A Comprehensive Survey on the Evolution and Deep Learning Techniques in Recommender Systems

Smitha S P ***Harishkumar K S***

smithasp.89@gmail.com harishkumar@presidencyuniversity.in

Abstract:

This survey paper provides a comprehensive review of the evolution, methodologies, and current state-of-the-art in recommender systems (RS), which are pivotal in managing the surge of data across various digital platforms, particularly in e-commerce. It traces the development of RS from early collaborative and content-based filtering techniques to sophisticated machine learning models that leverage deep learning, hybrid methods, and complex algorithms to enhance prediction accuracy and user interaction. Despite significant strides in the field, recommender systems continue to grapple with challenges including the cold-start problem, data sparsity, synonymy, and scalability, as well as ensuring user privacy in an increasingly security-conscious digital environment. The paper evaluates different RS architectures and their effectiveness across diverse sectors, highlighting how advancements like knowledge-based systems, attention mechanisms, and temporal models address specific shortcomings. We critically analyse existing gaps in technology and methodology, proposing future directions for research that include enhancing personalization features, improving adaptive learning capabilities, and integrating ethical AI practices.

Keywords.: Recommender system, Collaborative Filtering, Knowledge Filtering, Predictive Analytics, Neural Collaborative Filtering, Factorization Model.

1. INTRODUCTION

In this age of digitalization, the practice of making frequent purchases online has permeated people's daily lives. It might be challenging for customers to choose the most enticing goods from the wide selection that online stores provide. The digital era we live in today is characterized by a wealth of easily available material on the internet due to the broad availability of online services and recent technical breakthroughs. Customer reviews are appreciated for the variety of products and services that are offered online. Global advances in computer technology have led to the problem of internet data overload. Because there is so much stuff available online, it can be difficult to find relevant and useful information. Consumers may now be effectively steered to the right material utilizing more economical approaches thanks to technology improvements. Recommender systems have gained traction in the last few years. Recommender systems serve as information filters, presenting users with pertinent and useful content. Every page of the website has a suggestion system installed. On their homepages, e-commerce websites frequently include a "guess yourfavourite" function. Websites frequently use recommendation algorithms, which are used in a variety of industries including news, movies, videos, books, restaurants, and maps [1].The recommendation framework (RS) is intended to simplify the process of creating customized suggestions for clients while handling massive data sets with ease. A wide number of businesses, including publishing, movies, music, literature, essays, photography, financial services, and marriage, use random sampling, or RS. Researchers began creating suggestions in the middle of the 1990s by utilizing various filtering methods, such as collaborative and content-based filtration [2].

Considerable progress in recommendation science has been made in the last few years, with a particular focus on recommendation algorithm optimization.. Recommender systems have demonstrated superior effectiveness and customization compared to conventional keyword-based search methods in managing extensive amounts of data [3]. Numerous results are frequently returned by keyword-based searches, although many of them might not be appropriate or relevant. Researchers must use a filtering procedure in order to exclude relevant information from the search results. A significant degree of similarity will be shown in the search results when two researchers enter the identical question. The keyword-based search strategy has limitations when it comes to taking users' varied goals and interests into account. It is important to recognize the differences between the issues covered by the recommended results and the researchers' own interests. Short and controlled algorithms may be used to generate results, enabling customized and efficient recommender systems. Many recommendation algorithms have been developed as a result of recommender systems [4]. The subject of recommendation approaches encompasses four primary types: collaborative filtering (CF), content-based filtering (CBF), graph-based filtering (GB), and hybrid recommender system. Enhancing accuracy was the primary goal of the Recommendation System (RS),. Four new goals have evolved in the field of Recommender Systems (RS) with the increase in goods and users: diversity, serendipity, innovation, and coverage [5]. The recommended products for the receiver must be chosen at random in order to maintain serendipity. This product's principal aim is to surprise the customer. It is acceptable to use the phrases "novelty" and "serendipity" interchangeably [6-7]. An often-used filtering technique called the suggestion method was first created to prioritize email content based on user significance. It was observed that typical functionalities like messages, responses, and annotations were missing. For a variety of experimental tasks, including receiving, filtering, and searching an unbroken stream of electronic documents, the tapestry facilitates group filtering. Another technique to ascertain an individual's implicit evaluation is to look at how much time they spend on a certain webpage. The fast expansion of features and product choices has made it more difficult to choose the perfect product that fits a person's tastes [8].

Depending on the particulars of each case, accessible recommendation systems may or may not be accurate.To provide suggestions for products, the first generation of e-commerce systems use a variety of filtering techniques, including content-based, collaborative, and hybrid filtering. Third-generation systems use knowledge-based models that take human characteristics, emotional aspects, and demography into account.. For the purpose of efficiently handling the substantial rise in data volume [9], it is essential to incorporate user feedback that has been assessed by a recommendation engine. In order to overcome the limitations mentioned above, the estimation and prediction of product items are conducted by utilizing the user profile [10]. The aforementioned systems commonly encounter challenges such as effectively managing data sparsity, ensuring privacy and scalability, addressing the cold-start problem with new users and items, and accurately capturing user preferences and behaviors. In addition, the paper identifies significant research gaps and proposes potential avenues for future development [11].

*1.1. **Challenges associated with recommendation systems***

- **Cold start problem:** Users are not originally offered a new item when it is introduced to the recommendation system. It is challenging to forecast customer preferences with any degree of accuracy when ratings and reviews are missing for the item. The recommendation system could thus provide less precise options. When additional people or things are added to the system, a problem occurs.
- **Sparsity**:One problem that frequently occurs is the inability of many customers to rate or review the things they buy. This leads to an insufficiently detailed grading model. Data sparsity can cause a number of problems, such as making it challenging to pinpoint a particular user group with comparable ratings or interests.
- **Scalability:** Scalability of algorithms in recommendation systems using real-world datasets is a crucial factor to take into account. Because of the volume of dynamic data created by user-item interactions, such ratings and reviews, the aforementioned datasets provide a considerable scalability problem.

2. LITERATURE SURVEY

The scenarios that make up the recommender system are categorized based on the techniques used to produce recommendations for the e-commerce process. The recommender system generates or offers pertinent product recommendations based on algorithms, similarity coefficients, and user knowledge range.The taxonomy categorizes it into four distinct groups: content-based, knowledge-based, collaborative, and hybrid filtering. Figure 1 depicts the category utilized by the RS system for processing, as shown below:

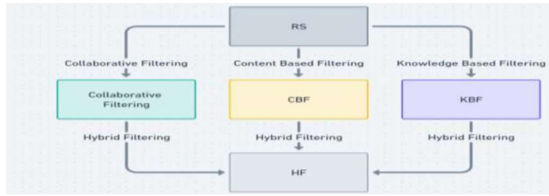

Figure 1: Types of recommendation system

- **Collaborative Filtering (CF)**

Collaborative filtering (CF) is a commonly employed technique in the field of recommender systems (RS). The process involves using historical data of users' indicated preferences for specific items to forecast future user preferences for similar goods [12]. When faced with challenges in finding new participants, user groups may adopt varying strategies for implementing the CF idea that are more effective and efficient. The complexity of the method arises from the lack of a precise definition of "collection" in terms of quality standards and personal preferences. The k-nearest neighbors (k-NN) model is employed by the Collaborative Filtering (CF) approach to assess the similarity of users' performances. The system records the elements selected by the user once the evaluation is completed. The rankings of the neighbors are utilized to determine the sequence of these items. The assessment procedure does not involve client participation. Recommender systems have become increasingly important in managing product data on websites, relying on user evaluations [13]. The recommendation system demonstrates exceptional performance across various contexts, such as e-learning and e-commerce. The RS system enables users to obtain comprehensive information about various items. The product classification is performed using a collaborative filtering-based recommendation system that utilizes the K-Nearest Neighbors algorithm and user input.

The k-NN approach will be employed to construct an information database that incorporates user similarities. A collaborative filtering technique was developed in [14] to determine user ratings for items by considering similarity. The MovieLens dataset is commonly utilized to evaluate and implement the generated model. The findings of the experimental study indicate that the collaborative filtering approach developed performs significantly better than the suggested benchmark. The effectiveness of a recommendation system is determined by the similarity metric used to measure user engagement. The Jaccard Uniform Operator Distance and the Pearson Correlation Coefficient are two factors considered when evaluating similarity. The objective of the sparse matrix is to optimize the efficient storage and retrieval of user-item ratings. The primary objective during the development of the recommendation system was to provide users with information about others who share similar interests. There was a strong focus on promoting active user engagement. A recommendation system for product rating based on a similarity model was developed in [15] to effectively identify potential similarities among consumers of a product. The model created utilizes global preference behaviour instead of relying solely on the local context information of the user rating. The experimental investigation utilizes three distinct real-time data sets, with a focus on analysing their similarity metrics.

- **Content-Based Filtering (CBF)**

Based on user experience, the Content-Based Filtering (CBF) approach is advised for presenting goods that are connected to comparable content. Examining the metadata is the first step in the analysis. Comparing the similarity of the current user data is the main goal of this inquiry. By using both historical and current data, the Collaborative Filtering (CBF) approach maximizes the suggestion and purchase process in a shorter length of time.The degree to which the content-based strategy is vulnerable depends on a number of factors, such as how well the material is described, how complicated the information has to be, and how hard it is to demonstrate user similarity. The development of a hybrid recommender system addressed the shortcomings of the aforementioned method. Collaborative and content-based filtering techniques are included into the system [15]. The process of evaluating product evaluations for similarity begins with the deployment of collaborative and content-based filtering algorithms. User similarity algorithms that are considered relevant are grouped and aggregated by the Collaborative Filtering (CBF) system.

- **Knowledge-based Filtering**

Explicit information is necessary for a knowledge-based algorithm with a recommendation system model to operate well. The knowledge-based system assesses the product's attributes and determines if it will be beneficial to the target audience by using a set of rules and algorithms. The recommendation system combines the calculated similarity score with the constraint-based knowledge option. This form's user preference is determined by a system that blends constraint-based and knowledge-based techniques. The repair system's goals are to fix data problems and provide recommendations to the user. This is accomplished by doing calculations based on predetermined parameters and constraints [16].The knowledge-based (KB) component contains the decisions about user requests that are derived from the decision-making model. Expert domain knowledge is often utilized by the Knowledge Base (KB) to compute and maintain recommendation-specific rules. The knowledge-based approach is intended to use observable conditions and system interaction. The prediction system makes use of the values that are produced by these observable circumstances. The main benefit of the Knowledge Base (KB) system is that it can be tailored to fit different sets of rules in the recommendation process system, including business rules. This cold-start issue is avoided by the knowledge-based recommender scheme, which uses user ratings to provide recommendations [17]. The knowledge-based approach's collection of clear fashion guidance information is one of its limitations. The case-based subscheme and the constraint-based sub-scheme make up the knowledge-based scheme.

- **Hybrid Filtering**

The main goal of hybrid recommendation systems is to integrate various strategies used in recommendation systems. The hybrid recommendation system is designed to leverage the strengths of both X and Y strategies to overcome the limitations. The challenge arises when novel product formulations do not have a rating, making it impossible to recommend them using the collaborative filtering approach. The implementation of the method for detecting and predicting new product descriptions is not restricted solely to the content-based system. A hybrid model-based system is developed through the integration of different techniques and approaches. The recommender system employs a hybrid recommender scheme to gather implicit input on manufacturing procedures. Content-based filtering (CBF) is a predictive method that focuses on the product's features to determine its rating. The identification of specific qualities of the new product that align with the user profile is challenging based on the product profile and user feedback. Consequently, providing a recommendation becomes difficult. The cold-start problem is a significant challenge encountered by the hybrid recommendation model. It pertains to the difficulty of providing necessary product information when there is no product rating available. The sequential suggestion approach has been commonly employed in research investigations. The FPMC system [14] utilizes factorization of the transition matrix of the underlying Markov Chain (MC) to generate personalized sequential behaviors for predicting items. This is achieved by leveraging the capabilities of Markov Chains (MC) and Matrix Factorization (MF). The Recurrent Neural Network (RNN) is currently utilized to effectively capture the long-term dependence of user-item interactions in a series. The achievement is facilitated through the implementation of various strategies. The session-based recommendation method utilizes the user-based RNN, GRU4Rec, and HRNN models to analyze user behavior and detect dependencies. A Gated Recurrent Unit (GRU) is utilized in the operation of these models. The GRU4RecC method is an enhancement of the GRU4Rec model. The performance of the original model is enhanced by employing a unique sampling technique and loss function in this strategy. The RRN approach utilizes the long short-term memory (LSTM) methodology [15] to calculate the dynamic embeddings of individuals and objects. Table 1 displays the survey table. Deep learning is used in the collaborative filtering process of the NeuralCF technology. It has grown in popularity and is being used in many different contexts right now [16]. The NeuralCF method replaces the inner product operation in matrix decomposition with a neural network, improving the model's ability to handle nonlinearity and integrate information. Parameter satisfaction, user modeling, machine learning data, and other AI approaches are used.

Table 1: Model and its Gaps

Model	Advantages	Disadvantages	Research Gaps
FPMC	Integrates MC with MF for tailored predictions.	May not capture complex sequential patterns as effectively as DNNs.	Need for deeper learning techniques to improve accuracy.
GRU4Rec, HRNN	Effective in modeling user behavior and dependencies for session-based recommendations.	Limited by session length and data sparsity issues.	Enhancements in handling longer sequences.
GRU4RecC	Enhances GRU4Rec with a unique loss function and sampling scheme.	May overfit on specific session patterns.	Methods to reduce overfitting and increase generalizability.
RRN	Employs LSTM to capture dynamic embeddings effectively.	High computational cost, complex to train.	Optimization of computational efficiency.
LSTM	Well-suited for long-term dependency modeling.	Struggles with very long sequences and forgetfulness of initial inputs.	Improvements in memory retention over longer sequences.
NeuralCF	Replaces inner product with NN, enhancing nonlinear feature combination.	May require extensive data to train effectively.	Techniques to reduce training data requirements.
AFM	Attention mechanism highlights important traits dynamically.	Complexity increases with number of traits.	Scalability and efficiency improvements.

One machine learning approach that regularly generates suggestions is the AdaBoost algorithm. Tags are used to contain user material. AFM, or the Attention Factorization Model, was created by [17] in order to address the necessity to account for the varied significance of various crossing attributes. An attention mechanism is incorporated into the AFM factorization model. A two-headed self-encoder model was introduced with the goal of learning representation from user comments and implicit feedback. The authors' suggested method is an easily scalable architecture neural network-based grading tool. The aim of this methodology is to faithfully capture the intrinsic properties of entities and people in complex nonlinear contexts. A unique deep attention-based sequential model was put out in an earlier work [18]. In order to properly discern an item's relevance with respect to the user's short- and long-term preferences, the model makes use of attention processes. Recommender systems make extensive use of attention methods. The authors described a collaborative sequential recommendation system with category awareness in their work [19]. The system employs a self-attentive methodology to scrutinize and detect recurrent patterns across several categories. After then, the system determines which particular transition patterns need to be taken into account by considering the categories of recent actions. The suggested method is a processes-based, self-attentive learning paradigm for knowledge representation. This paradigm makes use of knowledge graphs to enhance the communication of entity learning characteristics. According to reference, the self-attention mechanism helps the user recognize important information by reducing the impact of outside information on the attention mechanism. The model's goal is to have a thorough grasp of the direct links between each item and the underlying structure of the behavioral sequence.

3. Conclusion:

In conclusion, this survey has extensively explored the dynamic landscape of recommender systems, chronicling their evolution from simple algorithmic frameworks to sophisticated systems that integrate advanced machine learning techniques. Despite the remarkable progress in enhancing personalization and accuracy, several persistent challenges remain. Key among these are the cold-start problem, data sparsity, and maintaining user privacy and system scalability in diverse applications. The continuous growth in digital data further complicates these issues, necessitating innovative solutions.Future research should focus on developing robust models that not only address these challenges but also incorporate user feedback in real-time, adapt to changing user behaviours, and respect ethical standards in data usage. Moreover, integrating cross-disciplinary approaches from fields such as psychology and sociology could enrich the understanding of user interactions and improve the recommendation quality.

References:

1. G. Zhang, Y. Liu, and X. Jin, ‘‘A survey of autoencoder-based recommender systems,’’ Frontiers Comput. Sci., vol. 14, no. 2, pp. 430–450, Apr. 2020.
2. S. Zhang, L. Yao, A. Sun, and Y. Tay, ‘‘Deep learning based recommender system: A survey and new perspectives,’’ ACM Comput. Surv., vol. 52, no. 1, pp. 1–38, Jan. 2020.
3. R. Mu, ‘‘A survey of recommender systems based on deep learning,’’ IEEE Access, vol. 6, pp. 69009–69022, 2018.
4. D. Alahmadi and F. Alruwaili, ‘‘Deep learning for MOOCs course recommendation systems: State of the art survey,’’ Int. Trans. J. Eng. Manage. Appl. Sci. Technol., vol. 12, no. 11, 2021, Art. no. 12A11Q.
5. L. Salau, M. Hamada, R. Prasad, M. Hassan, A. Mahendran, and Y. Watanobe, ‘‘State-of-the-art survey on deep learning-based recommender systems for e-learning,’’ Appl. Sci., vol. 12, no. 23, p. 11996, Nov. 2022.
6. vanCapelleveen, G., Amrit, C., Yazan, D. M., &Zijm, H. (2019). The recommender canvas: a model for developing and documenting recommender system design. Expert systems with applications, 129, 97-117.
7. Haruna, K., Akmar Ismail, M., Damiasih, D., Sutopo, J., &Herawan, T. (2017). A collaborative approach for research paper recommender system. PloS one, 12(10), e0184516.
8. Mısır, M., &Sebag, M. (2017). Alors: An algorithm recommender system. Artificial Intelligence, 244, 291-314.
9. Tejeda-Lorente, Á.,Porcel, C., Peis, E., Sanz, R., & Herrera-Viedma, E. (2014). A quality based recommender system to disseminate information in a university digital library. Information Sciences, 261, 52-69.
10. Winkels, R., Boer, A., Vredebregt, B., & Van Someren, A. (2014, November). Towards a Legal Recommender System. In JURIX (pp. 169-178).
11. Ma, T., Zhou, J., Tang, M., Tian, Y., Al-Dhelaan, A., Al-Rodhaan, M., & Lee, S. (2015). Social network and tag sources based augmenting collaborative recommender system. IEICE transactions on Information and Systems, 98(4), 902-910.
12. Kiran, R., Kumar, P., & Bhasker, B. (2020). DNNRec: A novel deep learning based hybrid recommender system. Expert Systems with Applications, 144, 113054.
13. Sridevi, M., Rao, R. R., & Rao, M. V. (2016). A survey on recommender system. International Journal of Computer Science and Information Security, 14(5), 265.
14. Wang, D., Liang, Y., Xu, D., Feng, X., & Guan, R. (2018). A content-based recommender system for computer science publications. Knowledge-Based Systems, 157, 1-9.
15. Gulzar, Z., Leema, A. A., & Deepak, G. (2018). Pcrs: Personalized course recommender system based on hybrid approach. Procedia Computer Science, 125, 518-524.
16. M. A. Islam, M. M. Mohammad, S. S. S. Das, and M. E. Ali, ‘‘A survey on deep learning based Point-of-Interest (POI) recommendations,’’ Neurocomputing, vol. 472, pp. 306–325, Feb. 2022.
17. Y. Dai and J. Xu, ‘‘Study of online learning resource recommendation based on improved BP neural network,’’ Int. J. Embedded Syst., vol. 14, no. 2, p. 101, 2021.
18. T. K. Dang, Q. P. Nguyen, and V. S. Nguyen, ‘‘A study of deep learningbased approaches for session-based recommendation systems,’’ Social Netw. Comput. Sci., vol. 1, no. 4, p. 216, Jul. 2020.
19. K. Barbara and C. Stuart, ‘‘Guidelines for performing systematic literature reviews in software engineering,’’ Version 2.3, Dept. Comput. Sci., Keele Univ., Keele, U.K., Tech. Rep. EBSE-2007-01, 2007, pp. 1–26.

Smitha S P received her Bachelor's degree in Computer Science and Engineering from Visvesvaraya Technological University, Belgaum in 2011, Master's Degree in Computer Science and Engineering from Visvesvaraya Technological University, Belgaum in 2014 Currently, She is also pursuing her Ph.D. at Presidency University. Smitha S P is an Assistant Professor with 10.2 years of teaching experience, currently serving in the School of Computer Science and Engineering (SoCSE) at Presidency University, Bengaluru, since February 2023.

Dr.Harishkumar K.S. is an Assistant Professor Senior Scale in at the School of Computer Science Engineering, Presidency University with 8 years of postgraduate and undergraduate teaching experience. He has a Ph.D. in Data Science and research interests in AI, machine learning and deep learning. He has published more than 20 research papers in peer-reviewed journals and conferences including, Elsevier, IEEE, Springer and Scopus publications. His papers apply AI for problems like air quality prediction, plant disease detection etc. He has over 289 google scholar citations with 4 H-index, 193 Scopus index citations with 3 H-index and present guides for 5 research scholars and 1 scholar already awarded. He has authored a book chapter and edited book on AI trends. He has taught computer science subjects at universities like Mangalore University and Vijayanagara sri krishnadevaraya University. He has completed over 20 certifications in machine learning, Python, data science etc. He serves as a reviewer for 10 journals, is an editorial member for 2 journals and a technical committee member for multiple conferences. He has been an invited speaker to deliver talks on AI and big data analytics. He has strong academic credentials with rich research, teaching and reviewing experience in the fields of AI, data science and computer science.