

# Application of Data Science and Machine Learning Algorithms in Intelligent Recommendation System

Yuxuan Wang

School of Mathematical Sciences, University of Nottingham Ningbo China, Ningbo, 315100, China

---

## ABSTRACT

Intelligent recommender systems use data science and machine learning algorithms to significantly improve recommendation accuracy and personalization through preprocessing, feature engineering and user profile construction. The study proposes solutions for challenges such as data sparsity and cold start, and experimentally verifies the effectiveness of algorithms such as XGBoost. Meanwhile, the potential of deep learning models and integrated learning in recommendation systems is explored. The study shows that the integration of data science and machine learning algorithms strengthens the data processing and pattern recognition capabilities and promotes the optimization and innovation of recommendation algorithms.

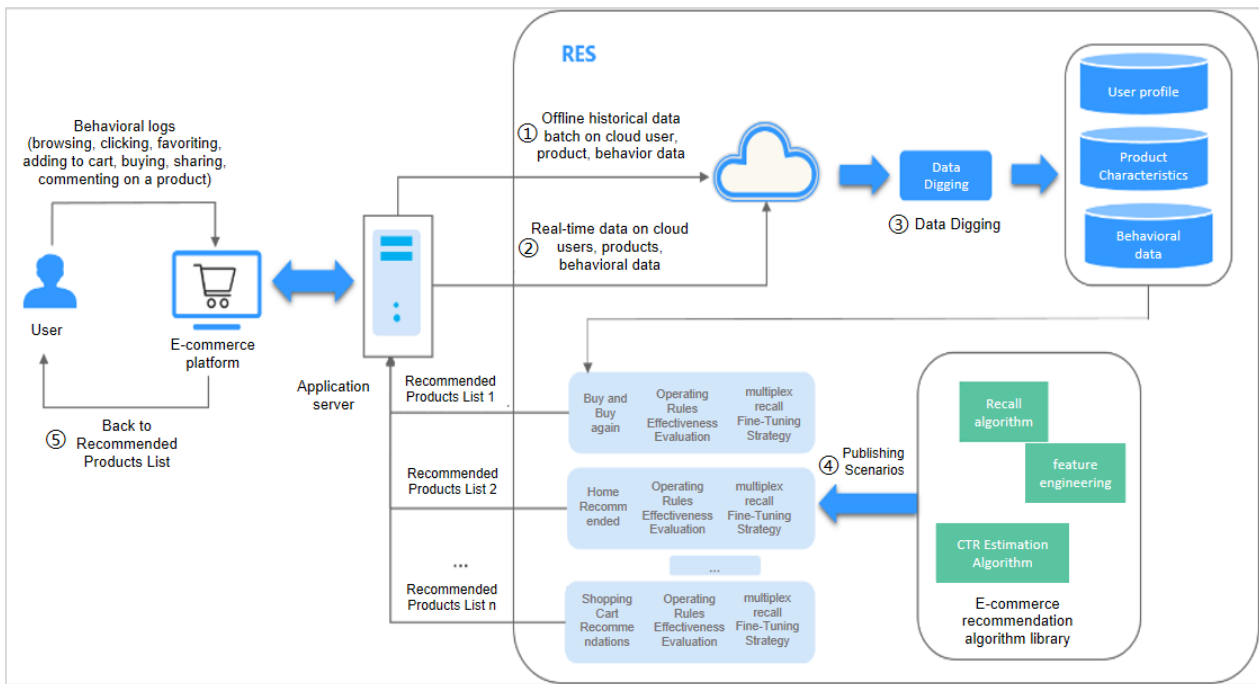
## KEYWORDS

Intelligent recommendation system; Data science; Machine learning; Deep learning; User profiling

---

## 1. INTRODUCTION

With the rapid progress of Internet technology, intelligent recommendation systems are widely used in e-commerce platforms and social media, providing personalized services by analyzing users' historical behaviors and interests and preferences, significantly enhancing user experience and platform stickiness (see Figure 1). Data science and machine learning algorithms play a central role in this process, empowering recommender systems with data processing and pattern recognition capabilities to optimize recommendation accuracy [1]. However, intelligent recommender systems still face challenges such as cold start, data sparsity, and dynamic changes in user interests. This study focuses on exploring the application of data science and machine learning in recommender systems, aiming to solve the above challenges, improve recommendation effect and accuracy, and promote technological development and platform competitiveness. By optimizing the algorithms and adapting to environmental changes in real time, it not only promotes the technical innovation of recommender systems, but also brings more accurate personalized service experience to users.



**Figure 1.** Application scenario of intelligent recommender system

## 2. OVERVIEW OF INTELLIGENT RECOMMENDATION SYSTEM

### 2.1. Definition and Characteristics of Intelligent Recommendation System

Intelligent recommender system, as an advanced information technology, its core lies in providing users with personalized content and services through in-depth analysis of multi-dimensional information such as users' historical behaviors and interests and preferences. Compared with traditional recommender systems, intelligent recommender systems pay more attention to the intelligence of the algorithm and the deep mining of data. Traditional recommender systems rely on explicit user feedback, such as ratings, reviews, etc., while intelligent recommender systems can more accurately capture the potential needs and interests of users, even in the absence of clear feedback from users, implicit feedback, such as browsing behavior, click streams, etc., to provide more accurate and personalized recommendation services. For example, e-commerce platforms use intelligent recommendation system, not only based on the user's purchase records, but also combined with their browsing history, search behavior and other multi-dimensional data, for the user to recommend products that meet their personalized needs, thus effectively improving the user experience and purchase conversion rate. Intelligent recommendation system is characterized by its powerful data processing capability, highly personalized service and real-time dynamic adjustment capability, which make the intelligent recommendation system show a wide range of application prospects and important research value in the modern internet environment [2].

### 2.2. Classification of Intelligent Recommender Systems

Intelligent recommender systems can be mainly categorized into content-based recommendation, collaborative filtering recommendation and hybrid recommendation according to their core algorithms and application scenarios (see Table 1) [3]. Content-based recommender systems analyze users' historical behavior and preferences to recommend content similar to their past behavior, such as Netflix, which analyzes users' viewing history to recommend similar types of movies and TV shows. Collaborative filtering recommenders analyze the behavior of user groups to find out user groups with similar interests and make recommendations based on the preferences of these groups,

e.g., Amazon's "Customers Also Bought" feature is a typical application. Hybrid recommender system combines the advantages of the first two, not only considering the user's personal history and behavior, but also integrating the behavior analysis of the user group, in order to improve the accuracy and diversity of the recommendation, such as Spotify's music recommendation algorithm combines a variety of recommender strategies to provide users with personalized music recommendation services. These classifications reflect the flexibility and diversity of intelligent recommendation systems in different application scenarios, and also reveal their potential for continuous development and optimization.

**Table 1.** Classification of intelligent recommender systems

Categorization	Principle	Source of information
Collaborative filtering	Assuming that people are similar and we all share similar behaviors then demand	User records and contextual parameters, group data and preferences are about the same
Content-based	Follow the change for the better by recommending similar items based on what users have liked in the past	User records and contextual parameters, product characteristics
knowledge-based	It is thought that there will be additional sources of information, i.e. explicit personalized knowledge	User records and contextual parameters, product characterization, knowledge modeling
Mixed Recommendations	Constructing a hybrid system that combines the advantages of different algorithms and models while overcoming their shortcomings	User records and contextual parameters, group data, product characteristics, knowledge models

### 2.3. Architecture and Core Modules of Intelligent Recommender System

The typical architecture of an intelligent recommendation system covers multiple aspects from data collection to recommendation generation, and its core modules include data collection, user modeling, item modeling and recommendation algorithms. The data collection module is responsible for obtaining user behavior data and item attribute data from various channels to provide the basis for subsequent modeling and recommendation. The user modeling module builds user profiles through in-depth analysis of the user's historical behavior, interests and preferences to capture the user's personalized needs. The item modeling module, on the other hand, models the attributes, contents, and other information of the items in order to better describe and recommend the items [4]. The recommendation algorithm module is the core of the intelligent recommendation system, which generates a personalized recommendation list for the user based on the results of user modeling and item modeling, using various algorithms such as collaborative filtering, content-based recommendation, and deep learning. Taking the e-commerce platform as an example, its intelligent recommendation system recommends goods that meet the personalized needs of users by collecting their behavioral data such as browsing, searching, and purchasing, combining their historical preferences and real-time behaviors, and using deep learning algorithms to recommend goods that meet their personalized needs, so as to achieve an accurate and personalized recommendation service. The design and implementation of this architecture and its core modules fully demonstrate the complexity and efficiency of the intelligent recommendation system in the modern Internet environment.

## **3. DATA SCIENCE AND INTELLIGENT RECOMMENDER SYSTEMS**

### **3.1. The Role of Data In Intelligent Recommender System**

Data plays a central role in intelligent recommender systems, and its quality and diversity have a decisive impact on the recommendation performance. Rich data sources enable the system to accurately capture user interests and behavioral patterns, and help e-commerce platforms build personalized recommendation services. However, data sparsity and cold-start problem limit the system performance, and academia and industry have responded to this problem by means of matrix decomposition, data filling, etc. Meanwhile, social networks and user registration information are used to assist in solving the cold-start problem and improve the accuracy of recommendation for new users [5]. The importance of data timeliness has prompted the system to update user profiles and models in real time to match dynamically changing user interests. In summary, optimizing data processing and ensuring data quality and timeliness not only strengthens the effectiveness of the intelligent recommendation system and user satisfaction, but also lays the cornerstone for the continuous innovation of personalized recommendation services, which foretells the broad prospect of data-driven recommendation systems.

### **3.2. Data Preprocessing and Feature Engineering**

Data preprocessing and feature engineering are the core steps in building an efficient intelligent recommendation model. In the preprocessing stage, data cleaning removes noise and outliers to ensure data quality, especially when dealing with user behavior data, and effectively filters invalid interaction records. Feature engineering, on the other hand, focuses on feature extraction and selection, refining the information useful for the recommendation task and screening the most influential features to improve model performance, reduce complexity and enhance the recommendation effect. In the practice of e-commerce platforms, well-designed feature engineering significantly improves recommendation accuracy and user satisfaction, and at the same time enhances the model's interpretability and generalization ability, captures the complex relationship between users and products, and promotes personalized and diversified recommendations. Therefore, in-depth investigation of preprocessing and feature engineering strategies is of great significance to promote the continuous progress of intelligent recommendation technology.

### **3.3. Data Science and User Profile Construction**

Data science plays a central role in constructing accurate user profiles to support personalized recommendation. As the cornerstone of personalized recommendation, the accuracy of user profiles directly affects the recommendation effect. Through the use of machine learning algorithms and big data analysis techniques and other data science methods, key features such as user preferences, behaviors and needs can be extracted from massive data, and a detailed user profile can be constructed (see Figure 2). E-commerce platforms use multi-dimensional data such as users' historical purchase records, browsing behaviors and search keywords to identify groups of users with similar characteristics through techniques such as clustering analysis and association rule mining, and build unique user profiles for each group. These profiles not only contain basic user attributes, such as age, gender and geographic location, but also cover deeper consumption habits and lifestyle information, enabling the recommendation system to more accurately match user needs and commodities to achieve personalized recommendations. At the same time, data science also helps the dynamic update and optimization of user profiles, capturing changes in user behavior in real time through real-time data stream processing and incremental learning algorithms, and adjusting user profiles in a timely manner to ensure the timeliness and accuracy of recommendations. With the progress of data science and technology, the construction of user profiles will be further refined, bringing greater innovation opportunities for personalized recommendation services.

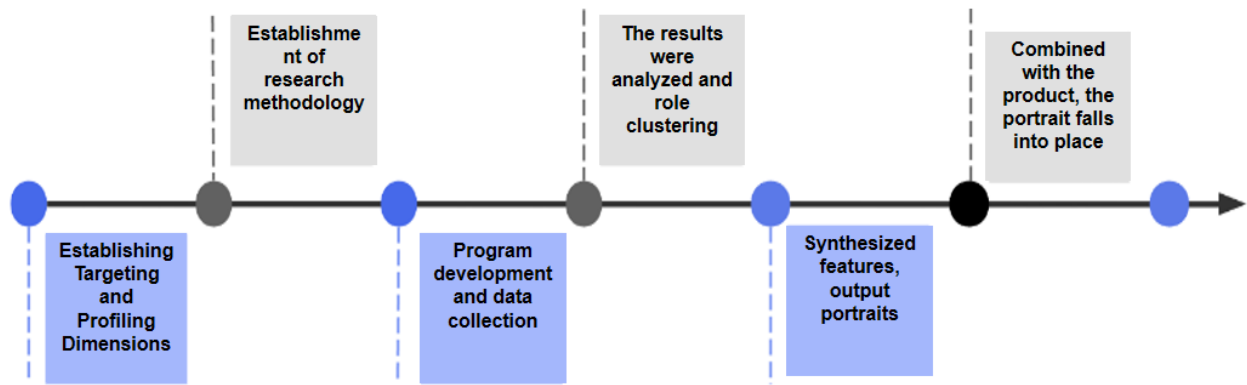


Figure 2. User profile construction process

## 4. APPLICATION OF MACHINE LEARNING ALGORITHMS IN INTELLIGENT RECOMMENDATION SYSTEMS

### 4.1. Application of Classical Machine Learning Algorithms in Recommendation and Its Experimental Evaluation

Classical machine learning algorithms occupy an important position in recommendation systems, in which XGBoost classifier, as the core algorithm, shows its powerful classification ability. The algorithm integrates several weak learners (mainly decision trees) through the gradient boosting strategy, and continuously optimizes the model prediction performance in the iterative process. To specifically evaluate the performance of XGBoost in recommender systems, a series of experiments were conducted in this study.

The experiments were conducted in a standard computing environment using a dataset from the Taobao online shopping mall, which contains 10,000 records where the number of clicks on messages per user ranges from 4 to 12. In order to eliminate messages with less than 4 clicks, we used 80% of the data as the training set and 20% as the test set (see Table 2).

The experimental results show that XGBoost achieves excellent recognition rates, up to 87%, and relatively short computation times, the shortest being 0.1267 s. This indicates that XGBoost has efficient prediction capabilities and good performance in recommender systems.

Table 2. Recognition rate and computation time

Test	Recognition rate/%	Calculation time/s	Test	Recognition rate/%	Calculation time/s
1	85	0.2061	6	86	0.1731
2	87	0.1783	7	87	0.1648
3	85	0.1654	8	86	0.2016
4	87	0.1267	9	85	0.1794
5	86	0.1789	10	8	0.1479

(Source of original data: Taobao User Behavior Dataset Dataset - DataFountain)

After the prediction process, the system takes the five closest predicted commodities and converts the label-coded items to real commodity numbers, and the conversion results are shown in Table 3. By comparing the recommended item numbers with the user's click list, it can be judged whether the recommended items accurately reflect the user's interests and preferences.

**Table 3.** Commodity recommendation results

User	Projected value	Actual value	Merchandise that is most similar	Real item number
1	765	765	775, 765	469979036, 476118140
2	789	788	789, 788	481211265, 481209638
3	790	790	780, 790	480614516, 478771178
4	757	758	748, 796	469978712, 469979179
5	72	143	64, 91	256900418, 256883617

The experiments fully demonstrate the value of XGBoost's wide range of applications in recommender systems and its ability to provide accurate and personalized recommendation services. Its powerful classification model, efficient prediction ability and good performance make XGBoost one of the indispensable classical machine learning algorithms in recommendation systems.

#### 4.2. Optimization of Deep Learning in Intelligent Recommender Systems

Deep learning models, such as RNN, LSTM, CNN and Transformer, significantly improve recommendation accuracy in intelligent recommender systems [6]. These models achieve dynamic and fine modeling of user preferences by capturing the temporal dependence of user behavior sequences and complex feature interactions, which enhances the personalization of recommendations. Meanwhile, integrated learning technology effectively overcomes the limitations of a single technology and improves the overall system performance by integrating multiple recommendation strategies, such as content-based, collaborative filtering and deep learning recommendation. The design of the hybrid recommendation system is based on this multifaceted fusion idea, aiming to achieve more comprehensive user preference modeling and more accurate product recommendation, opening up a new path for the development of personalized recommendation services.

#### 4.3. Integrated Learning and Hybrid Recommender System Pinning

As an effective means to improve the performance of machine learning models, the integrated learning technique shows significant potential for application in recommender systems. By combining the prediction results of multiple recommendation models, this technique can significantly reduce the bias and variance of a single model, thus improving the accuracy and robustness of the recommendation system [6]. The design idea of hybrid recommender system is based on this concept of integrated learning, aiming to overcome the limitations of a single recommendation technology and improve the overall performance of the recommendation system by combining multiple recommendation technologies. The hybrid recommender system not only integrates traditional content-based recommendation and collaborative filtering recommendation technologies, but also incorporates deep learning-based recommendation technologies to fully utilize the advantages of deep learning models in feature extraction and complex pattern recognition. The implementation methods include weighted fusion, integrated learning, hybrid ranking, collaborative training, and hybridization of policy level and feature level. These multivariate fusion strategies optimize the recommendation algorithms, further enhance the degree of personalization of the recommendations, and provide users with more accurate and diverse product recommendation services.

## 5. CHALLENGES AND SOLUTIONS

### 5.1. Challenges of Intelligent Recommender System

Intelligent recommender systems face multiple challenges in practical applications, among which data sparsity, cold-start problem, and scalability are particularly prominent. Data sparsity refers to the

existence of a large number of missing values in the user-item interaction matrix, which makes it difficult for recommendation algorithms to accurately capture user preferences and item characteristics. In the case of e-commerce platforms, for example, the relatively limited variety of items purchased by users makes the presence of a large number of unpurchased records in the user-item matrix increase the difficulty of recommendation. The cold-start problem, on the other hand, involves the difficulty of recommendation algorithms in generating accurate recommendation results due to the lack of sufficient interaction data when new users or new items join the system. In social media platforms, the information overload problem faced by new users when registering is a concrete manifestation of the cold-start problem. Scalability is also a major challenge for intelligent recommender systems. As the number of users and items grows rapidly, recommendation algorithms need to ensure the quality of recommendations while realizing efficient computation and storage. Together, these problems constitute a major obstacle to the practical application of intelligent recommender systems, and an effective solution is urgently needed.

## 5.2. Existing Solutions and In-Depth Analysis

Aiming at the challenges faced by intelligent recommender systems, academia and the industry have adopted a variety of strategies to deal with them. The cold-start problem is solved by utilizing the user's social network information or content-based recommendation methods (as practiced by short-video platforms such as Jittery Voice, which effectively alleviates the content recommendation problem for new users entering the platform for the first time for user interest initialization). The data sparsity problem, on the other hand, successfully extracts features from sparse user-item interaction data by virtue of matrix decomposition techniques, especially Singular Value Decomposition (SVD) and Non-negative Matrix Factorization (NMF), so as to improve recommendation accuracy [8]. Migration learning, as an emerging paradigm, shows strong potential in handling cold starts and data sparsity through cross-domain knowledge transfer. These solutions not only provide a way to alleviate the dilemmas in the practical application of recommender systems, but also open up new paths for subsequent research and development, reflecting the close integration of academic research and industry practice.

## 6. CONCLUSION

In this paper, the application of data science and machine learning algorithms in intelligent recommendation system and its effectiveness are discussed in depth. Through key technologies such as data preprocessing, feature engineering and user profile construction, the system effectively overcomes the challenges of data sparsity and cold start, and significantly improves the recommendation accuracy and personalization level. The deep integration of data science and algorithms not only enhances data processing and pattern recognition capabilities, but also promotes the optimization and innovation of recommendation algorithms. In the future, the integration of emerging algorithms such as deep learning and graph neural networks will further improve system performance and robustness, and the integration of cross-field technologies will expand application prospects. Meanwhile, with the rapid changes in user behavior and environment, future research should focus on the real-time adaptability, interpretability and privacy protection of algorithms to ensure the accurate and personalized service and healthy development of recommendation systems.

## REFERENCES

- [1] Hu Di. Research on intelligent product recommendation system based on machine learning [J]. *Wireless Connected Technology*, 2023, 20(16):18-21.
- [2] Ian Goodfellow, Yoshua Bengio, Aaron Courville. *deep learning* [M]. Beijing: People's Posts and Telecommunications Press, 2021.

- [3] Zhou Zhihua. Machine Learning [M]. Beijing: Tsinghua University Press, 2016.
- [4] Xiang Liang. Intelligent recommender system [M]. Beijing: People's Posts and Telecommunications Press, 2012.
- [5] Riabchuk V, Hagel L, Germaine F, et al. Utility-based context-aware multi-agent recommendation system for energy efficiency in residential buildings [J]. Information Fusion, 2024, 112102559-102559.
- [6] Hsieh S F. Applying “Two Heads Are Better Than One” Human Intelligence to Develop Self-Adaptive Algorithms for Ridesharing Recommendation Systems [J]. Electronics, 2024, 13(12):2241-2241.
- [7] Xiaoyan M. Cross-domain information fusion and personalized recommendation in artificial intelligence recommendation system based on mathematical matrix decomposition [J]. Scientific Reports, 2024, 14(1):7816-7816.
- [8] Rani S S, Shilpa P ,Menon G A .Enhancing Drug Recommendations: A Modified LSTM Approach in Intelligent Deep Learning Systems [J]. Procedia Computer Science, 2024, 233872-881.