0.959	0.25	0.934	0.95
Remove the knowledge map	0.897	0.285	0.878	0.858
Remove comment text	0.876	0.292	0.861	0.823
Remove coupling attention	0.912	0.272	0.905	0.885
Remove the graph network	0.931	0.264	0.915	0.912
Remove confrontation optimization	0.95	0.328	0.842	0.932

for computers, and 0.55 for home appliances at each stage, and small numerical errors at each time step.

3.2.5 Module Ablation Experiment and Contribution Analysis

An ablation study was conducted to determine the extent to which each core part of the proposed model contributed to its overall performance. The study progressively removed significant components from the model to quantify how each core module impacted the recommendation system's accuracy, diversity, and ability to overcome cold-starts. Based on the complete model, five variants were constructed: removing the knowledge graph embedding module to verify the enhancing effect of structured semantics on item representation; removing the comment text encoding module to evaluate the importance of unstructured semantics in understanding user intent; removing the interest-semantic coupling attention module to examine the gain of global semantic memory on dynamic interest modeling; removing the interest evolution graph network module to analyze the necessity of capturing higher-order interest evolution; and removing the adversarial optimization module to examine the moderating effect of multi-objective collaboration on diversity and long-tail recommendation. Experiments were conducted using NDCG@20, ILS, and coverage rate to measure long-tail coverage capability, and evaluated the NDCG@10 decay in a cold-start user scenario. The results are shown in Table 2.

As shown in Table 2, in terms of accuracy, removing the comment text module reduced NDCG@20 to 0.876, demonstrating the core role of unstructured semantics in understanding user intent; removing the knowledge graph module reduced it to 0.897, confirming the enhancing effect of structured semantics on item representation; the removal of coupled attention and graph networks reduced it to 0.912 and 0.931, respectively, indicating that global semantic memory and higher-order evolutionary capture have progressive gains for interest modeling. Regarding diversity, removing the adversarial optimization module increased ILS from 0.25 to 0.328, while coverage decreased from 0.934 to 0.842, proving that adversarial training is key to balancing accuracy and diversity. These data stem from the model design: heterogeneous multi-source information deep semantic unified representation, through complementary fusion of knowledge graph and comment text, constructing a semantic foundation independent of the number of interactions for cold start; time-aware dynamic interest coupling modeling, achieving accurate interest tracking through attention mech-

anisms and graph networks; and adversarial optimization, while ensuring accuracy, explores long-tail items, achieving dual optimization of ILS and coverage.

4. CONCLUSION

The aim of this research was to identify the limitations of conventional intelligent recommendation systems within the realm of large-scale data. In particular, these limitations arise from their inability to capture dynamic changes in user interest and their lack of strong semantic meaning. A new approach was proposed for the development of a deep semantic-based recommendation model, which utilizes a unified deep semantic representation that has been constructed by incorporating heterogeneous multi-source data. It consists of both structured and un-structured data and creates a very rich semantically consistent, unified deep semantic space that serves as a solid basis for accurate recommendations. A time-aware user dynamic interest coupling modelling component will also be developed to enable the effective understanding of users' dynamic interests, thus enabling the transition from being static to being dynamic. A mechanistic approach to optimizing for multiple objectives using collaborative adversarial coupling was developed with the intent of improving both the diversity and novelty of recommendations while maintaining accuracy, which addresses the 'Matthew Effect'. In terms of recommendation accuracy, Recall@20 reaches 0.937; NDCG@20 reaches 0.959, outperforming all compared models. In diversity evaluation, ILS is as low as 0.25, with a coverage rate as high as 0.934, alleviating the problem of recommendation homogenization. In the cold start scenario, the mean NDCG@10 values for cold start users and projects reach 0.95 and 0.968, respectively, demonstrating strong robustness and generalization ability. In conclusion, the combined algorithm proposed in this study incorporates high-level understanding of both the object being recommended and the requester's dynamic engagement with that object, thus solving the main problems of intelligent recommendation systems by improving accuracy, diversity and ability to adapt to cold starts, thereby providing better quality recommendation services to users.

FUNDING

This project was supported by General Research Projects in Philosophy and Social Sciences at Jiangsu University (Project number: 2024SJYB1378). This project was funded

by the Major Project of Philosophy and Social Science Research in Jiangsu Higher Education Institutions (Project number: 2025SJZDSZ030).

REFERENCES

- Mu, J. (2023). The application and effect of intelligent marketing technology and personalized recommendation system in e-commerce. *Frontiers in Computing and Intelligent Systems*, 5(1), 1–4. DOI:10.54097/fcis.v5i1.11533
- Ejjami, R. (2024). Enhancing user experience through recommendation systems: A case study in the ecommerce sector. *International Journal for Multidisciplinary Research*, 6(4), 6–25. DOI:10.36948/ijfmr.2024.v06i04.24598
- Khoubiri, N., Farhaoui, Y., & El Allaoui, A. (2023). Design and analysis of a recommendation system based on collaborative filtering techniques for big data. *Intelligent and Converged Networks*, 4(4), 296–304. DOI:10.23919/ICN.2023.0024
- Zhong, M., & Ding, R. (2022). Design of a personalized recommendation system for learning resources based on collaborative filtering. *International Journal of Circuits, Systems and Signal Processing*, 16(1), 122–131. DOI:10.46300/9106.2022.16.16
- Ko, H., Lee, S., Park, Y., & Choi, A. (2022). A survey of recommendation systems: recommendation models, techniques, and application fields. *Electronics*, 11(1), 141–189. DOI:10.3390/electronics11010141
- Wu, L., He, X., Wang, X., Zhang, K., & Wang, M. (2022). A survey on accuracy-oriented neural recommendation: From collaborative filtering to information-rich recommendation. *IEEE Transactions on Knowledge and Data Engineering*, 35(5), 4425–4445. DOI: 10.1109/TKDE.2022.3145690
- Wang, Z. (2023). Intelligent recommendation model of tourist places based on collaborative filtering and user preferences. *Applied Artificial Intelligence*, 37(1), 2203574–2203597. DOI:10.1080/08839514.2023.2203574
- Song, Y., & He, Y. (2023). Toward an intelligent tourism recommendation system based on artificial intelligence and IoT using Apriori algorithm: Y. Song, Y. He. *Soft Computing*, 27(24), 19159–19177. DOI:10.1007/s00500–023-09330–2
- Widayanti, R., Chakim, M. H. R., Lukita, C., Rahardja, U., & Lutfiani, N. (2023). Improving recommender systems using hybrid techniques of collaborative filtering and content-based filtering. *Journal of Applied Data Sciences*, 4(3), 289–302. DOI:10.47738/jads.v4i3.115
- Parthasarathy, G., & Sathiyadevi, S. (2023). Hybrid recommendation system based on collaborative and content-based filtering. *Cybernetics and Systems*, 54(4), 432–453. DOI:10.1080/01969722.2022.2062544
- Du, L., & Xu, Y. (2024). Development strategy of online English teaching based on attention mechanism and recurrent neural network recommendation method. *International Journal of Data Mining and Bioinformatics*, 28(2), 140–155. DOI:10.1504/IJDMB.2024.137748
- Liu, H., Zhang, Y., Li, P., Qian, C., Zhao, P., & Wu, X. (2023). DeepCPR: deep path reasoning using sequence of user-preferred attributes for conversational recommendation. *ACM Transactions on Knowledge Discovery from Data*, 18(1), 1–22. DOI:10.1145/3610775
- Zhou, R., Qian, H., Zhang, J., Liu, C., Wan, J., Ren, Y., ... & Zhang, S. (2021). Self-attention mechanism enhanced user interests modeling for personalized recommendation services in cyber-physical-social systems. *IEEE Transactions on Network Science and Engineering*, 9(6), 3835–3846. DOI:10.1109/TNSE.2021.3093782
- Yu, S., Guo, M., Chen, X., Qiu, J., & Sun, J. (2023). Personalized movie recommendations based on a multi-feature attention mechanism with neural networks. *Mathematics*, 11(6), 1355–1377. DOI:10.3390/math11061355
- Zhang, Q., Cao, L., Shi, C., & Niu, Z. (2021). Neural time-aware sequential recommendation by jointly modeling preference dynamics and explicit feature couplings. *IEEE Transactions on Neural Networks and Learning Systems*, 33(10), 5125–5137. DOI:10.1109/tnnls.2021.3069058
- Cui, C., Qin, J., & Ren, Q. (2022). Deep collaborative recommendation algorithm based on attention mechanism. *Applied Sciences*, 12(20), 10594–10609. DOI:10.3390/app122010594
- Farshidi, S., Rezaee, K., Mazaheri, S., Rahimi, A. H., Dadashzadeh, A., Ziabakhsh, M., ... & Jansen, S. (2024). Understanding user intent modeling for conversational recommender systems: a systematic literature review: S. Farshidi et al. *User Modeling and User-Adapted Interaction*, 34(5), 1643–1706. DOI:10.1007/s11257–024-09398-x
- Siro, C., Aliannejadi, M., & De Rijke, M. (2023). Understanding and predicting user satisfaction with conversational recommender systems. *ACM Transactions on Information Systems*, 42(2), 1–37. DOI:10.1145/3624989
- Jin, Y., Chen, L., Cai, W., & Zhao, X. (2024). CRS-Que: A user-centric evaluation framework for conversational recommender systems. *ACM Transactions on Recommender Systems*, 2(1), 1–34. DOI:10.1145/3631534
- Wu, Y., & Yusof, Y. (2024). Emerging Trends in Real-time Recommendation Systems: A Deep Dive into Multi-behavior Streaming Processing and Recommendation for E-commerce Platforms. *Journal of Internet Services and Information Security*, 14(4), 45–66. DOI:10.58346/jisis.2024.i4.003
- Sharma, S., Rana, V., & Kumar, V. (2021). Deep learning based semantic personalized recommendation system. *International Journal of Information Management Data Insights*, 1(2), 100028–100034. DOI:10.1016/j.jjime.2021.100028
- Feng, S., Cao, B., Xie, Z., Fu, Z., Peng, Z., & Kang, G. (2025). LLMSARec: large language model with semantic alignment for Web service recommendation. *International Journal of Web Information Systems*, 21(1), 37–53. DOI:10.1108/IJWIS-09–2024-0262
- Huijuan S. & Jian S. (2026). Construction of humanoid robot and generative artificial intelligence information system based on deep learning. *Engineering Intelligent Systems*, 34(3).
- Zhu, W. (2021). Topic recommendation system using personalized fuzzy logic interest set. *Journal of Intelligent & Fuzzy Systems*, 40(2), 2891–2901. DOI:10.3233/jifs-189329
- Chiang, J. H., Ma, C. Y., Wang, C. S., & Hao, P. Y. (2022). An adaptive, context-aware, and stacked attention network-based recommendation system to capture users' temporal preference. *IEEE Transactions on Knowledge and Data Engineering*, 35(4), 3404–3418. DOI:10.1109/TKDE.2022.3140387
- Suresh, A., & Carmel Mary Belinda, M. J. (2022). Online product recommendation system using gated recurrent unit with Broyden Fletcher Goldfarb Shanno algorithm. *Evolutionary Intelligence*, 15(3), 1861–1874. DOI:10.1007/s12065–021-00594-x
- Ji, S. An intelligent recommendation model for disclosure of electricity market entities based on association rule mining algorithm. *Electr Eng* 107, 13055–13068 (2025).
- Xia, H., Luo, Y., & Liu, Y. (2021). Attention neural collaboration filtering based on GRU for recommender systems. *Complex & Intelligent Systems*, 7(3), 1367–1379. DOI:10.1007/s40747–021-00274–4
- Bansal, S., & Baliyan, N. (2023). Remembering past and predicting future: a hybrid recurrent neural network based recom-

- mender system. *Journal of Ambient Intelligence and Humanized Computing*, 14(12), 16025–16036. DOI:10.1007/s12652-022-04375-x
30. Zheng, X., Li, X., Chen, Z., Sun, L., Yu, Q., Guo, L., & Luo, Y. (2024). Enhanced self-attention mechanism for long and short term sequential recommendation models. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 8(3), 2457–2466. DOI:10.1109/TETCI.2024.3366771
 31. Akkem, Y., & Biswas, S. K. (2024). Analysis of an intellectual mechanism of a novel crop recommendation system using improved heuristic algorithm-based attention and cascaded deep learning network. *IEEE Transactions on Artificial Intelligence*, 6(5), 1100–1113. DOI: 10.1109/TAI.2024.3508654
 32. Liang, W., Xie, S., Cai, J., Xu, J., Hu, Y., Xu, Y., & Qiu, M. (2021). Deep neural network security collaborative filtering scheme for service recommendation in intelligent cyber-physical systems. *IEEE Internet of Things Journal*, 9(22), 22123–22132. DOI: 10.1109/JIOT.2021.3086845
 33. Tahir, S., Hafeez, Y., Abbas, M. A., Nawaz, A., & Hamid, B. (2022). Smart learning objects retrieval for E-Learning with contextual recommendation based on collaborative filtering. *Education and Information Technologies*, 27(6), 8631–8668. DOI:10.1007/s10639-022-10966-0
 34. Jangid, M., & Kumar, R. (2025). Deep learning approaches to address cold start and long tail challenges in recommendation systems: a systematic review. *Multimedia Tools and Applications*, 84(5), 2293–2325. DOI: 10.1007/s11042-024-20262-3
 35. Shang, F., Zhao, F., Zhang, M., Sun, J., & Shi, J. (2024). Personalized recommendation systems powered by large language models: Integrating semantic understanding and user preferences. *International Journal of Innovative Research in Engineering and Management*, 11(4), 39–49. DOI:10.55524/ijirem.2024.11.4.6
 36. Xia Zhongxiu, Zhang Weiyu, Weng Ziqiang, & Guo Xinchao. (2024). A social recommendation system based on graph attention adversarial networks. *Computer Applications and Software*, 41(8), 289–297. DOI: 10.3969/j.issn.1000-386x.2024.08.042
 37. Yu, S., Duan, M., Wang, K., & Yang, S. (2025). RUP-GAN: A Black-Box Attack Method for Social Intelligence Recommendation Systems Based on Adversarial Learning. *Big Data Mining and Analytics*, 8(4), 820–836. DOI: 10.26599/BDMA.2025.9020002

