Contemporary Recommendation Systems on Big Data and Their Applications: A Survey

ZIYUAN XIA1, ANCHEN SUN2, (Member, IEEE), JINGYI XU3,4, YUANZHE PENG5, RUI MA6, AND MINGHUI CHENG7,8

1Antai College of Economics and Management, Shanghai Jiao Tong University, Shanghai 200030, China
2Department of Electrical and Computer Engineering, University of Miami, Coral Gables, FL 33146, USA
3Department of Architecture, Cornell University, Ithaca, NY 14853, USA
4HOK’s Miami Studio, Coral Gables, FL 33134, USA
5Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32611, USA
6Bascom Palmer Eye Institute, Miller School of Medicine, University of Miami, Miami, FL 33136, USA
7Department of Civil and Architectural Engineering, University of Miami, Coral Gables, FL 33146, USA
8School of Architecture, University of Miami, Coral Gables, FL 33146, USA

Corresponding authors: Anchen Sun (asun@miami.edu) and Minghui Cheng (minghui.cheng@miami.edu)

ABSTRACT This survey paper provides a comprehensive analysis of the evolution and current landscape of recommendation systems, extensively used across various web applications. It categorizes recommendation techniques into four main types: content-based, collaborative filtering, knowledge-based, and hybrid approaches, tailored for specific user contexts. The review spans historical developments to cutting-edge innovations, with a focus on big data analytics applications, state-of-the-art recommendation models, and evaluation using prominent datasets like MovieLens, Amazon Reviews, Netflix Prize, Last.fm, and Yelp. The paper addresses significant challenges such as data sparsity, scalability, and the need for diverse recommendations, highlighting these as key directions for future research. It also explores practical applications and the integration challenges of recommendation systems in everyday life, underscoring the potential of big data-driven advancements to significantly enhance real-world experiences.

INDEX TERMS Recommendation system, big data, machine learning, sustainability.

I. INTRODUCTION
In this survey, we scrutinize the escalating popularity and diverse applications of recommendation systems in web applications, a topic extensively explored by Zhou et al. [1]. These systems, a specialized category of information filtering systems, are designed to predict user preferences for various items. They play a pivotal role in guiding decision-making processes, such as purchasing decisions and music selections, as discussed by Wang et al. [2]. An exemplary instance of this application is Amazon’s personalized recommendation engine, which tailors each user’s homepage. Major corporations like Amazon, YouTube, and Netflix utilize these systems to enhance user experience and generate substantial revenue, as highlighted by Adomavicius and Tuzhilin [3].

Omura et al. [4]. Figure 1 from Entezari et al. [5] provides a visual representation of a modern recommendation system. Recommendation systems are integral to companies, significantly impacting revenue generation and competitive positioning, evidenced by research from Rismanto et al. [6], Cui et al. [7]. For example, Netflix’s “Netflix Prize” challenge aimed to enhance their existing recommendation algorithms, offering a substantial reward to spur innovative solutions. Additionally, modern recommendation systems are increasingly pertinent in the field of human-computer interaction (HCI), where they improve interaction efficiency through feedback mechanisms, a topic explored in several studies [8].

In the context of big data, recommendation systems are increasingly crucial, as detailed by Li et al. [9], [10]. These systems predict user interests in purchasing based on extensive data analysis, including purchase history, ratings, and reviews. There are four widely
Deep foundation systems is propelled by several emerging trends: society. In our journey towards a more informed and connected way we interact with technology, making it a pivotal element [17]. This data-driven approach is revolutionizing how we collect, process, and engage with vast amounts of data across all age groups, from children to seniors, empowering individuals with insights to better understand and analyze their behaviors and preferences [16], [17]. This data-driven approach is revolutionizing how we interact with technology, making it a pivotal element in our journey towards a more informed and connected society.

The integration of big data technologies with recommendation systems is propelled by several emerging trends: Deep Learning Personalization: As datasets grow and neural networks become more sophisticated, recommendation systems will increasingly harness deep learning to analyze complex, unstructured data from diverse sources, predicting user preferences with unprecedented accuracy [18], [19].

Multimodal Data Usage: Future recommendation systems will utilize various data types, including images, videos, and audio. This approach will particularly enhance the quality of recommendations in sectors like fashion and entertainment, where visual and contextual sensitivity is crucial [13].

Real-time Processing with Edge Computing: With the growing demand for immediate responses, edge computing will be adopted more widely, enabling faster, real-time data processing close to data sources and significantly reducing latency [20].

Privacy and Ethics: As data privacy concerns escalate, technologies like federated learning and differential privacy will become increasingly prevalent. These methods will enable recommendation systems to utilize large datasets while protecting user privacy [21].

Sustainability Focus: There will be a shift towards integrating sustainability metrics within recommendation systems, prioritizing products and services based on their environmental impact to promote eco-friendly choices [22].

In summary, the fusion of big data with recommendation systems will not only elevate personalization and operational efficiency but also adhere to privacy standards and enhance environmental sustainability, significantly advancing the impact of these technologies in daily life.

This survey paper contributes uniquely to the recommendation system research filed by:

- Evaluating contemporary recommendation systems and state-of-the-art (SOTA) models, assessing their performance across various well-known datasets;
- Exploring the distinctive technologies and addressing the challenges associated with implementing recommendation systems on big data and the future integration of it;
- Analyzing a range of real-world applications, detailing the technologies employed and the challenges encountered in these settings.

The structure of this paper is as follows: Section II provides a comprehensive exploration of recommendation systems, tracing their evolution from historical foundations to contemporary state-of-the-art methodologies, accompanied by a thorough examination of recent advancements within the sector. Section III addresses the specific challenges encountered in recommendation systems that leverage big data, such as data sparsity, scalability issues, and the imperative for diverse recommendations. This section also delves into potential solutions for overcoming these obstacles. Section IV is dedicated to the application of recommendation systems in real-life contexts, discussing their practical implications and the challenges associated with their integration. Finally, the paper culminates with a comprehensive summary presented in Section VII.
II. RECOMMENDATION SYSTEMS

Recommendation systems aim to predict users’ preferences for a certain item and provide personalized services [23]. This section will discuss several commonly used recommender methods, such as content-based method, collaborative filtering-based method, knowledge-based method, and hybrid-based method.

A. CONTENT-BASED RECOMMENDATION SYSTEMS

The main idea of content-based recommenders is to recommend items based on the similarity between different users or items [24]. This algorithm determines and differentiates the main common attributes of a particular user’s favorite items by analyzing the descriptions of those items. Then, these preferences are stored in this user’s profile. The algorithm then recommends items with a higher degree of similarity with the user’s profile. Besides, content-based recommendation systems can capture the specific interests of the user and can recommend rare items that are of little interest to other users. However, since the feature representations of items are designed manually to a certain extent, this method requires a lot of domain knowledge. In addition, content-based recommendation systems can only recommend based on users’ existing interests, so the ability to expand users’ existing interests is limited.

B. COLLABORATIVE FILTERING-BASED RECOMMENDATION SYSTEMS

Collaborative Filtering-based (CF) methods are primarily used in big data processing platforms due to their parallelization characteristics [25]. The basic principle of the recommendation system based on collaborative filtering is shown in Fig. 2 [26]. CF recommendation systems use the behavior of a group of users to recommend to other users [27]. There are mainly two types of collaborative filtering techniques, which are user-based and item-based.

- User-based CF: In the user-based CF recommendation system, users will receive recommendations of products that similar users like. Many similarity metrics can calculate the similarity between users or items, such as Constrained Pearson Correlation coefficient (CPC), cosine similarity, adjusted cosine similarity, etc. For example, cosine similarity is a similarity between two vectors. Let \( x \) and \( y \) denote two vectors, cosine similarity between \( x \) and \( y \) can be represented by

\[
\cos(\theta) = \frac{\sum_{i=1}^{n} x_i y_i}{\sqrt{\sum_{i=1}^{n} x_i^2} \sqrt{\sum_{i=1}^{n} y_i^2}}
\]

(1)

\[
r_{xy} = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^{n} (y_i - \bar{y})^2}}
\]

(2)

- Item-based CF: Item-based CF algorithm predicts user ratings for items based on item similarity. Generally, item-based CF yields better results than user-based CF because user-based CF suffers from sparsity and scalability issues. However, both user-based CF and item-based CF may suffer from cold-start problems [28].

C. KNOWLEDGE-BASED RECOMMENDATION SYSTEMS

The main idea of knowledge-based recommendation systems is to recommend items to users based on basic knowledge of users, items, and relationships between items [55], [56]. Since knowledge-based recommendation systems do not require user ratings or purchase history, there is no cold start problem for this type of recommendation [57]. Knowledge-based recommendation systems are well suited for complex domains where items are not frequently purchased, such as cars and apartments [58]. But the acquisition of required domain knowledge can become a bottleneck for this recommendation technique [38].

D. HYBRID-BASED RECOMMENDATION SYSTEMS

Hybrid-based recommendation systems combine the advantages of multiple recommendation techniques and aim to overcome the potential weaknesses in traditional recommendation systems [59]. There are seven basic hybrid recommendation techniques [45]: weighted, mixed, switching, feature combination, feature augmentation, cascade, and meta-level methods [60], [61]. Among all of these methods, the most commonly used is the combination of the CF recommendation methods with other recommendation methods (such as content-based or knowledge-based) to avoid sparsity, scalability, and cold-start problems [42], [44], [62].

E. CHALLENGES IN MODERN RECOMMENDATION SYSTEMS

1) SPARSITY, SCALABILITY, AND DIVERSITY

- Sparsity. As we know, the usage of recommendation systems is growing rapidly. Many commercial recommendation systems use large datasets, and the user-item matrix used for filtering may be very large and sparse. Therefore, the performance of the recommendation process may be degraded due to the cold start problems caused by data sparsity [63].
Scalability. Traditional algorithms will face scalability issues as the number of users and items increases. Assuming there are millions of customers and millions of items, the algorithm’s complexity will be too large. However, recommendation systems must respond to the user’s needs immediately, regardless of the user’s rating history and purchase situation, which requires high scalability. For example, Twitter is a large web company that uses clusters of machines to scale recommendations for its millions of users [43].

Diversity. Recommendation systems also need to increase diversity to help users discover new items. Unfortunately, some traditional algorithms may accidentally do the opposite because they always recommend popular and highly-rated items that some specific users love. Therefore, new hybrid methods need to be developed to improve the performance of the recommendation systems [64].

### Table 1. Summary of the modern recommendation systems methods.

| Recommendation Systems | Descriptive Key Points | Papers |
|-------------------------|------------------------|--------|
| Content-Based Filtering (User-based) | Recommends items by finding similar users. This is often measured by observing the items they like or rate similarly. | Musto et al. [29], Volkovs et al. [30], Mittal et al. [31], Almaguer et al. [32] |
| Content-Based Filtering (Item-based) | Recommends items that are similar to items the user has liked in the past, based on the user’s previous actions or feedback on items. | Tewari [33], Ajabgu [34], Almaguer et al. [32] |
| Collaborative Filtering | Recommend items to some users based on the other users behavior. | Zhang et al. [35], Bobadilla et al. [36], Bobadilla et al. [37] |
| Knowledge-Based Systems | Recommend items to users based on basic knowledge of users, items, and relationships between items. | Dong et al. [38], Gaarder et al. [39], Alamdar et al. [40], Cen et al. [41] |
| Hybrid Approaches | Combine collaborative and content-based methods | Hrnjica et al. [42], Shokeen et al. [43], Zagra et al. [44], Ibrahim et al. [45] |
| Demographic-based Approaches | Recommend items based on demographic factors of the user such as age, gender, or location. | Alamdar et al. [40], Renjit et al. [46], Quijano et al. [47] |
| Utility-based Systems | Recommend items based on a computation of their utility for the user. | Raichuk et al. [48], Deldjoo et al. [49], Shambour [50] |
| Session-based and Context-aware Recommendations | Recommend items use the current session data and context, which is often used in ecommerce environments. | Moreira et al. [51], Gwadabe et al. [52], Qiu et al. [53], Chen et al. [54] |

2) COLD-START PROBLEM
The cold start problem in recommendation systems arises when new users lack sufficient historical data to generate accurate recommendations. This challenge is particularly acute due to the sparsity of information available for new users [65]. Early researchers addressed this issue by proposing hybrid approaches that combine collaborative filtering with content-based data to mitigate the effects of information scarcity [66]. In 2017, a significant development was the introduction of Dropoutnet, which demonstrated that neural network models could be specifically trained to handle cold start scenarios using dropout techniques to optimize performance. This advancement highlighted the potential of neural networks in adapting to limited data availability for new users [67]. More recently, in 2021, a novel collaborative filtering ranking model called RBPR was introduced. This model merges the rating-oriented approach of Probabilistic Matrix Factorization (PMF) with the pairwise ranking-oriented approach of Bayesian Personalized Ranking (BPR), offering a refined method to address the cold start problem by leveraging the strengths of both techniques [68]. In 2022, further progress was made by integrating features such as visual appearance, audio, and motion information extracted from video content, which has proven effective in enhancing movie recommendations for new users. This approach demonstrates how multimodal data can be leveraged to improve the accuracy of recommendations in the absence of extensive user history, further advancing the field’s ability to tackle the cold start problem effectively [69].

### III. RECOMMENDATION SYSTEM BASED ON BIG DATA
Big data refers to the massive, high growth rate and diversified information [70], [71]. It requires new processing models to have stronger decision-making and process optimization capabilities [72]. Big data has its unique “4V” characteristics, as shown in Fig. 3 [73]: Volume, Variety, Velocity, and Veracity.

#### A. BIG DATA PROCESSING FLOW
Big data comes from many sources, and there are many methods to process it [74]. However, the primary processing...
of big data can be divided into four steps [75]. Besides, Fig. 4 presents the basic flow of big data processing.

- Data Collection.
- Data Processing and Integration. The collection terminal itself already has a data repository, but it cannot accurately analyze the data. The received information needs to be pre-processed [76].
- Data Analysis. In this process, these initial data are always deeply analyzed using cloud computing technology [77].
- Data Interpretation.

B. MODERN RECOMMENDATION SYSTEMS BASED ON THE BIG DATA

The shortcomings of traditional recommendation systems mainly focus on insufficient scalability and parallelism [78]. For small-scale recommendation tasks, a single desktop computer is sufficient for data mining goals, and many techniques are designed for this type of problems [79].

However, the rating data is usually so large for medium-scale recommendation systems that it is impossible to load all the data into memory at once [80]. Common solutions are based on parallel computing or collective mining, sampling and aggregating data from different sources, and using parallel computing programming to perform the mining process [81]. The big data processing framework will rely on cluster computers with high-performance computing platforms [82]. At the same time, data mining tasks will be deployed on a large number of computing nodes (i.e., clusters) by running some parallel programming tools [83], such as MapReduce [71], [84]. For example, Fig. 5 is the MapReduce in the Recommendation Systems.

In recent years, various big data platforms have emerged [85]. For example, Hadoop and Spark [71], both developed by the Apache Software Foundation, are widely used open-source frameworks for big data architectures [71], [86]. Each framework contains an extensive ecosystem of open-source technologies that prepare, process, manage and analyze big data sets [87]. For example, Fig. 6 is the ecosystem of Apache Hadoop [88].

Hadoop allows users to manage big data sets by enabling a network of computers (or “nodes”) to solve vast and intricate data problems. It is a highly scalable, cost-effective solution that stores and processes structured, semi-structured and unstructured data.

Spark is a data processing engine for big data sets. Like Hadoop, Spark splits up large tasks across different nodes. However, it tends to perform faster than Hadoop, and it uses random access memory (RAM) to cache and process data instead of a file system. This enables Spark to handle use cases that Hadoop cannot. The following are some benefits of the Spark framework:

- It is a unified engine that supports SQL queries, streaming data, machine learning (ML), and graph processing.
- It can be 100x faster than Hadoop for smaller workloads via in-memory processing, disk data storage, etc.
- It has APIs designed for ease of use when manipulating semi-structured data and transforming data.

Furthermore, Spark is fully compatible with the Hadoop eco-system and works smoothly with Hadoop Distributed File System (HDFS), Apache Hive, and others. Thus, when the data size is too big for Spark to handle in-memory, Hadoop can help overcome that hurdle via its HDFS functionality. Fig. 7 is a visual example of how Spark and Hadoop can work together. Fig. 8 is the the architecture of the modern recommendation system based on Spark.

IV. RECOMMENDATION SYSTEM APPLICATION

Recommendation systems have become ubiquitous across diverse sectors such as search engines, digital media platforms, and e-commerce sites on the internet [89], [90]. The progression of information technology has led to their significant evolution, embracing increasingly complex models. The emergence of big data has been a catalyst for the refinement of recommendation systems,
empowering them to offer more precise and holistic recommendations. Currently, the incorporation of big data into contemporary recommendation systems plays a crucial role in streamlining operations in e-commerce and e-governance, as well as in fostering sustainable living practices [91], [92]. This integration signifies a leap forward
A. RECOMMENDATION SYSTEM IN E-COMMERCE
Recommendation systems have evolved from specialized tools utilized by a select few e-commerce platforms to vital commercial assets that significantly transform the e-commerce landscape [93]. Major online platforms and applications, such as Amazon and TikTok, now harness the power of big data to refine their recommendation algorithms for users [94], [95]. Ansari et al. previously highlighted the need for advancements in data collection and analytics to expand the operational advantages in the marketing sector, prompted by the advent of new applications for information agents [89]. With the advancement of big data, the application of information agents has shifted towards providing accurate and tailored recommendations to customers on online markets and platforms through innovative models like topic modeling and sentiment analysis [96]. Therefore, in an era dominated by big data, it’s evident that recommender systems have become widely integrated into various aspects of e-commerce operations. This extensive adoption emphasizes the pivotal role these systems play in improving user experience, personalizing customer interactions, and boosting the overall efficiency of e-commerce platforms.

B. RECOMMENDATION SYSTEM IN E-GOVERNANCE
E-governance stands as a fundamental challenge in the realm of smart city initiatives, integrating information technology and big data in the public sector to elevate the delivery of services and information. This approach not only aims to enhance government transparency, accountability, and trustworthiness but also to engage citizens in the governance process [92]. In the digital era, particularly highlighted by the rise of epidemics, the demand for e-governance in society is on the rise. This underscores the necessity for governments to establish adept online information systems to meet the goals of effective e-governance [97]. Businesses, including pharmaceutical companies, are also recognizing the need for digital governance. For instance, they can implement recommendation systems, built on blockchain and machine learning technologies, to streamline drug shipment monitoring [98]. Additionally, these systems have applications in demand-side management, like energy management, where they utilize big data analytics to identify residential users’ preferences for energy-efficient appliances [99]. Thus, the deployment of recommendation systems, bolstered by big data technologies, plays a crucial role in improving the digital governance landscape.
framework, optimizing business processes, and facilitating efficient energy management in the digital age.

C. RECOMMENDATION SYSTEM IN SUSTAINABLE LIFESTYLE

In the context of an academic discourse exploring the impact of recommendation systems on fostering a sustainable lifestyle, the digital transformation catalyzed by the Internet revolution has significantly facilitated the transition towards e-commerce and a lifestyle centered around digital interactions, concurrently with a heightened emphasis on sustainability across environment, societal, and governance sectors [100]. A Recommender System called sustain.AI [101] is also proposed to analyze Sustainability Reports. This shift has intensified the demand for environmentally friendly practices in daily activities, such as online shopping, dietary habits, and transportation, prompting both organizations and individuals to prefer products and services with reduced environmental footprints. Addressing the nexus of technological innovation, consumer behavior, and environmental responsibility, recent studies advocate the integration of green marketing strategies with online retail platforms through the deployment of sophisticated recommendation systems [102]. These systems are instrumental in refining the digital shopping journey, aiming to promote sustainable living by prioritizing options with lower environmental impacts, thus facilitating a shift towards sustainability among both providers and consumers. The work of Lee and Huang [103] and Zhang et al. [104], highlights the critical capacity of recommendation systems to advocate for environmentally sustainable choices, representing a crucial advancement in embedding eco-sustainability within the digital commerce ecosystem.

Furthermore, green building emerges as a significant aspect that profoundly influences our connection to sustainable living [105]. With the advancement of recommendation systems, big data, and the Internet of Things (IoT), several challenges associated with green building can be addressed through the integration of recommendation systems and machine learning technologies. These technologies hold the potential to enhance various facets of green building, including:

- Enhancing energy efficiency within buildings, [106], [107]
- Facilitating the selection and acquisition of green buildings, [108]
- Implementing Building Information Modeling (BIM) and Life Cycle Assessment (LCA) for sustainable construction, [109], [110]
- Innovating in the realm of green building design [111].

Such integrations not only underscore the potential of technological innovations in promoting energy efficiency and supporting sustainable building practices but also mark a significant stride towards leveraging recommendation systems in the pursuit of a greener lifestyle.

D. RECOMMENDATION SYSTEM IN HEALTHCARE

The healthcare industry is experiencing a paradigm shift with the incorporation of cutting-edge technologies, and recommendation systems have emerged as a pivotal component in this transformative landscape. Leveraging machine learning algorithms and data analytics, recommendation systems in healthcare offer a versatile range of applications [112], from improving clinical decision-making to enhancing patient engagement. Lodhi et al. [113] introduces a knowledge graph-driven methodology for creating highly personalized nutritional recommendations, addressing the variations in individual dietary needs by integrating data on health, lifestyle, and diet, and demonstrating its effectiveness through a case study. This subsection explores the diverse facets of recommendation systems in healthcare and their impact on various stakeholders within the ecosystem.

1) CLINICAL DECISION SUPPORT

One of the primary applications of recommendation systems in healthcare is in clinical decision support. These systems analyze electronic health records (EHRs), medical literature, and patient data to assist healthcare professionals in making informed decisions about diagnostics, treatment plans, and interventions. By providing relevant and evidence-based information, recommendation systems contribute to more accurate and personalized patient care, potentially reducing diagnostic errors and improving overall healthcare outcomes [114], [115].

2) PATIENT-CENTERED CARE

In the era of patient-centered care, recommendation systems play a crucial role in tailoring healthcare services to individual patient needs. These systems analyze patient preferences, demographics, and health histories to generate personalized recommendations for treatment options, preventive measures, and lifestyle modifications. By fostering patient engagement and empowerment, recommendation systems contribute to a more collaborative and effective healthcare relationship between providers and patients [116], [117].

3) RESOURCE OPTIMIZATION

Recommendation systems help optimize healthcare resources by streamlining processes and improving operational efficiency. For instance, in hospital management, these systems can suggest optimal bed allocation, appointment scheduling, and resource utilization based on historical data and real-time information. By minimizing bottlenecks and enhancing resource allocation, recommendation systems contribute to cost-effectiveness and improved service delivery [118], [119].

4) TELEHEALTH AND REMOTE MONITORING

With the rise of telehealth and remote monitoring, recommendation systems support healthcare providers in delivering virtual care. These systems analyze patient-generated health
data from wearable devices, monitoring tools, and tele-health platforms to provide timely recommendations for interventions, medication adjustments, or lifestyle modifications. This real-time support contributes to proactive healthcare management, particularly for patients with chronic conditions [120], [121].

5) COLLABORATIVE HEALTHCARE NETWORKS
Recommendation systems facilitate collaboration and knowledge sharing among healthcare professionals through the creation of collaborative networks. By analyzing expertise, research interests, and clinical experiences, these systems connect healthcare professionals for consultations, research collaborations, and second opinions. This fosters a culture of continuous learning and knowledge dissemination within the healthcare community [122], [123].

As recommendation systems continue to evolve, their integration into healthcare processes holds immense potential for improving patient outcomes, enhancing operational efficiency, and advancing the overall quality of healthcare services. However, challenges such as data interoperability, privacy concerns, and algorithmic transparency must be addressed to ensure the responsible and ethical deployment of recommendation systems in the complex healthcare landscape. Ongoing research, interdisciplinary collaboration, and stakeholder engagement are crucial for harnessing the full benefits of recommendation systems in healthcare.

V. DATASETS FOR RECOMMENDATION SYSTEMS
Datasets play a pivotal role in the development, evaluation, and benchmarking of recommendation systems. The choice of dataset depends on factors such as the application domain, the type of recommendation task, and the specific research objectives. In this section, we provide a summary of some commonly used datasets, as shown in Table 2, in recommendation systems research:

A. MovieLens
The MovieLens dataset is one of the most widely used datasets for collaborative filtering-based recommendation systems. It contains movie ratings provided by users, along with movie metadata such as titles, genres, and release years [144]. MovieLens datasets are available in various sizes (e.g. 100K, 1M, 10M, 20M) ranging from small-scale datasets suitable for initial experimentation to large-scale datasets for comprehensive evaluations.

B. AMAZON REVIEWS
Amazon provides extensive datasets comprising user reviews and ratings for a wide range of products available on its platform, including books, electronics, clothing, and more [145]. These datasets offer rich information about user preferences, product attributes, and user-item interactions, making them valuable resources for research in recommendation systems, sentiment analysis, and e-commerce analytics.

C. NETFLIX PRIZE DATASET
The Netflix Prize dataset is a large-scale dataset of movie ratings collected from Netflix users. It was released as part of the Netflix Prize competition, a renowned competition aimed at improving the performance of recommendation algorithms [146]. The dataset contains millions of ratings provided by users over several years, along with additional contextual information such as user demographics and temporal dynamics.

D. LAST.FM DATASET
The Last.fm dataset comprises music listening histories of users on the Last.fm platform, including details about artists, albums, and user preferences [147]. It offers insights into user behavior and preferences in the context of music consumption, making it valuable for research in music recommendation systems, personalized playlists, and music discovery applications.

E. YELP DATASET
The Yelp dataset consists of user reviews and ratings for businesses across various categories, including restaurants, hotels, and local services [148]. It provides rich information about user opinions, business attributes, and geographic locations, enabling research in recommendation systems, sentiment analysis, and location-based services.

F. OTHER DATASETS
In addition to the datasets mentioned above, several other datasets are available for recommendation systems research, covering diverse domains such as news articles, social media interactions [149], academic papers, and online retail transactions. These datasets offer valuable opportunities for exploring different recommendation scenarios, addressing various challenges, and advancing the state-of-the-art in recommendation technology.

In summary, the availability of diverse datasets facilitates the development and evaluation of recommendation systems across different application domains. Researchers can leverage these datasets to build robust, scalable, and personalized recommendation algorithms, ultimately enhancing user experiences and driving innovation in the field of recommendation systems.

VI. OPEN QUESTIONS AND FUTURE RESEARCH DIRECTIONS
Despite substantial progress in the development of recommendation systems using big data, several critical questions remain unresolved, presenting numerous opportunities for further research. This section highlights key areas that offer promising avenues for future exploration.

A. INTERPRETABLE RECOMMENDER SYSTEMS
Future research could focus on developing recommendation models that not only provide accurate predictions but
TABLE 2. Datasets and compared methods for recommendation systems. The metrics for different models like recall and nDCG evaluate the effectiveness of each top model [143].

| Dataset       | Top Models                          | Metrics         | Extra Training Data | Year |
|---------------|-------------------------------------|-----------------|---------------------|------|
| MovieLens 100K| GHRS [124]                          | 0.887 (RSME)    | N                   | 2021 |
|               | GLocal-K [125]                      | 0.889 (RSME)    | N                   | 2021 |
|               | MG-GAT [126]                        | 0.890 (RSME)    | Y                   | 2020 |
| MovieLens 1M  | GLocal-K [125]                      | 0.823 (RSME)    | N                   | 2021 |
|               | Sparse FC [127]                     | 0.824 (RSME)    | N                   | 2018 |
|               | CF-NADE [128]                       | 0.829 (RSME)    | N                   | 2016 |
| MovieLens 10M | Bayesian timeSVD++ flipped [129]    | 0.749 (RSME)    | N                   | 2019 |
|               | Bayesian timeSVD++ [129]            | 0.752 (RSME)    | N                   | 2019 |
|               | Bayesian SVD++ [129]                | 0.750 (RSME)    | N                   | 2019 |
|               | scaled-CER (Item cold-start) [69]   | 0.035 (MAP@15)  | N                   | 2025 |
| MovieLens 20M | VASP [130]                          | 0.448 (nDCG@10) | Y                   | 2021 |
|               | H-Vamp Gated [131]                  | 0.445 (nDCG@10) | Y                   | 2019 |
|               | RecVAE [132]                        | 0.442 (nDCG@10) | N                   | 2019 |
| Amazon-Book   | SSCF [133]                          | 0.066 (nDCG@20) | N                   | 2022 |
|               | SANS [134]                          | 0.064 (nDCG@20) | N                   | 2023 |
|               | BCPM-LM [135]                       | 0.064 (nDCG@20) | N                   | 2022 |
| Netflix       | H-Vamp Gated [131]                  | 0.409 (nDCG@10) | N                   | 2019 |
|               | RecVAE [132]                        | 0.394 (nDCG@10) | N                   | 2019 |
|               | EASE [136]                          | 0.393 (nDCG@10) | N                   | 2019 |
| LastFM        | Ekar [137]                          | 0.248 (HR@10)   | N                   | 2019 |
|               | HAKG [138]                          | 0.093 (nDCG@20) | N                   | 2022 |
|               | KONN-LS [139]                       | 0.370 (Rec@10)  | N                   | 2019 |
| Yelp          | DoRec [140]                         | 0.142 (nDCG)    | N                   | 2019 |
|               | RSPM-EM [141]                       | 0.160 (nDCG@20) | N                   | 2022 |
| Gowalla       | MGDCF [142]                         | 0.159 (nDCG@20) | N                   | 2024 |
|               | UltraGCN [143]                      | 0.158 (nDCG@20) | N                   | 2021 |

also offer transparent explanations for their recommendations [150]. Techniques such as model-agnostic explanation methods, rule-based systems, or attention mechanisms can be explored to enhance interpretability. Memory Network-Based Interpreter [151] and experience-driven interpreter [152] may offer viable solutions to this challenge.

B. FAIRNESS AND BIAS
There is a pressing need to create fairness-aware recommendation algorithms that mitigate biases and ensure equitable treatment across diverse user groups [153]. This could involve developing fairness metrics, fairness-aware loss functions, or debiasing techniques tailored for recommendation systems. Especially for debiasing, Chen et al. [154] introduces every clear and they collect latest approaches for debiasing in different strategies at https://github.com/jiawei-chen/RecDebiasing.

C. CONTEXT-AWARE RECOMMENDATIONS
Advanced research may explore leveraging deep learning architectures, such as recurrent neural networks (RNNs) or transformers, to effectively capture and utilize contextual information in recommendation models [155]. Exploring multi-modal learning to incorporate various types of contextual data, such as textual, visual, or temporal cues, could significantly enhance recommendation accuracy.

D. COLD-START PROBLEM
Addressing the cold-start problem could involve exploring transfer learning approaches, where knowledge learned from related domains or auxiliary data sources is transferred to alleviate the cold-start problem [156]. Integrating multi-modality features is a potential solution for this direction [69]. Hybrid recommendation techniques combining collaborative filtering, content-based methods, and knowledge graphs could also be investigated to handle cold-start scenarios more effectively.

E. LONG-TAIL RECOMMENDATIONS
Future efforts could focus on creating specialized algorithms for long-tail recommendations, such as mixture models, probabilistic graphical models, or ensemble methods tailored to capture rare item preferences [157]. Techniques like active learning or diversity-promoting algorithms could improve coverage of long-tail items.

F. DYNAMIC AND ADAPTIVE RECOMMENDATIONS
Future research might investigate reinforcement learning frameworks for recommendation systems, where agents learn optimal recommendation policies through interaction with users and feedback [158]. Reinforcement learning frameworks, particularly those incorporating deep learning, have demonstrated marked improvements over traditional methods in recommendation systems by effectively capturing...
complex user behaviors and preferences. They excel in environments where large datasets and dynamic interaction are involved [159], leading to increased personalization and user satisfaction. Industries like e-commerce, entertainment, and social media have notably benefited [160], where these systems offer enhanced user engagement by dynamically adapting to user interests. Moreover, specific applications like course recommendations in MOOCs and personalized advertising showcase the potential of reinforcement learning to handle multifaceted tasks and achieve superior outcomes compared to traditional, static methods [161]. Additionally, Online learning algorithms that continuously update in response to evolving user preferences could also be explored to keep recommendations relevant.

G. Privacy-Preserving Recommendations

Research could explore techniques such as federated learning, where recommendation models are trained across decentralized data sources without sharing raw user data [162]. Differential privacy mechanisms could also be integrated into recommendation algorithms like federated search techniques [163] to ensure individual user privacy while maintaining recommendation utility.

H. Multi-Stakeholder Recommendations

Advanced research may focus on developing multi-objective optimization frameworks for recommendation systems, considering the diverse interests of users, providers, and advertisers simultaneously [164]. Game-theoretic approaches could be employed to model interactions among stakeholders and design recommendation strategies that optimize collective utility while addressing conflicting objectives.

I. Future Exploration

In conclusion, the field of recommendation systems on big data is ripe with opportunities for future exploration and innovation. By addressing these open questions and exploring new research directions, the academic and technological communities can advance the state-of-the-art in recommendation technologies, unlocking new possibilities for personalized user experiences across various application domains.

VII. Conclusion

This survey paper offers an exhaustive overview of recommendation systems, a technological innovation that has seen widespread adoption in various web-based applications in recent years. The primary objective of modern recommendation systems is to furnish users with personalized suggestions for online products or services, employing a range of techniques including content-based, collaborative filtering, knowledge-based, and hybrid approaches to cater to diverse requirements across different scenarios.

The manuscript delves into a comprehensive historical review and a critical examination of the state-of-the-art methodologies in recommendation systems, with a special emphasis on the pioneering developments brought about by the advent of big data analytics. Notably, this paper highlights the utilization of prominent datasets such as MovieLens, Amazon Reviews, Netflix Prize, Last.fm, and Yelp in evaluating recommendation algorithms. Additionally, this paper scrutinizes the prevalent challenges encountered in contemporary recommendation systems, such as data sparsity, scalability issues, and the need for diversity in recommendations, proposing these hurdles as fertile grounds for future research endeavors.

Simultaneously, this survey extends its analysis to the application of recommendation systems within various life-related domains, including marketing, governance, medical, health, and the promotion of sustainable lifestyles. This exploration aims to provide a foundational understanding of how recommendation systems intersect with everyday life, highlighting their significance in shaping user experiences and influencing societal trends. Through this holistic review, the paper endeavors to present a nuanced perspective on the evolution of recommendation systems and their growing impact on digital consumer culture and beyond.

References

[1] F. Zhou, B. Luo, T. Hu, Z. Chen, and Y. Wen, “A combinatorial recommendation system framework based on deep reinforcement learning,” in Proc. IEEE Int. Conf. Big Data (Big Data), Dec. 2021, pp. 5733–5740.
[2] H. Wang, N. Lou, and Z. Chao, “A personalized movie recommendation system based on LSTM-CNN,” in Proc. 2nd Int. Conf. Mach. Learn., Big Data Bus. Intell. (MLDBDI), Oct. 2020, pp. 485–490.
[3] G. Adomavicius and A. Tuzhilin, “Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions,” IEEE Trans. Knowl. Data Eng., vol. 17, no. 6, pp. 734–749, Jun. 2005.
[4] T. Omura, K. Suzuki, P. Siriaraya, M. Mittal, Y. Kawai, and S. Nakajima, “Ad recommendation utilizing user behavior in the physical space to represent their latent interest,” in Proc. IEEE Int. Conf. Big Data (Big Data), Dec. 2020, pp. 3143–3146.
[5] N. Entezari, E. E. Papalexakis, H. Wang, S. Rao, and S. K. Prasad, “Tensor-based complementary product recommendation,” in Proc. IEEE Int. Conf. Big Data (Big Data), Dec. 2021, pp. 409–415.
[6] R. Rismanto, A. R. Syulistyo, and B. P. C. Agusta, “Research supervisor recommendation system based on topic conformity,” Int. J. Mod. Educ. Comput. Sci., vol. 12, no. 1, pp. 26–34, Feb. 2020.
[7] Z. Cui, X. Xu, X. Xue, X. Cai, Y. Cao, W. Zhang, and J. Chen, “Personalized recommendation system based on collaborative filtering for IoT scenarios,” IEEE Trans. Services Comput., vol. 13, no. 4, pp. 685–695, Jul. 2020.
[8] F. Ali, D. Kwak, P. Khan, S. H. A. Ei-Sappagh, S. M. R. Islam, D. Park, and K.-S. Kwak, “Merged ontology and SVM-based information extraction and recommendation system for social robots,” IEEE Access, vol. 5, pp. 12364–12379, 2017.
[9] B. Li, A. Maalla, and M. Liang, “Research on recommendation algorithm based on e-commerce user behavior sequence,” in Proc. IEEE 2nd Int. Conf. Inf. Technol., Big Data Artif. Intell. (ICIBA), vol. 2, Dec. 2021, pp. 914–918.
[10] X. Li and F. Sun, “Sports training recommendation method under the background of data analysis,” in Proc. Int. Conf. High Perform. Big Data Intell. Syst. (HPBD&IS), Dec. 2021, pp. 12–16.
[11] T. Nunnoonda, “A real-time recommendation engine using lambda architecture,” Artif. Life Robot., vol. 23, no. 2, pp. 249–254, Jun. 2018.
[12] J. Xiao, M. Wang, B. Jiang, and J. Li, “A personalized recommendation system with combinational algorithm for online learning,” J. Ambient Intell. Humanized Comput., vol. 9, no. 3, pp. 667–677, Jun. 2018.
[13] Z. Huang, X. Xu, J. Ni, H. Zhu, and C. Wang, “Multimodal representation learning for recommendation in Internet of Things,” IEEE Internet Things J., vol. 6, no. 6, pp. 10675–10685, Dec. 2019.
[14] H. Zhang, T. Huang, Z. Lv, S. Liu, and Z. Zhou, “MCRS: A course recommendation system for MOOCs,” Multimedia Tools Appl., vol. 77, no. 6, pp. 7051–7069, Mar. 2018.

[15] I. Benouaret and S. Amer-Yahia, “A comparative evaluation of top-N recommendation algorithms: Case study with total customers,” in Proc. IEEE Int. Conf. Big Data (Big Data), Dec. 2020, pp. 4499–4508.

[16] A. Sun, J. J. Londoño, B. Elbaum, L. Estrada, R. J. Lazo, L. Vitale, H. G. Villasanti, R. Fusaroli, L. K. Perry, and D. S. Messinger, “Who said what? An automated approach to analyzing speech in preschool classrooms,” 2024, arXiv:2401.07342.

[17] J.-H. Huh and T.-J. Kim, “A location-based mobile health care facility search system for senior citizens,” J. Supercomput., vol. 75, no. 4, pp. 1831–1848, Apr. 2019.

[18] M. Naumov et al., “Deep learning recommendation model for personalization and recommendation systems,” 2019, arXiv:1906.00091.

[19] J. Schneider and M. Vlachos, “Personalization of deep learning,” in Proc. 3rd Int. Data Sci. Conf. (iDSC), Cham, Switzerland: Springer, Jan. 2021, pp. 89–96.

[20] L. Ni, H. Lin, M. Zhang, and J. Zhang, “Hybrid filtrations recommendation system based on privacy preserving in edge computing,” Proc. Comput. Sci., vol. 129, pp. 407–409, Aug. 2018.

[21] S. Milano, M. Taddeo, and L. Floridi, “Ethical aspects of multi-stakeholder recommendation systems,” Inf. Soc., vol. 37, no. 1, pp. 35–45, Jan. 2021.

[22] P. Merinov, “Sustainability-oriented recommender systems,” inProc. 31st ACM Conf. User Model., Adaptation Personalization, Jun. 2023, pp. 296–300.

[23] B. Yi, X. Shen, H. Liu, Z. Zhang, W. Zhang, S. Liu, and N. Xiong, “Deep matrix factorization with implicit feedback embedding for recommendation system,” IEEE Trans. Ind. Inform., vol. 15, no. 8, pp. 4591–4601, Aug. 2019.

[24] P. Lops, M. De Gemmis, and G. Semeraro, “Content-based recommender systems: State of the art and trends,” in Recommender Systems Handbook, Springer, 2011, pp. 73–105.

[25] M. Elahi, F. Ricci, and N. Rubens, “A survey of active learning in collaborative filtering recommender systems,” Comput. Sci. Rev., vol. 20, pp. 29–50, May 2016.

[26] B. Alhijawi and Y. Kilani, “Using genetic algorithms for measuring the similarity values between users in collaborative filtering recommender systems,” in Proc. IEEE/ACIS 15th Int. Conf. Comput. Inf. Sci. (ICIS), Jun. 2016, pp. 1–6.

[27] L. A. Hassanich, C. A. Joudeh, J. B. Abdo, and J. Demerjian, “Similarity measures for collaborative filtering recommender systems,” in Proc. IEEE Middle East North Afr. Commun. Conf. (MENACOMM), Apr. 2018, pp. 1–5.

[28] F. Zhang, T. Gong, V. E. Lee, G. Zhao, C. Rong, and G. Qu, “Fast algorithms to evaluate collaborative filtering recommender systems,” Knowl.-Based Syst., vol. 96, pp. 96–103, Mar. 2016.

[29] C. Musto, G. Semeraro, M. De Gemmis, and P. Lops, “Learning word embeddings from Wikipedia for content-based recommender systems,” in Proc. Eur. Conf. Inf. Retr. Cham, Switzerland: Springer, Jan. 2016, pp. 729–734.

[30] M. Volkovs, G. W. Yu, and T. Poutanen, “Content-based neighbor models for cold start in recommender systems,” in Proc. Recommender Syst. Challenge, Jul. 2017, pp. 1–6.

[31] D. Mittal, S. Shandilya, D. Khirwar, and A. Bhise, “Smart billing using content-based recommender systems based on fingerprint,” in ICT Applications and Analysis, Cham, Switzerland: Springer, 2020, pp. 85–93.

[32] Y. Pérez-Almaguer, R. Yera, A. A. Alzahrani, and L. Martínez, “Content-based group recommender systems: A general taxonomy and further improvements,” Expert Syst. Appl., vol. 184, Dec. 2021, Art. no. 115444.

[33] A. S. Tewari, “Generating items recommendations by fusing content and user-item based collaborative filtering,” Proc. Comput. Sci., vol. 167, pp. 1934–1940, Jun. 2020.

[34] C. Ajaegbu, “An optimized item-based collaborative filtering algorithm,” J. Ambient Intell. Humanized Comput., vol. 12, no. 12, pp. 10629–10636, Dec. 2021.

[35] J. Bobadilla, S. Alonso, and A. Hernando, “Deep learning architecture for collaborative filtering recommender systems,” Appl. Sci., vol. 10, no. 7, p. 2441, Apr. 2020.

[36] J. Bobadilla, F. Ortega, A. Gutiérrez, and S. Alonso, “Classification-based deep neural network architecture for collaborative filtering recommender systems,” Int. J. Interact. Multimedia Artif. Intell., vol. 6, no. 1, pp. 1–68, 2020.

[37] M. Dong, X. Zeng, L. Koehl, and J. Zhang, “An interactive knowledge-based recommender system for fashion product design in the big data environment,” Inf. Sci., vol. 540, pp. 469–488, Nov. 2020.

[38] A. Gazdar and L. Hidri, “A new similarity measure for collaborative filtering based recommender systems,” Knowl.-Based Syst., vol. 188, Jan. 2020, Art. no. 105058.

[39] P. M. Alamdari, N. J. Navimipour, M. Hosseinzadeh, A. A. Safaei, and A. Darwesh, “A systematic study on the recommender systems in the e-commerce,” IEEE Access, vol. 8, pp. 115694–115716, 2020.

[40] F. Cena, L. Console, and F. Vermero, “Logical foundations of knowledge-based recommender systems: A unifying spectrum of alternatives,” Inf. Sci., vol. 546, pp. 60–73, Feb. 2021.

[41] B. Hrnjica, D. Music, and S. Softic, “Model-based recommender systems,” in Trends in Cloud-Based IoT, Springer, 2020, pp. 125–146.

[42] J. Shokeen and C. Rana, “A study on features of social recommender systems,” Artif. Intell. Rev., vol. 53, no. 2, pp. 965–988, Feb. 2020.

[43] A. Zagranovskia and D. Mitura, “Designing hybrid recommender systems,” in Proc. 4th Int. Syst. Practical Conf., Mar. 2021, pp. 1–5.

[44] A. J. Ibrahim, P. Zira, and N. Abdulganiyyi, “Hybrid recommender for research papers and articles,” Int. J. Intell. Syst., vol. 10, no. 9, 2021.

[45] S. Renjith, A. Sreekumar, and M. Jathavedan, “An extensive study on the evolution of context-aware personalized travel recommender systems,” Inf. Process. Manage., vol. 57, no. 1, Jan. 2020, Art. no. 102078.

[46] L. Quijano-Sánchez, I. Cantador, M. E. Cortés-Cediel, and O. Gil, “Recommender systems for smart cities,” Inf. Syst., vol. 92, Sep. 2020, Art. no. 101545.

[47] V. Riabchuk, L. Hagel, F. Germaine, and A. Zharova, “Utility-based context-aware multi-agent recommendation system for energy efficiency in residential buildings,” Infusion, vol. 112, Dec. 2024, Art. no. 102559.

[48] Y. Deldjoo, M. Schell, P. Cremonesi, and G. Pasi, “Recommender systems leveraging multimedia content,” ACM Comput. Surv., vol. 53, no. 5, pp. 1–38, Sep. 2021.

[49] Q. Shambour, “A deep learning based algorithm for multi-criteria recommender systems,” Knowl.-Based Syst., vol. 211, Jan. 2021, Art. no. 106545.

[50] R. Cabezas, J. G. Ruiz, and M. Leyva, “A knowledge-based recommender system,” Inf. Fusion, vol. 84, Jun. 2020, pp. 115694–115716, 2020.

[51] J. K. Tarus, Z. Niu, and G. Mustafa, “Knowledge-based recommendation: A review of ontology-based recommender systems for e-learning,” Artif. Intell. Rev., vol. 50, no. 1, pp. 21–48, Jan. 2018.
A. V. Dev and A. Mohan, “Recommendation system for big data applica-
tions based on set similarity of user preferences,” in Proc. \textit{Int. Conf. Next
Gener. Intell. Syst. (ICNGIS)}, Sep. 2016, pp. 1–6.

J. Chen, K. Li, H. Rong, K. Bilal, N. Yang, and K. Li, “A disease diagnosis
and treatment recommendation system based on big data mining and
cloud computing,” \textit{Inf. Sci.}, vol. 435, pp. 124–149, Apr. 2018.

Z. Wan, “Research on e-commerce recommendation system based on
big data technology,” \textit{J. Phys. Conf. Ser.}, vol. 1883, no. 1, 2021,
Art. no. 012159.

A. Banu and S. Banu, “Keyword based movie recommendation service
using MapReduce,” \textit{Int. J. Eng. Res.}, vol. 5, no. 1, pp. 1–3, Jan. 2016.

J. D. West, I. Wesley-Smith, and C. T. Bergstrom, “A recommendation
system using multi-criteria recommender systems,” in \textit{Proc. Int. Conf.
Cyberworlds (CW)}, Sep. 2016, pp. 155–158.

G. George and A. M. Lal, “Review of ontology-based recommender sys-
tems in e-learning,” \textit{Comput. Educ.}, vol. 142, Dec. 2019, Art. no. 103642.

J. Feng, Z. Xia, X. Feng, and J. Peng, “RBPR: A hybrid model for the new
user cold start problem in recommender systems,” \textit{Knowl.-Based Syst.},
vol. 214, Feb. 2021, Art. no. 106732.

S. D. Kadam, D. Motwani, and S. A. Vaidya, “Big data analytics-recommendation system with Hadoop framework,” in \textit{Proc. Int. Conf. Inventive Comput. Technol. (ICICT)}, vol. 3, Aug. 2016,
pp. 1–5.

D. P. Acharya and K. Ahmed, “A survey on big data analytics: Chal-
lenge, open research issues and tools,” \textit{Int. J. Adv. Comput. Sci. Appl.},
vol. 7, no. 2, pp. 511–518, 2016.

X. Zhou, W. Liang, K. I. Wang, R. Huang, and Q. Jin, “Aca-
demic influence aware and multidimensional network analysis for
research collaboration navigation based on scholarly big data,” \textit{IEEE
Trans. Emerg. Topics Comput.}, vol. 9, no. 1, pp. 246–257, Jan. 2021.

P. R. M. Rao, S. M. Krishna, and A. P. S. Kumar, “Privacy preservation
 techniques in big data analytics: A survey,” \textit{J. Big Data}, vol. 5, no. 1,
pp. 1–12, Dec. 2018.

X. Wu, X. Zhu, G.-Q. Wu, and W. Ding, “Data mining with big data,”
\textit{IEEE Trans. Knowl. Data Eng.}, vol. 26, no. 1, pp. 97–107, Jan. 2014.

C. K. Emami, N. Cullot, and C. Nicolle, “Understandable big data: A
survey,” \textit{Comput. Sci. Rev.}, vol. 17, pp. 70–81, Aug. 2015.

H.-Y. Lin and S.-Y. Yang, “A cloud-based data mining data information
agent system based on big data analysis technology,” \textit{Microelectron. Rel.},
vol. 97, pp. 66–78, Jun. 2019.

Y. Cheng and X. Bu, “Research on key technologies of personalized edu-
cation resource recommendation system based on big data environment,”
\textit{J. Phys., Conf. Ser.}, vol. 1437, no. 1, Jan. 2020, Art. no. 012024.

K. A. Farhan, F. Nafis, B. Aghoutane, A. Yahiyouy, J. Riffi, and
A. Sabri, “Hybrid recommender system for tourism based on big data
and AI: A conceptual framework,” \textit{Big Data Mining Anal.}, vol. 4, no. 1,
pp. 47–55, Mar. 2021.

A. V. Dev and A. Mohan, “Recommendation system for big data applica-
tions based on set similarity of user preferences,” in \textit{Proc. Int. Conf. Next
Gener. Intell. Syst. (ICNGIS)}, Sep. 2016, pp. 1–6.

J. Chen, K. Li, H. Rong, K. Bilal, N. Yang, and K. Li, “A disease diagnosis
and treatment recommendation system based on big data mining and
cloud computing,” \textit{Inf. Sci.}, vol. 435, pp. 124–149, Apr. 2018.
A. Coulter and J. Ellins, “Effectiveness of strategies for informing, educating, and involving patients,” *Inf. Fusion*, vol. 72, pp. 1–21, Aug. 2021.

M. T. Siddique, P. Koukars, D. Ioanidou, and C. Tjortjis, “SmartBuild RecSys: A recommendation system based on the smart readiness indicator for energy efficiency in buildings,” *Algorithms*, vol. 16, no. 10, p. 482, Oct. 2023.

B. J. Gogoi, “Green building features and factors affecting the consumer choice for green building recommendation,” *Int. J. Civil Eng. Technol.*, vol. 9, no. 6, pp. 127–136, 2018.

J. Veselka, M. Nehasilova, K. Dvostaková, P. Ryklová, M. Volf, J. Ružička, and A. Luptíček, “Recommendations for developing a BIM for the purpose of LCA in green building certifications,” *Sustainability*, vol. 12, no. 15, p. 6151, Jul. 2020.

M. Cheng and H. Gao, “Data-driven life-cycle risk assessment of bridge networks using Bayesian network,” in *Life-Cycle of Structures and Infrastructure Systems*. Boca Raton, FL, USA: CRC Press, 2023, pp. 1953–1960.

J. D. Gil, “Recommender system for sustainable alternatives of building materials: Sustainability-focused decision making in the construction industry,” Master’s thesis, LUT School Eng. Sci., Lappeenranta–Lahti Univ. Technol. LUT, Finland, 2023.

A. Sun, Y. Tao, M.-L. Shyu, S.-C. Chen, A.Blizzard, W. A. Rothenberg, D. Garcia, and J. F. Jent, “Multimodal data integration for interactive and realistic avatar simulation in augmented reality,” in *Proc. IEEE 22nd Int. Conf. Inf. Reuse Integ. Data Sci. (IRI)*, Aug. 2021, pp. 362–369.

A. B. Lodhi, M. Bilal, H. S. M. Bilal, K. U. Khan, F. A. Satti, S. Khalid, and S. Lee, “PNRG: Knowledge graph-driven methodology for person-alized nutritional recommendation generation,” in *Proc. Int. Conf. Smart Homes Health Telematics*. Cham, Switzerland: Springer, Jan. 2023, pp. 230–258.

A. Sun, E. J. Franzmann, Z. Chen, and X. Cai, “Deep contrastive learning for predicting cancer prognosis using gene expression values,” *Briefings Bioinf.*, vol. 25, no. 6, p. 544, Sep. 2024.

K. G. Shojaiania, A. Jennings, C. R. Ramsay, J. M. Grimshaw, J. L. Kwan, and L. Lo, “The effects of on-screen, point of care computer reminders on processes and outcomes of care,” *Cochrane Database Systematic Rev.*, vol. 2021, no. 3, p. 153, Jul. 2020.

A. Coulter and J. Ellins, “Effectiveness of strategies for informing, educating, and involving patients,” *BMJ*, vol. 335, no. 7609, pp. 24–27, Jul. 2007.

R. L. Street, G. Makoul, N. K. Arora, and R. M. Epstein, “How does communication heal? Pathways linking clinician–patient communication to health outcomes,” *Patient Educ. Counseling*, vol. 74, no. 3, pp. 295–301, Mar. 2009.

X. Li and H. Wang, “A survey on resource allocation in cloud computing: Issues and solutions,” *J. Netw. Comput. Appl.*, vol. 80, pp. 90–106, Jan. 2017.

S. C. Brailsford, J. Vissers, and H. Buxton, “A comprehensive review of patient simulation in nursing education,” *Simul. Gaming*, vol. 40, no. 4, pp. 529–565, 2009.

F. Fatehi and R. Wootten, “A clinician’s guide to the use of artificial intelligence in telemedicine—A focus on triage services,” *Stud. Health Technol. Inform.*, vol. 252, pp. 46–51, Aug. 2018.

E. R. Doresey and E. J. Topol, “State of telehealth,” *New England J. Med.*, vol. 375, no. 2, pp. 154–161, 2016.

D. H. Al-Shammary, A. A. Rahman, and O. A. Moti, “Collaborative learning in healthcare: A systematic review,” *J. Med. Sci.*, vol. 43, no. 9, p. 279, Oct. 2019.

W. S. Choi and M. J. Rho, “A systematic review of research trends and future directions in gamification,” in *Proc. Extended Abstr. CHI Conf. Hum. Factors Comput. Syst.*, 2018, p. 511.

Z. Z. Darban and M. H. Valipour, “GHRS: Graph-based hybrid recommendation system with application to movie recommendation,” *Expert Syst. Appl.*, vol. 200, Aug. 2022, Art. no. 116850.

S. C. Han, T. Lim, S. Long, B. Burgstaller, and J. Poon, “GLocal-K: Global and local kernels for recommender systems,” in *Proc. 30th ACM Int. Conf. Inf. Knowl. Manage.*, 2021, pp. 3063–3067.
Z. Xia et al.: Contemporary Recommendation Systems on Big Data and Their Applications: A Survey

[151] N.-T. Tran and H. W. Lauw, “Memory network-based interpreter of user preferences in content-aware recommender systems,” ACM Trans. Intell. Syst. Technol., vol. 14, no. 6, pp. 1–28, Dec. 2023.

[152] X. Huang, T. Chyczewski, Z. Xia, R. Kunz, and X. Yang, “Distilling experience into a physically interpretable recommender system for computational model selection,” Sci. Rep., vol. 13, no. 1, p. 2225, Feb. 2023.

[153] M. Zehlike, C. Castillo, and M. Lalmas, “Fairness-aware recommendation in social networks,” ACM Trans. Inf. Syst., vol. 36, no. 3, pp. 1–33, 2018.

[154] J. Chen, H. Dong, X. Wang, F. Feng, M. Wang, and X. He, “Bias and debias in recommender system: A survey and future directions,” ACM Trans. Inf. Syst., vol. 41, no. 3, pp. 1–39, Jul. 2023.

[155] G. Adomavicius and A. Tuzhilin, “Context-aware recommender systems,” AI Mag., vol. 32, no. 3, pp. 67–80, 2011.

[156] S. J. Pan and Q. Yang, “Transfer learning for collaborative filtering,” in Proc. 32nd Int. ACM SIGIR Conf. Res. Develop. Inf. Retr., 2009, pp. 101–108.

[157] M. Ge, C. Delgado-Battenfeld, and D. Jannach, “The long tail of recommender systems and how to leverage it,” in Proc. ACM Conf. Recommender Syst., Oct. 2008, pp. 11–18.

[158] A. Zela, A. Bansal, M. Farajtabar, B. Lakshminarayanan, Y. Cheng, D. Schuurmans, and G. E. Hinton, “Reinforcement learning for bandit neural architecture search,” 2019, arXiv:1909.03574.

[159] Y. Zhang, C. Zhang, and X. Liu, “Dynamic scholarly collaborator recommendation via competitive multi-agent reinforcement learning,” in Proc. 11th ACM Conf. Recommender Syst., Aug. 2017, pp. 331–335.

[160] R. Takanobu, T. Zhuang, M. Huang, J. Feng, H. Tang, and B. Zheng, “Aggregating e-commerce search results from heterogeneous sources via hierarchical reinforcement learning,” in Proc. World Wide Web Conf., May 2019, pp. 1771–1781.

[161] J. Zhang, B. Hao, B. Chen, C. Li, H. Chen, and J. Sun, “Hierarchical reinforcement learning for course recommendation in MOOCs,” in Proc. AAAI Conf. Artif. Intell., Jul. 2019, vol. 33, no. 1, pp. 435–442.

[162] J. Li, Q. Li, and Z. Gao, “Privacy-preserving collaborative filtering using randomized perturbation-based approach,” Inf. Sci., vol. 478, pp. 54–71, Apr. 2019.

[163] A. Garba, S. Wu, and S. Khalid, “Federated search techniques: An overview of the trends and state of the art,” Knowl. Inf. Syst., vol. 65, no. 12, pp. 5065–5095, Dec. 2023.

[164] H. Abdollahpour, R. Burke, and B. Mobasher, “Multistakeholder recommendation: Survey and research directions,” in Proc. 13th ACM Conf. Recommender Syst., 2019, pp. 334–338.

ZIYUAN XIA received the B.A. degree in economics from the University of California, Davis, in 2019, and the M.S. degree in management of technology from New York University, in 2021. He is currently pursuing the Ph.D. degree with the Antai College of Economics and Management, Shanghai Jiao Tong University. His research interests include ESG, strategic management, and managerial application of machine learning.

ANCHEN SUN (Member, IEEE) received the B.S. degree in marine and atmospheric science and computer science and the M.S. degree in electrical and computer engineering from the University of Miami, Coral Gables, FL, USA, in 2017 and 2020, respectively, where he is currently pursuing the Ph.D. degree with the Department of Electrical and Computer Engineering. His research interests include deep learning, bioinformatics, artificial intelligence, and engineering management. He received the Elihu L. and Joyce Jury Award in 2024.

JINGYI XU received the B.A. degree from the School of Architecture, University of Miami, in 2020, and the M.S. degree in advanced architectural design from Cornell University, in 2021. She is currently an architectural design professional with Hellmuth, Obata + Kassabaum. Her research interests include architectural design, data-driven methodologies, sustainable green building practices, and computational approaches to architecture.

YUANZHE PENG is currently pursuing the Ph.D. degree with the Department of Electrical and Computer Engineering, Herbert Wertheim College of Engineering, University of Florida, Gainesville, FL, USA. His research interests include edge computing, wireless communications, and networking, with particular emphasis on the optimization of multimodal federated learning, as well as privacy and security in the Internet of Things.

RUI MA received the Ph.D. degree in electrical and computer engineering from the University of Miami. He is currently a Postdoctoral Researcher with the Bascom Palmer Eye Institute, Miller School of Medicine, University of Miami. His research interests include developing advanced deep learning and reinforcement learning algorithms to enhance the quality of visual field tests, predict visual field progression, and analyze multimodal ophthalmic imaging data.

MINGHUI CHENG received the B.E. degree in civil engineering from Hunan University, China, in 2016, and the Ph.D. degree in structural engineering from Lehigh University, in 2021. He is currently an Assistant Professor jointly appointed by the Department of Civil and Architectural Engineering and the School of Architecture, University of Miami. Previously, he was a Postdoctoral Associate with the Systems Engineering, Cornell University. His research interests include the digital twin-enabled smart city and sustainability-informed life-cycle management of structures and infrastructure systems with an emphasis on surrogate modeling, optimization, and machine learning.

* * *