

Eliciting Auxiliary Information for Cold Start User Recommendation: A Survey

Nor Aniza Abdullah ^{1,*}, Rasheed Abubakar Rasheed ², Mohd Hairul Nizam ¹ and Md Mujibur Rahman ¹

¹ Department of Computer System & Technology, Faculty of Computer Science and Information Technology, Universiti Malaya, Kuala Lumpur 50603, Malaysia; hairulnizam@um.edu.my (M.H.N.); mujib.rahman.bd@gmail.com (M.M.R.)

² Department of Software Engineering, Faculty of Computer Science and Information Technology, Bayero University Kano, Kano State 3011, Nigeria; rarasheed.se@buk.edu.ng

* Correspondence: noraniza@um.edu.my

Abstract: Recommender systems suggest items of interest to users based on their preferences. These preferences are typically generated from user ratings of the items. If there are no ratings for a certain user or item, it is said that there is a cold start problem, which leads to unreliable recommendations. Existing studies that reviewed and examined cold start in recommender systems have not explained the process of deriving and obtaining the auxiliary information needed for cold start recommendation. This study surveys the existing literature in order to explain the various approaches and techniques employed by researchers and the challenges associated with deriving and obtaining the auxiliary information necessary for cold start recommendation. Results show that auxiliary information for cold start recommendation is obtained by adapting traditional filtering and matrix factorization algorithms typically with machine learning algorithms to build learning prediction models. The understanding of similar or connected user profiles can be used as auxiliary information for building cold start user profile to enable similar recommendations in social networks. Similar users are clustered into sub-groups so that a cold start user could be allocated and inferred to a sub-group having similar profiles for recommendations. The key challenges of the process for obtaining the auxiliary information involve: (1) two separate recommendation processes of conversion from pure cold start to warm start before eliciting the auxiliary information; (2) the obtained implicit auxiliary information is usually ranked and sieved in order to select the top rated and reliable auxiliary information for the recommendation. This study also found that cold start user recommendation has frequently been researched in the entertainment domain, typically using music and movie data, while little research has been carried out in educational institutions and academia, or with cold start for mobile applications.

Keywords: recommender systems; cold start; auxiliary information; machine learning algorithms

Citation: Abdullah, N.A.; Rasheed, R.A.; Nizam, M.H.; Rahman, M.M. Eliciting Auxiliary Information for Cold Start User Recommendation: A Survey. *Appl. Sci.* **2021**, *11*, 9608. <https://doi.org/10.3390/app11209608>

Academic Editor: Luis Javier García Villalba

Received: 7 September 2021

Accepted: 10 October 2021

Published: 15 October 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Recommender systems in large numbers are in operation today. These systems are based on diverse techniques, approaches, and are used to provide recommendations in various disciplines and application domains. The development of these recommender systems, as well as their improvement for various domains and purposes has given rise to an active area of scientific research [1]. This development and evolvement of recommender systems is based on the continuing evolution of artificial intelligence, information retrieval, machine learning, statistical methods, data mining, etc. [2]. Recommender systems are highly successful on e-commerce web sites capable of recommending the products that customers would probably like.

Recommender systems are software tools that recommend users to items or items of interest to users based on their preferences [3]. These systems used either implicit or

explicit data/information to determine users' preferences. Explicit information involves data collection from surveys or questionnaire to understand users' preferences, while implicit information retrieval is basically an indirect method which involves culling information from online platforms such as purchasing history, demographics, likes and other traces of users' online activity. Recommender systems have boosted the existence and success of several application domains such as e-commerce, online entertainment, tourism industry, etc., in which customers are recommended and become more attractive to these products or materials due to the accuracy of recommendation offered by these software systems.

Recommender systems use algorithms; these algorithms are typically based on filtering strategies that utilize the available information about items or users to provide recommendations. Content based and collaborative filtering are the two most commonly used filtering strategies [3]. Content based is based on the analogy that if a user prefers an item, then the user would most likely prefer similar items of similar characteristics. There are two main collaborative filtering approaches: latent factor model and neighborhood-based approaches [4,5]. The neighborhood-based approach is based on recommending items to a group of users sharing similar characteristics with the notion that users having similar characteristics would most likely have similar preferences. In this approach, a product or item is selected for recommendation if one neighborhood member is determined for a user and an item liked by the neighbors is selected for recommendation. On the other hand, the latent factor model-based approach attempts to determine factors that help in understanding users' psychology and personalities to determine the suitable item for recommendation. Both these filtering strategies train models based on prior information which are used for the item recommendation usually in the form of a ratings matrix which contains the ratings given by different users to various items. Moreover, collaborative filtering exhibit better performance and ease of use compared to content-based filtering [3–6]. Although, combining both content-based and collaborative strategies as a hybrid form have been a promising approach in affording the strengths of both filtering strategies [3,7,8].

When a new user interacts with a recommender system or a new item is added to the repository of a recommender system, the system faces a challenge in giving recommendations due to the unavailability of prior information about the user or the item; this problem is called cold start problem [9–11] in recommender systems. There are two main types of cold start problems in recommender systems: new user cold start problem and new item cold start problem. In the new user cold start problem, the recommender system faces a problem in offering recommendations when the new user is introduced to the system as it has no prior information about the user. In the new item cold start problem, the recommender system has no ratings for a newly added item, so it faces difficulty in recommending the right item for a user, or the right user for a particular item. Between these two types of cold start problems, the new user cold start problem has been more widely studied due to the challenge of accurately recommending an item to a new user [9,12]. Similarly, there are also two types of cold start problem in terms of data sparsity: Pure cold start, and warm start. Pure cold start or cold start refers to a situation in which there are no available ratings or past information about a user or item, while warm start refers to a situation whereby the available ratings, information or preference is sparse and is insufficient or could not be used to provide recommendation. However, researchers have used several terminologies such as “complete cold start” or “pure cold start” to indicate the complete absence of ratings; or the use of “incomplete cold start” or “warm start” to indicate the existence of very little information or preference for recommendation (typically rating by a single user).

Over the years, researchers in various domains have leverage the use of implicit data in addressing new user cold start problems. Researchers have devised the use elicit implicit data from various sources such as social network streams like Twitter, Facebook, background information, etc., to provide new user with better recommendations. One of

the most common approaches to mitigate cold start for user is by adapting or coupling the traditional collaborative filtering techniques with other techniques most notably machine learning algorithms. Although there have been influential studies that reviewed and studied cold start problems in recommender systems, these studies have a limited scope in explaining the state of recommending items to new user in the recommender system domains. For example, the study of [13] explains and reviews the literature on the use of social network data to mitigate cold start problem. The comparative study of [14] categories cold start literature based on the use of additional data sources; selection of the most prominent groups of analogous users; and enhancement of prediction using hybrid methods; the study of [15] reviews sparsity issues in recommender systems.

In addition, other key influential studies in recommender systems such as [16] have only focused on reviewing cross domain recommendation by identifying the most widely used cross-domain recommender systems building block definitions, algorithms adopted, challenges, common features between them. In addition, due to the significant attention and popularity of deep learning in several research fields for its effectiveness in information retrieval, recommender systems researchers and practitioners have progressively adopted deep learning approaches in recommending suitable items to users. For instance, the study of [17] reviews recent research efforts on deep learning-based recommender systems.

Despite complementary or auxiliary information being the integral information needed for a cold start recommendation, recommender system literature lacks a detailed explanation of the process of reaching and obtaining such auxiliary information typically by integrating or coupling various filtering strategies with supporting machine learning and other algorithmic approaches. This study is entirely different from previous related studies as it focuses on explaining and providing an insight into the detailed process of reaching and extracting the auxiliary information needed for cold start recommendation. This study also aims to explain the various approaches and techniques employed by researchers in offering recommendations to cold start users, using various algorithms and techniques; the various application domains for cold start recommendation and identifying areas for future research.

Arguably, one of the key functions and benefits of review studies is to highlight and enable researchers with a clear understanding of the research activities in a particular domain; and to guide them in pursuing relevant studies in a preferential manner. This reduces overlapping, distortion and makes research a lot easier. Therefore, this study contributes to the cold start recommender system as well as the recommender system as a whole in guiding researchers the way forward and to better understand the type and how cold start is mitigated. Other key contributions of this study are as follows:

- This study offers a novel explanation and insight into the detailed process of reaching and extracting the auxiliary information needed for cold start recommendation.
- This study categorizes and presents a taxonomy of the approaches used for eliciting auxiliary information needed for cold start recommendation.
- This study shows and identifies several challenges associated with eliciting auxiliary information for a cold start recommendation. One of the most noticeable challenges from the results is that deriving the auxiliary side information usually involves two distinct processes.
- This study illustrates the three (3) main processes or strategies used in obtaining auxiliary information by adapting traditional filtering and matrix factorization algorithms, typically with machine learning algorithms to build learning prediction models. The understanding of similar or connected user profiles can be used as auxiliary information for building cold start user profile to enable similar recommendation in social networks. Similar users are clustered into sub-groups so that a cold start user could be allocated and inferred to a sub-group having similar profiles for recommendations.

- This study presents the bigger picture of research productivity and outcomes in cold start recommender systems. This study identified the need of key areas that cold start recommendation research is yet weak especially with the use of deep learning approaches, so that researchers and recommender system practitioners could further understand the technicality of recommendation to a new user and serve as motivation in pursuing further investigations.

2. Methodology

2.1. Literature Search

The main goal of this study is to understand how researchers elicit and obtain the auxiliary information used for cold start recommendations. Due to the limitations associated with the traditional filtering strategies for example, researchers have adapted these approaches to attain auxiliary information needed for cold start recommendation. As such, we need to understand how these traditional filtering strategies such as collaborative filtering are modified, adapted to elicit cold start user preferences and recommendation in various application domains. We also need to understand other approaches, techniques or algorithms that researchers have used in mitigating cold start recommendation apart from the commonly used collaborative filtering techniques. We also need to understand the challenges associated with obtaining the auxiliary information for cold start. Finally, we also intend to identify the limitations of the cold start sub-domain in recommender systems and suggest future research direction.

In order to achieve these goals, we explore the existing literature from various electronic databases, we first identified the databases containing recommender systems literature which include (ACM Digital Library, Science Direct, Springerlink, IEEEExplore, Web of Science (WoS) and Google scholar portal) for extracting these studies. We formed a combination of keywords inspired by the studies of [13,18] using any of the combination of “cold start” and “recommendation”, “recommender systems”, and carefully keyed in into each of the databases to retrieve the relevant literature. The database selections and keyword formation are inspired by influential recommender systems studies such as (see [13,18]). We carefully checked and examined the title and abstract of each of the extracted studies to make sure that each study has a well-defined methodology and have proposed and evaluated an approach that mitigates user cold start problems in recommender systems. We have only considered peer reviewed indexed studies published in recognized publication outlets so as to ensure the quality of our studies, as well as ensuring that the recommendation process uses implicit data. Figure 1 below is the publication trend from the last seven (7) years from (January 2014 to December 2020). We decided on the last seven years inspired from similar recommender system studies of [13,18].

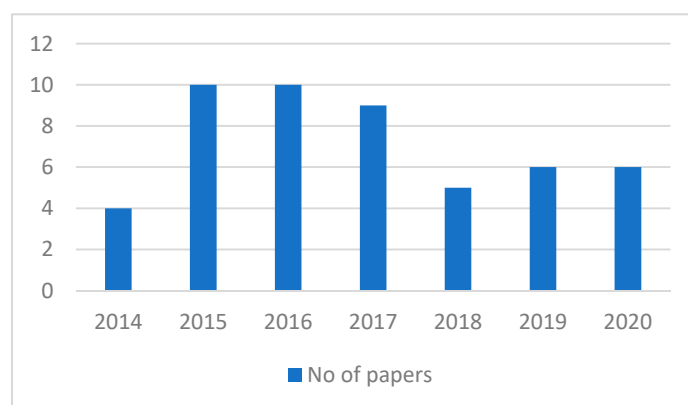


Figure 1. Publication trend.

However, this work is not limited to the 50 selected papers we have extracted. The literature extracted from the bases for building our foundation, arguments, interpretations, reasoning and conclusions. Therefore, we have cited numerous empirical and review papers in recommender systems domain as well as other domains based on relevance and applicability to provide basis for our findings, suggestion and concise explanations. We carefully read through all the selected papers to understand the approaches adopted in mitigating user cold start problem. We also looked closely and understand the type of implicit auxiliary data/information used and how this auxiliary data is derived and incorporated into the recommendation process. We also identified and understand the challenges associated with cold recommendation.

2.2. Categorization

The obtained studies/results were carefully grouped and categorized into categorical themes. We noticed that these studies can be categorized into four (4) main themes with each having sub-categories. The two popular mainstream approaches—model-based approaches and memory-based approaches—were also categorized based on studies that elicit auxiliary information by building a social network user profile; we also categorized other studies as “others” which used unfamiliar approaches.

3. Results

There are two mainstream approaches to recommender systems, namely model-based approach and memory-based approach.

3.1. Adapting Model Based Approaches

Model based recommendation involves building a model based on the dataset of ratings, in which information from dataset is extracted and used as a model for making recommendations. Model based recommendation approaches have been adapted for cold start recommendations through the use of machine learning and other non-machine learning algorithms as categorized below.

3.1.1. Adapting Traditional Filtering Strategies

Collaborative filtering is a well-developed framework that utilizes the history of different users to provide personalized recommendation. Over the years, collaborative filtering model has been extensively studied and revised, and researchers have proposed various enhancements and modification of the model to suit various recommendation needs in varying application domains [19,20]. The main drawback of collaborative filtering is the cold-start problem in providing accurate recommendation to new users because traditional filtering strategies/techniques—collaborative filtering and content-based filtering have limited scope in the cold start scenario due to lack of descriptive information about the new user/item. One of the ways to use these traditional filtering strategies for cold start scenario is to modify or adapt them to work on sparse data and pure cold start situations. This adaptation usually involves coupling them with algorithms and other strategies to work on cold start recommendation. Collaborative filtering algorithms are combined with classification algorithms such as C4.5 and naive Bayes algorithm, for finding a neighborhood for a new user [9]. The use of collaborative filtering for cold start recommendation has been widely adopted in recent decades, for example, the study of [21] built error reflected model and applying it to collaborative filtering recommendation for cold start recommendation using user neighborhood and user similarity matrix. The study of [22] proposed unique approach which is a weighted linear combination of simpler similarity measures that is optimized using neural network machine learning algorithm.

a. *Coupling collaborative filtering with machine learning algorithms:*

Mitigating the cold start problem ordinarily consider traditional factorization algorithms as the baseline model or the initial step in acquiring the additional auxiliary information to build the cold start recommendation. This is because traditional factorization algorithms usually capture the fine-grained link or relationships between users or items, then these links or models are coupled with various algorithms or approaches-machine learning algorithms being the most popular, to find the neighbor of the items, users or links in order to extract the implicit auxiliary information from the same network/domain or from an entirely different domain for cold start recommendation. From Table 1, coupling of collaborative filtering with other approaches to infer auxiliary or side information which is used for cold start user recommendation can be categorized into two main sub-categories.

Table 1. Adapting traditional filtering strategies.

S/N	Articles	Platform, Data and Domain	Summary of Approaches
<i>Coupling collaborative filtering with machine learning algorithms</i>			
1	[9]	Movielens (Movies)	Collaborative filtering algorithms combined with naive bayes and C4.5 classification algorithms
2	[23]	Movielens (Movies)	Model based collaborative filtering by a context aware to elicit auxiliary data
3	[24]	Netflix, jetstar, Movielens (Movies)	User demographic information together with collaborative filtering to address cold start problem using a SCOAL (Simultaneous co-clustering and learning) algorithm
4	[25]	Not stated	Adapting collaborative filtering using explicit trusted user neighbor data
5	[26]	Netflix (Movies)	Collaborative filtering merged with machine learning algorithms-deep learning algorithm (deep neural network) SADE, timeSVD++
6	[27]	Opinions	Collaborative filtering merged with K-nearest neighbors
7	[28]		Collaborative filtering with word2vec unsupervised machine learning algorithm
<i>Coupling collaborative filtering with other approaches</i>			
1	[29]	Tancent Weibo (Video)	Collaborative filtering merged with a hybrid random walk (HRW) method
2	[30]	Tancent Weibo (Video)	Collaborative filtering merged with a social group-based algorithm
3	[31]	Ontological approach	Collaborative filtering adapted with an ontological approach
4	[32]	WSDream (Web services)	A collaborative filtering web service QoS prediction approach that utilizes geographical neighborhoods knowledge

One of the most effective approaches for obtaining and eliciting auxiliary information in recommender systems for cold start recommendation is to couple collaborative filtering with machine learning algorithms. Researchers have used various supervised, semi-supervised and unsupervised machine learning approaches coupled together with traditional collaborative filtering to obtain relatively reliable new user inference or preferences which is used for the cold start recommendation. From Table 1, the study of [9] involves modelling collaborative filtering algorithms with C4.5 and Naïve Bayes classification algorithms to find a neighborhood for a new user. First, the system trains the classification model with the new users' demographic information, which then classifies the new users to demographics-based groups. The process involves adopting a mechanism that considers users' demographic data, based on the similarity techniques, the system then finds the users' auxiliary information for identifying their neighbors preferences which leads to cold start recommendation. Neighborhood recommendation was built based on the idea that users having similar characteristics and common backgrounds are more likely to have similar preferences. Hence, each cold start user is categorized to a group with members having similar characteristics, and accordingly the

ratings from a cold start user group category is used for inferring the new users cold start recommendation. One of the main strengths of the study in relation to handling cold start situation is that it mitigates the new user cold start problem effectively. The study also involves semantic similarity metrics in the calculation process and does not depend on any complex calculations. Similarly, the study of [24] leverage the use of user demographic information together with collaborative filtering to curl new user preferences implicitly so as to provide suitable recommendation using an algorithm called SCOAL (simultaneous co-clustering and learning). First, users with common interest are identified and grouped into sub-groups, then prediction models are built simultaneously specifically for each sub-group. The second step involves identifying which sub-group a new cold start user fits or should be allocated to base on the prediction models generated by the SCOAL. In this case, the new user is offered recommendation based on the data of the new user allocated group.

Researchers in recommender system domain have also devised the use of multiple models approach to extract implicit auxiliary information for solving the problem of cold start recommendation, since it is challenging to find useful information from a single source. Integration is usually performed by allowing the models to train each other in a “semi-supervised manner”, using an implemented “co-training algorithm”. The study of [23] adopted a state-of-the-art model-based collaborative filtering algorithm by a context-aware approach to incorporate auxiliary information about both the users and items in the recommendation system. Based on the context modelling data, multiple supervised learning models are constructed to complement and train each other for offering a new user suitable recommendation in recommender system. These models train each other and new user information could be generated from the model training and used for cold start recommendation. This significantly alleviates the cold start problem.

Over the past two decades, deep learning has received huge attention in various research domains [33–35]. Deep learning approaches have been applied in various capacities due to its capability in solving many complex tasks while providing start-of-the-art results. The two arguably most popular deep learning architectures are convolutional neural network CNN, which is a technique commonly used for image processing applications, and recurrent neural network RNN are mostly used in recommender systems, text and speech applications are the most popular deep learning methods currently in used. In [26], collaborative filtering is merged with deep learning algorithm (deep neural network) in forming a cold start recommendation. Deep learning machine learning approach are used for obtaining content features from content descriptions. Here, the content features are used as key components of the recommendation models which implicates the prediction of the unknown rating and the training models for cold start recommendation.

The use of traditional collaborative filtering with trusted neighbors have been effective in resolving data sparsity and cold start problems which the traditional collaborative filtering suffer [25]. Specifically, from the study of [25], the implicit data for cold start recommendation was generated through merging trusted neighbors to complement and represent active users preferences based on the similar users that can be identified. The approach involves the use of explicit user’s social trust information to specify other users as trusted neighbors. The approach involves merging trusted neighbors’ ratings of an active user according to the extent to which the trusted neighbors are similar to the active user. The set of merged ratings is then used to represent the active user’s preferences as auxiliary information used to adapt the traditional collaborative filtering for cold start recommendation. One of the main strengths of the study was effectively complementing user rating profiles based on the ratings of trusted neighbors—an effective trust-aware recommender system. Although the recommendation generated is reliable, it only covers the smallest portion of items, since only the ratings of the users who have a large number of trusted neighbors and high rating correlations are possible

to be predicted. Therefore, for better recommendation, users have to provide explicit trust information for the merging process which users may not be willing to reveal.

Adapting traditional filtering strategies can also be applicable for cold start item recommendation using unsupervised machine learning approaches such as the study of [28] that proposed a content embedding based hybrid recommender model which involve the combination and adaptation of traditional collaborative filtering together with machine learning algorithms for item cold start recommendation. The study proposed HRS-CE framework using Word2Vec for obtaining description of items using content embedding methods. Word2vec is an unsupervised methodology for building word embeddings. It uses a neural network model to learn word associations from a large corpus of text. The word2vec algorithm once trained can detect synonymous words or suggest additional words for a partial sentence. These suggestions and predictions are used as the auxiliary information to offer recommendation to cold start users. This approach utilizes item-detailed descriptions, as opposed to using meta-data such as tags, keywords; this helps in capturing the deep semantics of item descriptions result in richer information about cold start users and consequently more accurate predictions.

b. Coupling Collaborative filtering with other algorithms:

Obtaining and eliciting auxiliary information used for cold start user does not only involve coupling collaborative with machine learning algorithms. Researchers have adopted and developed various other algorithms to obtain additional user inferences and ratings for user cold start recommendation. The study of [29] adapted collaborative filtering with random walk-based algorithm modelled as hybrid-walk recommendation approach to infer auxiliary information for cold start recommendation from a different domain—the cross domain. A cross domain recommendation typically involves the transfer of knowledge from a source domain to a target domain (transfer learning to bridge two domains by the set of user nodes). Thus, [29] leverage the use of social ties as the fundamental bridge to connect item domains in social networks that determine behavioral consistency and popularity factors for the cold start recommendation as the implicit link. The study successfully transferred useful knowledge from a social network domain to a target domain which offers recommendation for both data sparsity and pure cold start recommendation. Because it involves more data from heterogenous domain that provide more auxiliary information for cold start recommendation, it provides better and reliable cold start recommendation as it also bridges cross-domain knowledge through social information. One of the main issues with the random walk recommendation algorithm is that it must be tuned considering information about the nature of the application.

The study of [31] adapted collaborative filtering using an ontological approach for cold start recommendation. Ontology seeks the classification and explanation of entities. Ontology looks at the things the data are about and uses them as the basis for the structure of the data. This is achieved through the identification of key users within a social network. The implicit auxiliary information involves structuring data of key users of a network in an ontological way. The approach involves identification of users' best entry points into the network to explore its shared resource content. The study is based on an importance measure which combines several aspects characterizing the individuals within the social network. One of the shortcomings of this study is that it was not experimented in real world social network such as Facebook or Twitter.

The study of [30] infer recommendation to cold start users by coupling social group-based algorithm with traditional collaborative filtering to produce personalized video recommendations in cold start. The approach is similar to [9,24] in which users in a social network are grouped into friends based on their similarities, then a new user is inferred or offered recommendation based on the group that the cold start user is best affiliated to. The proposed algorithm was found to not only improves the click-through rate, but more diverse videos are recommended. However, one of the main drawbacks associated with

the social group-based algorithm is that the collaborative filtering is not outperformed by the social group-based algorithm in all cases. Generation of results is difficult for highly active users by the social group-based algorithm since many videos in the group video candidate pool have been viewed previously by the active users. In a pure cold start problem, the model leverage factor models in order to boost performance. Thus, it cannot perform well in warm start but only in pure cold start.

The study of [32] proposed a collaborative filtering Web service QoS prediction approach which utilizes the knowledge of geographical neighborhoods for effectively mitigating data sparsity and cold-start user recommendation. The use of geographical information as complementary information have led to the approach achieving complete prediction accuracy. On the other hand, the approach is inapplicable on different scenarios unless the method is extended. In addition, the prediction accuracy needs significant improvement by considering the influences of contextual information.

3.1.2. Adapting Matrix Factorization Model

The use of matrix factorization has been one of the most common approaches for providing both user and item recommendations in recommender systems as well as mitigating cold start issues recommendation in cold start situations. Matrix factorization is a family or class of frequently used techniques in collaborative filtering. Users and items characteristics are transformed by the matrix factorization into latent factor space and predicts the rating of users concerning items by computing the similarity between the user's interest and the target item [36,37].

The adaptation of matrix factorization collaborative filtering technique has been one of the main approaches used in addressing cold start problems in various domains. Researchers have adapted or modified the matrix factorization to suit various application domains and contexts.

Similarly, researchers have used matrix factorization in combination with other approaches or machine learning techniques/algorithms for recommending items to new users. The use of matrix factorization has been a constant theme across for cold start recommendation especially in e-commerce and entertainment (movies and music) domains. Researchers have also provided recommendations for cold start users by leveraging cross domain data. Cross-domain recommendation refers to the use of representations of users and items in one domain to make recommendations in another domain [16]. These representations are considered the auxiliary information for the cold start recommendation. Adapting matrix factorization approaches can be used for cross-domain recommendation in such a way that aggregated information of user preferences from both domains are combined, however matrix factorization approaches have mostly been used in knowledge transfer and linkage approaches [38,39] for cross domain recommendation by leveraging the relatively richer information such as ratings from the source domain for improving cold start recommendation accuracy in the target domain. Therefore, cold start recommendation accuracy highly depends on the mapping accuracy of the latent factors across the source and target domains.

a. *Coupling Matrix Factorization with machine learning algorithms:*

From Table 2, [40] tackled cold start recommendation through coupling matrix factorization with linear regression and regression tree supervised machine learning algorithms to input the missing values into the rating matrix. The approach combines attribute selection, local learning and value aggregation into a single approach to solve the problem of cold start recommendation. Attribute selection involves selecting a small relevant portion of a data that can be sufficient for predicting the larger data. This reduces complexity by removing irrelevant and redundant attributes from a training data. Local learning involves models built only with the neighboring training data. The attribute value data is used as the auxiliary information to transform pure cold start into warm

start, then auxiliary information is obtained using the Euclidean distance formula to complete the missing values of the matrix from warm start to recommendation.

Table 2. Adapting matrix factorization model.

S/N	Articles	Platform, Data and Domain	Summary of Approaches
<i>Coupling matrix factorization with machine learning algorithms</i>			
1	[40]	Not stated	Matrix factorization coupled with regression tree and linear regression supervised learning models
2	[41]	Movielens	Matrix factorization with Markov chain (Markov random field prior (mrf-MF))
3	[42]	Social media	Matrix factorization with Markov chain
4	[43].	Music	Merging Markov chain with matrix factorization model
5	[44]	Siena weibo (e-commerce)	Cross site (<u>cross domain</u>). Matrix factorization with modified gradient boosting trees and recurrent neural networks
6	[45]	Not stated	Cluster-based matrix factorization model with K means clustering algorithm for cold start cross domain recommendation
7	[46]	Videos & DVDs, Books, and Music	Cross-domain recommendation mechanism based partial least squares regression analysis. PLSR coupled with matrix factorization
8	[47]	Not stated	Cross-domain and cross-system recommendations coupling deep neural network with matrix factorization models
<i>Coupling matrix factorization with other approaches</i>			
1	[12]	DBpedia	Adapting matrix factorization with open linked data (MF-LOD)
2	[48]	Not Stated	algorithmic framework based on matrix factorization
3	[49]	Social media (Twitter)	ISoNTRE—is a framework called intelligent Social Network Transformer into Recommendation Engine is combined with matrix factorization
4	[50]	Douban movie iMDB (Movies)	CHRS uses a two-step approach to integrate the complementary in both the target and source networks for cold start recommendation utilizing the rich information from these two paths
5	[51]	Social media platforms (e-commerce)	Cross-domain cold start recommendation based on mapping user feature based matrix factorization from social media and product features from e-commerce website
6	[52]	Yelp; movielens (Movies)	Developed KAMF that use social links between users for cold start recommendation
7	[53].	Movielens (Movies)	Matrix factorization with a content
8	[54]	Tancient Weibo (Movie)	Neighborhood and matrix factorization (hybrid—memory based and model based)
9	[55]	Tencent Weibo (microblogs)	Adapted matrix factorization adapted by coupling social behavior and social trust data

Markov's chain has been used coupled with matrix factorization to infer and accurately predict new users' preferences. The study of [41] enhanced matrix factorization prediction capability by integrating an n-dimensional Markov random field prior (mrf-MF) to cope with three types of cold-start problem: recommending new users to existing items; recommending new items to existing users; recommend new items to new users. First, a specific neighbors system for user attributes such as age, occupation of users and genre, release year of items is defined followed by conditional distribution of latent profiles. Based on the k-nearest neighbors system, the conditional distribution of latent profiles of users/items in n-dimensional Markov random field is defined. The proposed model could not deal with sparse data or warm cold scenario where very sparse training data is available. Similarly, the study of [43] mitigate the cold start problem for next-song recommendation by merging Markov chain with matrix factorization model. The approach captures the content-based transition preference by mining both sequential behavior and content feature simultaneously. These extracted data is used as the auxiliary

information for cold start recommendation. Due to the difficulty of obtaining the correct side information for cold start recommendation, recommender systems researchers have used arguably a hybrid of machine learning approaches by combining unsupervised and semi-supervised machine learning algorithms in different steps to arrive at accurate recommendation for user cold start. The study of [42] adopted the FPMC framework which combines matrix factorization and Markov's chain. The approach involves the use of node2vec semi-supervised algorithm to pre-train the representation of users. Then inferences of network neighbors with high similarities are considered as the auxiliary information which are selected and integrated into the user cold start recommendation. The recommendation of this approach is relatively weak in dealing with cold start users having fewer social links.

The study of [47] proposed a deep framework for both cross-domain and cross-system recommendations which is based on coupling deep neural network with matrix factorization models. The process of obtaining and transferring knowledge from the source domain to the target domain typically in a cross-domain recommendation involves three processes. First the user and item latent factor matrices is obtained using matrix factorization. The second phase, the benchmark factor matrices are generated by combining the latent factor matrices according to the sparsity degrees of individual users and items, then the latent factor matrices are mapped to fit the benchmark factor matrices. Based on the affine factor matrices learned from the second prediction phase, the users' ratings on all items in the target domain or system and recommend which matched items to target users is predicted. Similarly, The study of [45] used a modified or adapted matrix factorization-clustering based matrix factorization model to construct a rating matrix that includes available ratings on both of the domains for a cross domain recommendation, after which a k-means algorithm is used to categorize users and items after mapping the matrix into a lower-dimensional latent space. The matrix factorization models generate user and item latent factors and then employs the deep neural network to map the latent factors across domains or systems for cold start recommendation. [44] used linked users between two sites (e-commerce sites and social networking sites) who have made purchases on e-commerce websites and has social networking sites to map and elicit user preferences for cold start recommendation. The approach involves two distinct processes (two-step process). The first step involves learning both product and user embeddings from e-commerce websites data using machine learning algorithms-gradient boosting trees (supervised) and a recurrent neural network (supervised) for transforming users' social networking features into user embeddings. The second phase involve adapting matrix factorization to a feature-based matrix factorization which these user embeddings for cold start product recommendation.

Typically, matrix factorization is adapted or coupled with additional algorithm to predict cold start user preferences, however, there are instances in which a particular technique or algorithm is adapted or extended so as to be able to couple or integrate with matrix factorization. In this case, matrix factorization is the secondary method or mechanism to trigger cold start recommendation. A typical example can be found from the study of [46] that adapted partial least squares regression (PLSR) machine learning technique to enable coupling with matrix factorization for cross domain cold start recommendation. Although, PLSR can be applied to matrices without missing values, the rationale of adapting PLSR with matrix factorization is to train the data that reflect user preferences based on records. As such, the two proposed recommendation models from the study-PLSR-Latent and PLSR-CrossRec were able to purely use source-domain ratings in predicting pure cold start users' ratings for users that have never rated any item in the target domains. The basic challenge associated with PLSR-based methods is that it first needs to complete the rating matrices before the recommendation process commences, and as such, may not lead to satisfying performance.

b. Coupling Matrix factorization with non-machine learning algorithms:

Basically, researchers leverage the handiness and utility of matrix factorization coupled with machine learning algorithm or other strategies or techniques such as cross domain recommendation for tackling cold start issues in various domains for user cold start recommendation. From Table 2, it can be seen that researchers have used matrix factorization heterogeneously and diversely coupled with diverse set of algorithms or modified in various ways to cope and suit their cold start recommendation pursuits. The study of [49] proposed a framework called iSoNTRE (the intelligent Social Network Transformer into Recommendation Engine) and combines the framework together with matrix factorization to offer cold start recommendation by using users' information on online social networks. The framework begins by transforming data on social network into useful information, and then the recommendation core—matrix factorization uses this information as auxiliary information to offer recommendation to cold start users. The approach sounds promising in mitigating both cold start users and item problems. However, the data evaluation is shallow using relatively undersized social network data. In addition, the data was only tested on a single social network stream (Twitter). [48] introduced a novel matrix factorization model in which the cold start users and items form the rating matrix, then are excluded and recover the obtained sub-matrix in a perfect way. Then, the recovered sub-matrix along with available similarity side information about users and items are utilized to transduct the knowledge to cold-start users/items for user cold start recommendation. Transduction is a word commonly used in biological sciences to explain the transfer of a genetic material within similar organisms but via a different type of specie. One of the main problems of this approach is that users can only be offered effective recommendation only when there is sufficient side or complementary information. Its optimization is only in pure cold start situation. The study of [53] presents a novel method for learning latent factor representation for videos based on modelling the emotional connection between user and videos. Emotion modelling deals with estimating the likelihood of emotional response that an item would generate by its users. This estimated response data is the auxiliary information used for cold start user representation. The approach ensures cold start video recommendation with no user prior collaborative information. On the other hand, the visualCLiMF is not particularly effective in a pure cold start situation for recommendation to users. The study of [54] incorporates content-based information and social information to collaborative filtering to build model (i.e., combination of memory based and model based approaches) to build content association between items in selected tag-keywords and tags using user item ratings. The information generated from relational user interest by tag keyword relation matrix is considered as the auxiliary information for cold start user recommendation.

The study of [52] used a two-step approach. The first step involves integrating items attribute information into an improved and revised matrix factorization called kernel-based attribute-aware matrix factorization model KAMF. KAMF typically exploits the rich attributes of items and users using nonlinear interactions among attributes. This rich attributes of items and users is the auxiliary information but, in this case, it only provides the baseline for the recommendation process for the second step. The second step involves evaluating the model parameters of KAMF using an extension of KAMF incremental algorithm to which is further adapted for addressing the user cold-start problem using social links data between users. Although KAMF can deal with cold-start problem by nature, however KAMF's main weakness is that it focuses on predicting item rating first, then using those rating to recommend cold start users to these items. The study of [55] used social trust and social behavior data in microblogs to adapt matrix factorization using an algorithm called trust and behavior-based singular value decomposition (TBSVD) for eliciting user preferences. First, implicit trust is calculated based on user interaction behavior including comment, mention and retweet, while explicit trust is based on the direct connections between users. This implicit and explicit trust information are used as auxiliary information needed for cold start recommendation. Then, an extended trust matrix is constructed combining both implicit trust and explicit trust which

is then used to build the model using matrix factorization technique. The technique was only evaluated in microblogs.

Leveraging linked open data has been a common approach for building recommendation capabilities in recommender systems due to the recent advancement in semantic web technology which has enabled data to be represented strategically for improving recommender systems. This advancement enables systems to significantly and better understand user and item preferences and features based on their contexts and domains. The increase in publication of linked open data has enabled to link user and item entities from varying knowledge sources by connecting different homogenous data or information in a single global data space. Over the years, linked open data has been used for both cold start [56] and non-cold start [57,58] recommendation. The study of [12] from Table 2 leverage implicit feedback data and linked open data based on similarity measure for user cold start recommendation. Matrix factorization is coupled with linked open data (MF-LOD) considered as the auxiliary information to enhance the matrix factorization model called singular value decomposition (SVD++) for cold start recommendation.

The study of [51] proposed microblogging reviews based on cross lingual sentimental classification model for mitigating cross-site cold start product recommendation. An adapted matrix factorization called feature-based matrix factorization method is used in the mapping of user features from social media and product features from e-commerce website. Then sentimental analysis is applied on two languages reviews (English-Hindi) as well as the extracted texts from audio-video reviews. One of the main drawbacks of the study is that the approach is not fully evaluated. More experiments needed to validate the approach.

However, researchers have also provided cross recommendations from within the same domain, in which two different social network-source and target networks are mapped to offer recommendations. The study of [50] adopts a across heterogeneous network approach (CHRS) which uses a two-step approach to integrate the auxiliary information in both of the source and target networks from within the same network for user cold start recommendation. The approach utilizes matrix factorization to incorporate item similarity information in the source network. This approach is somewhat similar to cross domain approaches, but happens within the same network. The approach leverages the rich information from meta-paths for calculating the movie similarities from multiple types of relation information. Using a unidirectional way, the movie similarities from these rich meta-paths is the auxiliary information. However, the study of [59] is similar to [50] in the sense that [59] employs the use of inter domain cross recommendation within the same music domain recommendation.

In addition, we have also identified a number of studies that adapted matrix factorization model for cross domain item recommendation such as the study of [38] that leverage the use of matrix factorization models for cross-domain collaborative filtering in bridging between items liked by users in different domains. The study developed a hybrid factorization matrix factorization model that jointly exploit user preferences and item metadata for cross-domain recommendation.

In summary, researchers have adapted and leverage matrix factorization to mitigate cold start recommendation through adapting and coupling with various machine learning and techniques. Although, the study of [60] used a collective matrix factorization method using tag information to solve the sparsity problem. Results shows that the approach generates more precise prediction than general collaborative filtering suffering from the cold start problem. However, it is primarily focused on sparse data and not pure cold start user recommendation.

3.2. Adapting Memory Based Approaches

Memory-based algorithms also referred to as neighbourhood-based collaborative filtering algorithms were among the oldest approaches developed for collaborative filtering [61]. The neighbourhood-based approaches are generally based on the k-nearest

neighbour rule (KNN), and provide recommendations by aggregating the opinions of a user's k-nearest neighbours [62]. Two stages are involved in neighbourhood-based methods: the neighbour selection and the rating prediction. In cold start scenario, researchers have modified and adapted k-nearest neighbour by coupling them with various algorithms to obtain auxiliary information needed for cold start recommendation. Similarly, researchers have leveraged the use of tags to establish neighbourhood link for cold start recommendation.

a. Tags-based elicitation approach:

Tags are usually associated with supplementing and offering valuable information that are used in recommendation. This is tags summarize item properties as well as describing user preferences through tagging behaviors.

From Table 3, researchers have leveraged the use of tags in recommender systems in identifying neighborhoods of a cold start user. Tags carry vital information and defines the behavior or preferences of the tag recommendation or user. Although the use of tags has been popular in recommender systems, tags have also been used in eliciting implicit information for user cold start recommendation.

Table 3. Adapting memory-based approaches.

S/N	Articles	Platform, Data and Domain	Summary of Approaches
1	[63]	Facebook (movies)	Proposed a similarity metric for grouping users of the same interest for improving cold start movie recommendation through the use of KNN and K Means algorithms
2	[64]	Movielens, Elo7 (Movies)	Tag recommendation methods in a cold start scenario through syntactic and neighbourhood-based attributes.
3	[65]	Youtube (videos)	Video recommendation based on visual tags; visual description of the videos is used for the automatic annotation of the tags using K-Nearest Neighbor algorithm.
4	[66]	Mobile app	Predicting the next app to be used which involve a technique based on a set of features representing the real-time spatiotemporal contexts sensed by the home screen app for user cold start
5	[67]	Mobile app	Provides cross domain cold start recommendation for mobile apps from a mobile application domain to a news domain.
6	[68]	Mobile app	A modified incremental k-nearest neighbors (IkNN) and K-clustering algorithm for predicting next app for cold start
7	[69]	Mobile app	Location features and user interests are transferred from app usage data for cold-start location recommendation,

For example, the study of [64] proposed a syntactic and neighborhood-based attributes by using syntactic patterns of the text associated with web objects which can be exploited to identify and recommend tags to new users in a cold start scenario. However, this approach was not tested in pure cold start recommendation. Similarly, the study of [65] used k-nearest neighbor algorithm for cold start recommendation. The process involves eliciting implicit data with the help of tags for user cold start recommendation using visual tags that are automatically annotated to videos based on visual description of the videos. Such features are used as the auxiliary information, automatically extracted and added to the video items cold start recommendation. Research in recommender system has shown the reliability and effectiveness of building models capable of learning the correlation between video tags and visual features. This approach could be used in a pure cold start item-a new video item with no tag information, then personalized user recommendation for users is generated by exploiting the visual tags. Thus, this promising technique is capable of generating suitable recommendation in both warm and pure cold start recommendation. This has been a promising approach for mitigating cold start in video. Although, this technique was only tested in video recommendation, it would be

interesting to explore testing this approach in other similar domains such as music. The study of [63] used k-nearest neighbor to establish a similarity metric for organizing users in the same interest group for movie recommendation. This has been a recurring technique or strategy across cold start recommendation in which users are grouped based on their similar characteristics such that a cold start user is affiliated to a subgroup and offered recommendation based on the features of the sub-group. The data or information from a subgroup is used as auxiliary information for offering recommendation to the cold start user.

b. Cold start recommendation in mobile environment:

Over the years, there has been a steady increase of research activities in mobile environment. Although, recommendation in mobile environment have largely been researched and recommendations are proposed by the use of contexts information (see [70–75]), mobile devices and their associated apps require contextual user information to suggest the correct apps based on their location and contexts. Cold start recommendation in mobile environment can arguably be associated with various and diverse set of possibilities for exploiting and obtaining auxiliary information. Due to users' constant dynamics of changing personal, social and location context these instances could be leveraged to determine a more accurate and personalized recommendation to a user bases on users' location, activity, past records in a location, or even emotional state. For instance, a recent study of [76] integrates cloud computing and mobile technology for providing smart market recommendations for context-aware mobile recommendations. This is usually done by cross platform recommendation or culling information from other social network platforms, especially Twitter. Our study has shown that the use of cross domain data transfer and the use of app data for cross platform mobile apps recommendations have been a rich approach for tackling cold start recommendation in mobile environments. From Table 3, it is obvious that research has also employed other techniques of accessing side information to provide cold start users with relevant apps such as the use of privacy data or preferences. There are two types of user cold starts in mobile app which are app cold start that happens as soon as a cold start user install a new app on a mobile device, while the other is user cold start happens when a user installs and opens a cold start (new) on the screen. User cold start is critical because if it is not addressed, it leads to the risk of app abandonment which necessitates offering personalized recommendation from users' interactions with the home screen apps.

The advancement of technology and mobile technological devices have made it possible for recommendations to be done based on geographical location, context on mobile devices and mobile applications. Recommending a particular app on mobile devices is becoming increasingly challenging due to the increasing number of apps installed. Consequently, it is important to be able to quickly and accurately predict the next app to be used. Users arguably have difficulty in identifying apps that are relevant to their interests due to the large number of mobile applications (apps) readily at their disposal. Recommender systems that use previous ratings can address such problem (i.e., collaborative filtering) once apps have sufficient ratings from past users. However, newly apps that are added to repositories or newly released apps face the problem of cold start, due to collaborative filtering not having any user ratings on those apps to infer recommendations. Research on app usage data has been relatively few which throws the doubt on its accessibility and seems to limit the utility. However, it is encouraging that the situation is now changing.

Although, there have been huge progress in research about app recommendation in mobile environment focusing on different dimensions and aspects of users' mobile application. For example, the study of [77] that leverage the unique properties in the app domain and explored the effectiveness of using version features in app recommendation. Another example is the study of [78] that users' privacy preferences and interest-functionality interactions to perform personalized app recommendations. The study of

[79] that leverage users' and apps' data on multiple platforms to enhance the recommendation accuracy, which address the problem of cross-platform app recommendation. [80] that used social network information from Twitter by culling user preferences from social network planforms (Twitter). Very few studies have actually focused on users' cold start app recommendation in mobile environment. The few studies that precisely focused on user cold start app recommendation use the neighborhood approach by typically adapting the k-nearest neighbors approach. From Table 3, the study of [68] proposed an approach for predicting the next app that a new user is going to use by using incremental k-nearest neighbor algorithm (an adapted K nearest neighbor) together with a traditional collaborative filtering to form a recommender system called Predictor. Predictor is an efficient dynamic collaborative filtering fusion algorithm that provides app cold start prediction. First, a prediction model is formed and trained as a basic k-nearest neighbor model to elicit the implicit data for next app recommendation for cold start user.

Secondly, the study of [67] proposed a solution called neighborhood-based transfer learning for cross domain cold start recommendation for mobile apps from a mobile application domain to a news domain. Cross domain typically uses rich information from a side domain to mitigate cold start. This recommendation is built on the notion of users with similar app-installation behaviors are likely to have similar tastes in news articles, and then transfer the knowledge of neighborhood of the cold start users from a mobile app domain to a news app domain. As such, this information is used as the auxiliary information for cold start app recommendation. Although this approach suffers from an issue in the sense that cross domain recommendation is not sufficient to provide cold pure start recommendation. Data from the app domain to the news domain is not fully sufficient for pure cold start recommendation. Thirdly, the study of [69] proposed a transfer learning based generative model for personalized location recommendation problem, which transfers user interest and location features (as complementary information) from app usage data to help cold-start location recommendation. This technique is a typical cross domain recommendation approach. Results of the study shows the significance of using app usage information for location-based recommendation, and as such mitigate user cold start recommendation. The study of [66] proposed an approach for predicting the next app that a new user is going to use. The technique adopted is based on a set of features representing the real-time spatiotemporal contexts sensed by the home screen app for cold start. The next app prediction uses the spatiotemporal contexts of real date and time information by the home screen apps as the auxiliary information for next app cold start recommendation. One of the main weaknesses of this study lies in improving the app usage prediction accuracy because prediction accuracy to decline over time, this is because some algorithms such as k-nearest neighbors does not consider the increasing amount of training data available over time. The other is that the remodeling time is increased because some algorithms consider the aggregation of training data over time, and they rebuild their recommendation models using all historical data once the amount of new data has reached a certain limit.

3.3. Building Social Network Profile for Cold Start Users

Users' relationships in social networks have been exploited to attain auxiliary information for recommendation process. Studies such as [81,82] have leveraged social network and rating preferences for deriving and eliciting the social relations auxiliary information. However, it is often challenging to accurately exploit the social relations and correctly determine the user preference from both social and rating information for cold start recommendation.

For a given user, "connected users" are all those users with whom a user is connected to or shares a social relation to, while "similar users" are those users who have similar preferences with a particular user. Although, all connected users may not be similar users. In fact, research has shown that the intersection between similar users and connected

users is less than 10% [83]. In addition, not all similar users are connected [84]. For collaborative filtering and latent factor model based, there are different ways to utilize social data. Some solutions factorize both the social matrix and the ratings matrix to obtain user's features, while some of them build a rough user profile through analyzing similar implicit users profile data [83]. Researchers have also obtained users missing ratings by aggregating ratings of their similar users [83], thus, these aggregated ratings or inferences of similar users are used as the auxiliary information for cold start recommendation.

Although social recommender systems are well suited for the new user cold start problem, there are some noticeable challenges such as a user having large number of social relations which poses challenge in providing cold start recommendation. Secondly, casual relationships for example introduce noise which leads to misleading relations. Furthermore, very few social relations do not also help in providing such cold start recommendation and may cause failure in recommendation. A typical example is the study of that [85] from Table 4 in which social ties of users with their friends are classified. This was done by calculating the affinity strength between users and friends through adapting the coefficient of Jaccard in the social network topology. These data obtained is used as the auxiliary information to offer recommendation of interest to cold start users. One of the main issues with the approach weak personalization, since some users prefer items consumed by weak ties over those by strong ties, other users may prefer in the opposite or different way. A recent study of [86] used social media data to create user profile behavior classification, then used random forest and classification tree algorithm machine learning algorithms to provide recommendation. Although, the study is weak in establishment of relation between user profile and item features.

Table 4. Building social network profile for cold start user.

S/N	Articles	Platform, Data and Domain	Summary of Approaches
1	[27]	Opinions	Collaborative filtering merged with k-nearest neighbors
2	[85]	DBLP, Ciao, Douban, Opinions (Movies, books, music)	The coefficient of Jaccard is adapted as an intrinsic feature of the social network. A customized ranking model of items are formed for generating recommendation to cold start user
3	[86]	Twitter	Leveraging the use of social media (Twitter) data to create a behavioral profile and classify users based on their behavior. Decision tree classifier and random forest machine learning techniques
4	[87]	Social networks	Geographical distance and social network correlation are leveraged on location based social network LBSNs for building cold start user profile
5	[88]	Facebook (music, books)	Proposed three approaches to mitigating, which are: personality-based active learning, personality-based matrix factorization and personality-based cross-domain recommendation
6	[89]	Not stated	Social network data is used with collaborative filtering to exploit community preferences to address cold start recommendation

The use of implicit social network information has been the most prominent type of implicit data used in matrix factorization-based approaches and has been the dominant and one of the most common approaches adopted by researchers in providing recommendation to new users of a system. Another promising approach used by researchers for providing recommendations to cold start users involve building social network profile for cold start recommendation using social network information. Since there are various types of information available in social networks, this information can be used in different sort of ways to use them especially in tracing users unstated information or preferences. Therefore, leveraging social networks information for alleviating cold start recommendation have proven to be an effective approach due to user's social profile containing various information such as opinions, friends, likes and dislikes background information or demographics, etc.

Information and activities that users share on social networks can help define their preferences. Such information includes, for example, friends, favorite sport, favorite movie, educational background, etc., which can be used to directly or indirectly to understand users' preferences as such can be used as auxiliary information cold start recommendation. In a direct way, the system could understand the users' preferred items and therefore recommend similar items. Several studies have shown the richness of leveraging these indirect user data which can help draw reliable and correct inferences about user and consequently understanding users personality (see [88,90]). Although one of the major downsides of some of these studies such as [88] in Table 4 is that the data used are offline data from a social media platform such as Facebook. It is not the true representation of cold users, which makes it difficult for the system to infer the users' true preferences (see Table 4). The main strengths of this approach is that learned user profile is generic as such, can be used for recommendations across multiple domains [88,90]. On the other hand, the data are unary (like, dislike) which is a key drawback of the approach. [87] leverage the use of geographical distance and social network correlation is leveraged on LBSNs for user cold start recommendation. The approach use social network information on location based social network as auxiliary information for cold start recommendation. This was achieved by forming a geo-social correlation model to extract the social network information of users based on various correlation strength in different geo-social circles. This information is used in building social network user profile based on a cold start user geographical location and context.

Forming Users Social Circle

Eliciting auxiliary information has always been challenging. Researchers have explored various ways to understand a user for offering suitable and preferred recommendation. Friendship plays an important role in recommendation because users' preferences are normally influenced by their circle of friends. Groups of friends or connected friends tend to have similar preferences and therefore neighborhood could easily be formed for a user using social relations such as trust relations, follow relations and group membership. While traditional recommender system only utilizes the ratings matrix, social recommender system uses the knowledge of a user's social relations together with the user preference information by using a social relations matrix that captures the relationships between different users [83] to offer personalized cold start recommendation. A user's social relations information together with the ratings information by using a social relation is the key auxiliary information required for cold start recommendation.

Techniques typically used by traditional recommender systems such as collaborative filtering and neighborhood based are deficient in forming effective neighborhoods coverage of a cold start user due to the lack of prior ratings. Moreover, traditional recommender systems offer recommendations that are limited to those which are discovered by the neighbors, one of the advantages of user social circle formation is that relationships between user circles in different neighborhoods could be exploited and leveraged in improving the accuracy and richer information for cold start user so that the limitation of using only traditional filtering strategies are overcome [84].

From Table 4, the study of [89] used social network data with collaborative filtering to exploit community preferences derived from social network containing all users, which helps to establish a tentative users social circle and consequently address cold start recommendation. In addition, personalized recommendations are offered to active users based on the predicted ratings. Using the Katz similarity which is a collaborative filtering approach for measuring regular equivalence in social networks for selecting k-nearest neighbors of cold start users, the study of [27] was able to elicit and use the implicit trust relationships data based on the social connections between users as the implicit auxiliary data for cold start recommendation. One limitation of the study was that the evaluation is

only done using recommender accuracy, optimizing of non-accuracy measures has been closely tied to user satisfaction.

3.4. Others

This section explains other unfamiliar approaches and techniques used by researchers to offer or provide recommendation for cold start users. From Table 5, we have found different sorts of studies that employ varying approaches for mitigating user cold start recommendation. One of the eye-catching studies is the study of [91] that proposed a hybrid model by combining the strength of item response theory (IRT) models with machine learning (classification and regression trees) for predicting students skills or ability to address the cold start problem for learners in their learning systems in online learning environments. The study demonstrates the application of machine learning in combination with item response theory to alleviate the effect of cold start in adaptive learning environments. Although fully implicit data are not used (partial implicit data), we intend to articulate and highlight this study due to lack of cold start recommendation approaches in learning and educational environments.

Table 5. Others.

S/N	Articles	Platform, Data and Domain	Summary of Approaches
1	[30]	Tancient weibo (video)	Social-group-based algorithm for video recommendation
2	[59]	Music	Inter domain cross recommendation of using representation of music taste and past music listening for user cold start recommendation
3	[91]	ERS	Cold start recommendation in online learning adaptive systems using machine learning (classification and regression trees)
4	[92]	(Movies (yahoo movies))	Probabilistic approach incorporating deep neural network (DNN)
5	[93]	Not stated	Probabilistic approach
6	[94]	Facebook (CARS). Location based)	Rule mining and community-based knowledge for user context-aware recommender systems
7	[95]	Twitter	Conversational recommender system (critique-based recommendation)
8	[96]	Movielens (Movies)	hybrid model with machine learning algorithms to provide recommendation for crowd workers
9	[97]	Movie	Movie videos representation using gradient boosting tree

Another interesting approach involves the use of probabilistic approach (see [92,93]) coupled with deep learning approach-neural network. A probabilistic model based on uncertainty rules to allow new users to informally infer their own recommendations. Such uncertain rules involve for example “if a user likes item A, this user will probably like item B”. As such, cold start users are offered recommendation using inferred probabilistic information as the auxiliary information. One of the typical shortcomings with probabilistic recommendation is that the models built are not as accurate as the successful matrix factorization models when dealing with ordinary registered users, and as such, their cold start recommendation are not deemed accurate and reliable especially in pure cold start cases. A similar approach from the study of [94] alleviates cold start recommendation by hybridizing rule mining and community-based knowledge. For rule mining, association rules are frequently used in providing useful recommendations.

In this study, the top rules are ranked based on the highest score using the Apriori algorithm and FP Growth for cold start user recommendation. As such, the topmost popular rules are the auxiliary information used for cold start recommendation. It is clear that the recommendation does not involve integrating or the use of the total or complete implicit auxiliary information elicited, researchers usually rank the auxiliary data in accordance with their reliability in providing the accurate or ‘closer to the truth’ recommendation to cold start user. This is true from the study of [30] in Table 5 that used a social group-based algorithm for video recommendation based on social group

information. Candidate videos from a single group and multiple groups which a user is affiliated to could be ranked by the method. Thus, the social group information is used as an auxiliary information for recommending videos to new users. This was achieved through the use of video-ranking algorithms to rank the videos within a group and outside a group which a cold start user is affiliated to. Thus, the ranking information is used as the auxiliary information for cold start recommendations. The top-ranking data elicited are strongly considered and included in the recommendation process while the bottom or low-ranking data are usually ignored to avoid risk of flop or unreliable recommendations. Another interesting approach for tackling cold start user recommendation is the study of [95] presented an interesting conversational recommender system study that offers recommendation to users based on conversation (critique-based recommendation) to tackle cold start. The recommendation approach entails the rich and informative conversation nature between friends on social network based on critique. Ease of feedback and information content are the two spectrum ends.

3.5. Pictorial Summary of Findings

Figures 2–4 depicts the summary of the main findings of this study on how the auxiliary side information is elicited.

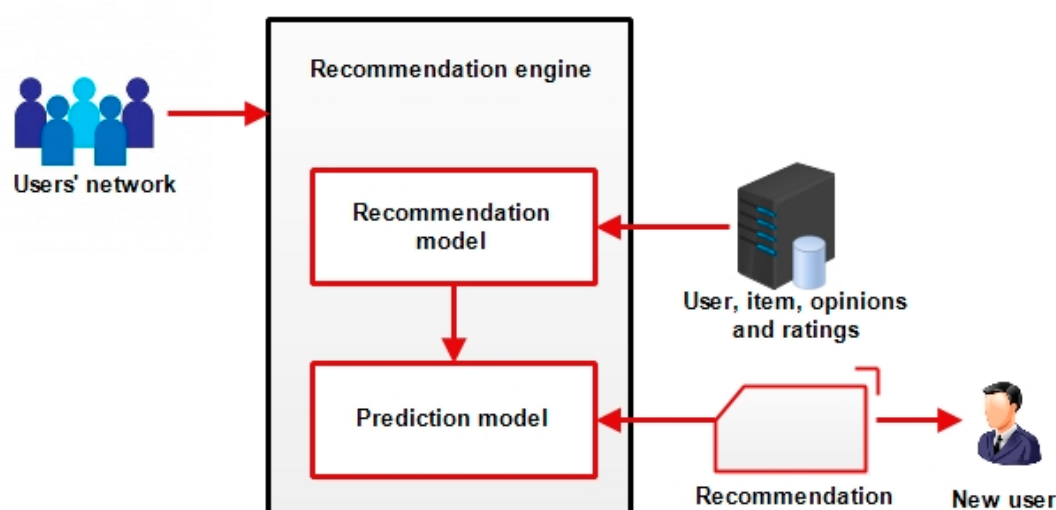


Figure 2. Adapting traditional filtering and matrix factorization algorithms typically with machine learning algorithms to build learning prediction models.

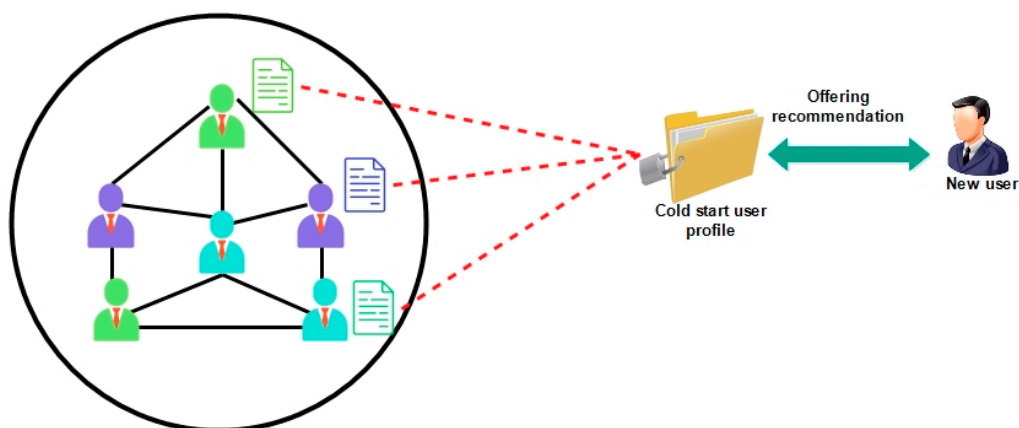


Figure 3. Understanding of similar or connected user profiles used as auxiliary information for building cold start user profile to enable similar recommendation.

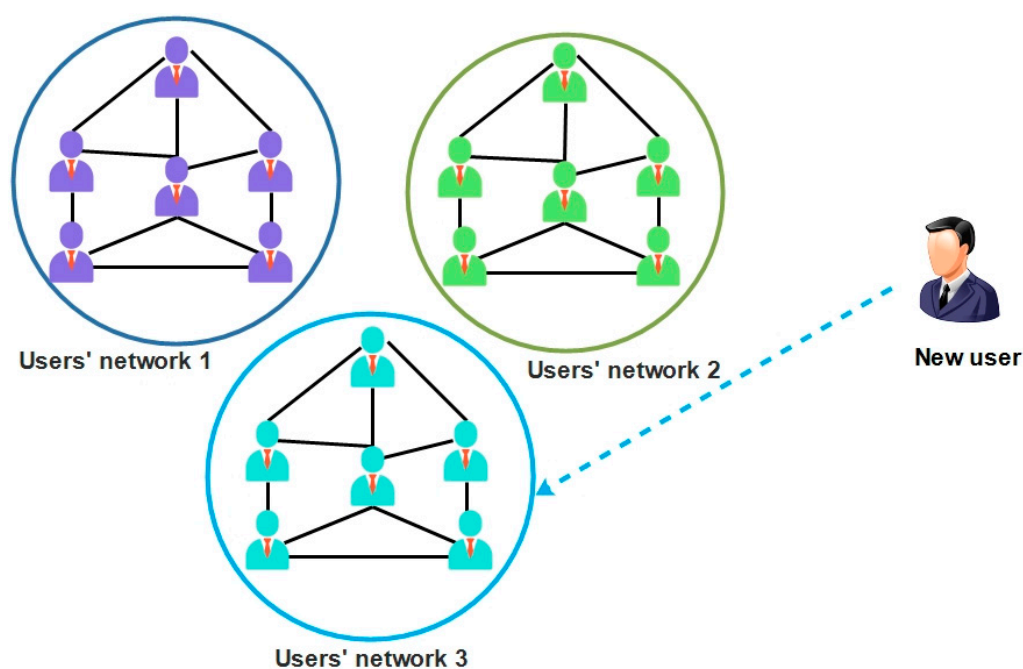


Figure 4. Similar users are clustered into sub-groups so that a cold start user could be allocated and inferred to a sub-group having similar profiles for recommendations.

4. Discussion

This study presents a survey explaining existing approaches and techniques used in mitigating a cold start user recommendation using implicit data depicted in the taxonomy (Figure 5). The use of implicit data or indirect method of understanding a new user in order to offer suitable and correct recommendations have always been the best and swiftest approach as opposed to eliciting user preferences explicitly. Furthermore, this study goes on to explain how researchers use these approaches to arrive at and elicit the auxiliary information necessary for cold start user recommendation. The implicit auxiliary information elicited for a cold start user recommendation include behavioral aspects of users such as the studies that used emotions, conversation (critique based) feedback; other user information includes social ties, likes, user reviews, ontology, user demographics, contextual data, tag information, DBpedia data, etc., in order to understand the users' preferences through the users' behavior, context and thereby offer users relevant and accurate recommendation.

As cold start recommendation has been notoriously challenging, since the system is deprived of the knowledge or preferences of a new user, researchers have used various clever techniques to attain additional information about the new user implicitly. These approaches involve improving filtering strategies such as collaborative filtering or matrix factorization; or collaborative filtering and matrix factorization coupled with machine learning algorithms (supervised, unsupervised) to form models for training and eventually attaining and eliciting the additional necessary information. Our study has shown various supervised, unsupervised and semi supervised machine learning algorithms coupled with collaborative filtering or matrix factorization used to form models to cold start recommendation.

Our study has shown and identified several challenges associated with eliciting auxiliary information for a cold start recommendation. One of the most noticeable challenges from the results is that deriving the auxiliary side information usually involves two distinct processes. This has been a common theme in adapting collaborative filtering and matrix factorization with machine learning algorithms, because eliciting auxiliary information involves forming and training the models to obtain the auxiliary information,

then incorporating the obtained information into recommendation processes. Leveraging cross domain auxiliary information has always proven to be an effective approach in obtaining accurate and spot-on auxiliary information, however, the knowledge transfer process comes at a cost, such that the auxiliary information from a target domain must be interpreted and carefully linked to align with the source domain recommendation practice. In addition, eliciting auxiliary information from a target domain becomes more challenging when there is discrepancy or too much data or noise in one of the sources or target domains that requires extra effort to transfer the correct cold start user preferences to the source domain.

However, some recommendation processes do not specifically involve two distinct processes of eliciting the auxiliary information and then incorporating it into the recommendation process but rather involve the process of converting a cold start recommendation process to a warm start. With this process, few sparse user inferences are obtained, but they are insufficient in making a recommendation; then the second step involves obtaining additional auxiliary information to converting the warm start to generate sufficient information for recommendation. Thus, a cold start recommendation could involve significantly additional amount of time and effort needed compared to a non cold start recommendation.

This study has explained an interesting theme in which most memory-based approaches involve the use of k-nearest neighbor algorithms, for cold start. This involves adapting or modifying these approaches to elicit auxiliary information through tags. This study has found that a cold start recommendation in mobile applications usually involves the use of neighborhood-based approaches. This study has also shown that the use of cross domain data transfers and the use of app data for cross platform mobile apps recommendations have been promising for tackling cold start recommendation in mobile environments. One of the shortcomings and areas of future research in mobile environments is location-based cold start app recommendation because relatively few studies have explicitly focused on location based cold start app recommendation. Furthermore, research is warranted on leveraging contextual information for next app recommendation in mobile app recommendation.

Our study has shown how social network data and platforms are leveraged in eliciting auxiliary information for cold start users. Some approaches do not literally extract these data explicitly and integrate them into the recommendation process, but rather the information or data from similar users or connected users are studied or considered in forming social network profile for a new user so that a new user could be offered recommendations based on the nature of his/her profile. In this case, auxiliary information does not involve explicitly extracting the auxiliary data/information from the social network, but social network profiles are used as inspirations and references to build cold start user profiles based on similar or connected users as depicted in Figure 2. Arguably, the social network profiles being considered and recognized for the cold start recommendation are the implicit auxiliary information.

One of the noticeable findings of this study is that researchers commonly perform a selection or sieving process such as ranking the information or inferences based on certain rules. This is because auxiliary information is arguably not fully reliable and accurate as the information are merely predictions and data from indirect sources. Hence, the information is usually deduced and interpreted by using various techniques; therefore, researchers select and sort these data according to their reliability.

The adoption of deep learning approaches has skyrocketed in the past two decades from various research disciplines [98,99]. The sheer benefits of using deep learning approaches in recommender system has been impressive. This is because deep learning is capable of capturing the non-trivial and non-linear relationships between users and items effectively. Deep learning allows the codification of more complex abstractions as data representations in the higher layers [17], it also handles the complex and sophisticated relationships within data from varied data sources such as with textual, contextual and

visual embeddings [17]. Despite all the revolutionary advantages of deep learning in modern computation and in particular recommender systems, relatively few studies have used deep learning approaches for cold start recommendation [47,92]. It would be interesting to explore deep learning-based models capability of embedding the auxiliary information into a deep feature which would possibly lead to better performance than matrix factorization or traditional filtering strategies. Therefore, further research is warranted in the use of deep learning approaches for cold start recommendations.

Figure 5 below depicts the various approaches adopted in eliciting auxiliary information for cold start user recommendation.

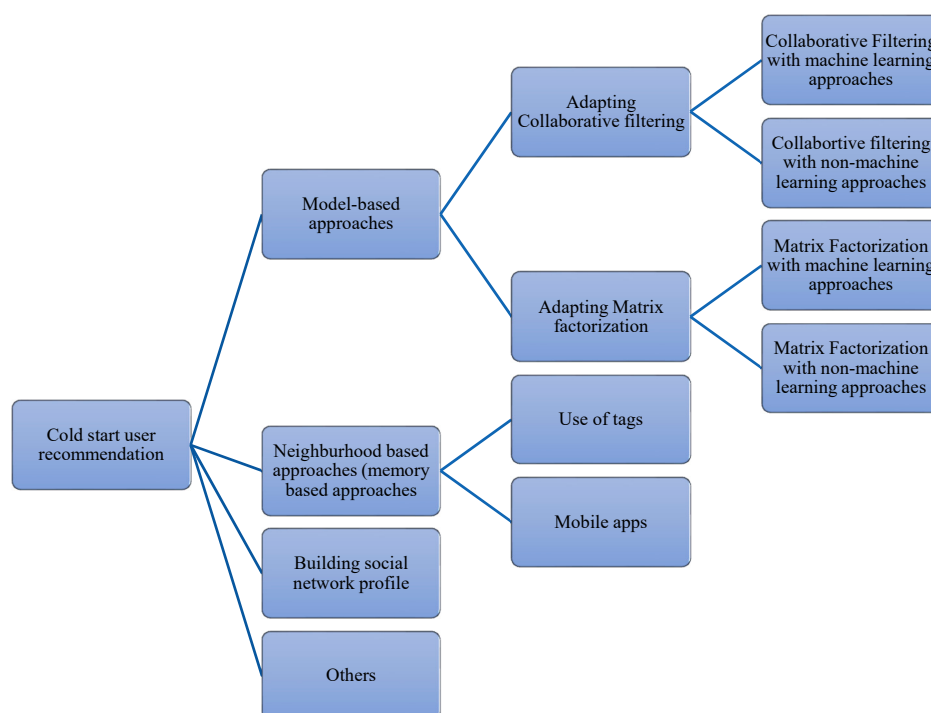


Figure 5. A taxonomy of the approaches for deriving and eliciting cold start user implicit data.

Despite the growth in research activities and the need to provide recommendations to users based on the context in which several review and empirical studies have shown its benefits and accuracy in providing better recommendation to users (see [18,100–103]), we have only found a few studies that offer cold start context-aware recommendations. These are mostly through mobile apps. Our study has clearly shown the need of using possibly the existing and common approaches of matrix factorization based cold start recommendation or neighborhood recommendation to look into offering better recommendation based on users' contexts in cold start situations. Arguably, implicit auxiliary information can be obtained and elicited contextually to offer rich auxiliary information needed for a cold start recommendation in mobile applications. One of the possible reasons why researchers have not focused on context-aware cold start recommendations might be due to the difficulty and challenges associated with identifying the precise and accurate context of cold start users, so as to utilize the accurate auxiliary information, plus merging or coupling the traditional collaborative filtering or its derivatives-matrix factorization models with this contextual information for context-based cold start recommendations.

Another important finding of this study is the use of tag-based recommendations. Although tags are usually associated with certain problems and drawbacks, such as personalization, trust, task categorization (identifying characteristics of tags, its meaning, its origin, etc.) (see [104]), many researchers have leveraged the use of tag information on

recommender systems to improve the performance of traditional recommendation techniques for mitigating cold start recommendation. Because tags carry user information and preferences, the use of tags improve the quality and prediction accuracy of a recommender system. Relevant recommendations are provided through the use of tag information with suggesting the tags that are related to the content of a target object [105,106]. Consequently, once a user or new user's tags or item's keywords are known, it becomes easier to suggest relevant or preferred items to that user using the known tag-keyword relations [107]. Commonly, researchers use matrix factorization matrices using tags to offer cold start recommendation. With tag information, a user-tag matrix that represents users' preferences about tags and convert sparse user-item matrix into dense user-item matrix is constructed. Results shows that the approach generates more precise prediction than general collaborative filtering suffering from the cold start problem. Our results show that the use of tag information have been valuable in providing recommendation to new users. further research is warranted on using tags for user cold start recommendation from domains other than entertainment (music and video) and e-commerce.

Another interesting finding of our study involves educational recommender systems (ERS). Over the years, educational and blended learning institutions have intended to improve their learning systems and the whole online environments or components in which students use technologies for their learning process by designing and customizing the learning systems to offer the suitable recommendations to their students. Educational recommender system has received relatively less attention compared to other research domain such as entertainment (music, movies) or e commerce. Recommending learning materials to students have shown to positively impact students learning outcomes (see [108–112]). The recent study of [113] that recommend learning contents to be adjusted has shown that the future of e-learning heavily relies or is associated with recommender systems in which learning systems would be designed to recommend or provide learners with personalized learning materials based on their contexts, skill, and various other information that defines or identifies a learner. However, our study has found only a single study that proposed an approach for cold start recommendation in educational settings (see [91]). It is alarming to see that new students (cold start students) in various educational institutions are not offered the preferred learning materials based on their preferences. Arguably, educational recommender systems can leverage the sheer amount of students' side information ranging from students' enrolment data, background information, educational history, etc., as auxiliary information for building cold start recommender system to support and ease new students into their studying. Similar approach of using implicit background data for students personalization of students learning materials (see [114]) and scaffolding students peer learning (see [115,116]) have recently been used in various educational settings to support students learning. This way, students' explicit information is indirectly used as an implicit auxiliary information for cold start recommendation. As such, research is strongly warranted to offer recommendation of learning materials as well as peers to support them in their studying.

Regardless of the learning environment in which learners learn, self-regulated learning have been one of the most talked about competences that today's learners must possess in order to flourish [117]. ERS could be used to motivate learners to study continuously in order to obtain better academic results, as well as motivating new students to kick off a good start and to encourage working continuously. It would be interesting to further investigate the correlation between ERS for cold start recommendation with self-regulated learning. In addition, research is warranted not only in ERS but also in scholarly aware recommender system as several influential studies over the years have shown the significance and sheer benefits of recommending scholarly materials to students (see [102,118,119]).

Lastly, emoji recommendation has been one of the recent attractive research dimensions [120] in text classification and recommender systems. Future research should

explore the possibility of leveraging emojis for cold start recommendations with consideration for personal, point of interest (see [121]) and/or contextual data.

5. Conclusions

Cold start has been one of the most challenging and difficult issues to address in recommender systems domains. This study explains the various strategies and techniques employed by researchers in tackling cold start user recommendations from various application domains. This study has looked at and discussed an entirely new dimension in recommender systems. We learned about the difficulties associated with obtaining the correct and reliable auxiliary information in order to recommend preferred items to a cold start user. We believe this study to be the first to explain the processes of identifying, arriving at and extracting the implicit auxiliary information necessary for cold star user recommendations.

Our study has identified and described the various ways that researchers form models, the steps they take for these approaches and the categorization of the various approaches to related themes. We have learned that researchers and recommender system practitioners usually couple traditional filtering matrix or matrix factorization with machine learning algorithms to train data and obtain auxiliary information for cold start users. We have also learned that researchers have used several other algorithms to adapt the popular collaborative filtering, as well as matrix factorization to elicit auxiliary information for cold start recommendation. We learned that cold start recommendation in mobile apps is usually carried out using neighborhood approaches through cross domain recommendations. We have learned that recommendation in mobile app environments and mobile devices is usually associated with contextual or location-based recommendations, and usually involve k-nearest neighbor approach. We have learned that most of the cold start recommendation approaches are carried out in the entertainment and e-commerce research domain (movies and music) and has been less researched in the educational domain. This study has exposed and explained the challenges associated with eliciting auxiliary side information for cold start recommendation, so that researchers and recommender system practitioners could further understand the technicality of a recommendation to a new user. Hopefully, this study will enlighten and add massive knowledge to recommender system researchers and practitioners on a more in-depth dimension or aspect of cold start recommendations, and more importantly encourage research activities concerning cold start recommendations, which have been one of the inherent challenges that recommender systems face.

Author Contributions: All the authors (N.A.A., R.A.R., M.H.N. and M.M.R.) have made substantial contributions in proposing the review, writing and revising the manuscript. All the authors have also approved the final version of the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Ministry of Higher Education, Malaysia under the Impact-Oriented Interdisciplinary Research Grant Programme (IIRG) (Grant. No.: IIRG030C-2019).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Manouselis, N.; Drachsler, H.; Verbert, K.; Duval, E. *Recommender Systems for Learning*; Springer Science & Business Media: New York, NY, USA, 2012.
2. Melville, P.; Sindhvani, V. Recommender systems. *Encycl. Mach. Learn.* **2010**, *1*, 829–838.
3. Ricci, F.; Rokach, L.; Shapira, B. Introduction to recommender systems handbook. In *Recommender Systems Handbook*; Springer: Berlin/Heidelberg, Germany, 2011; pp. 1–35.

4. Desrosiers, C.; Karypis, G. A comprehensive survey of neighborhood-based recommendation methods. In *Recommender Systems Handbook*; Springer: Berlin/Heidelberg, Germany, 2011; pp. 107–144.
5. Koren, Y.; Bell, R. Advances in collaborative filtering. In *Recommender Systems Handbook*; Springer: Berlin/Heidelberg, Germany, 2015; pp. 77–118.
6. Sun, M.; Li, F.; Lee, J.; Zhou, K.; Lebanon, G.; Zha, H. Learning multiple-question decision trees for cold-start recommendation. In Proceedings of the Sixth ACM International Conference on Web Search and Data Mining, Rome, Italy, 4–8 February 2013; pp. 445–454.
7. Miranda, T.; Claypool, M.; Gokhale, A.; Mir, T.; Murnikov, P.; Netes, D.; Sartin, M. Combining content-based and collaborative filters in an online newspaper. In *Proceedings of the ACM SIGIR Workshop on Recommender Systems*; Worcester Polytechnic Institute: Worcester, MA, USA, 1999.
8. Montaner, M.; López, B.; De La Rosa, J.L. A taxonomy of recommender agents on the internet. *Artif. Intell. Rev.* **2003**, *19*, 285–330.
9. Lika, B.; Kolomvatsos, K.; Hadjiefthymiades, S. Facing the cold start problem in recommender systems. *Expert Syst. Appl.* **2014**, *41*, 2065–2073.
10. Nadimi-Shahraki, M.-H.; Bahadorpour, M. Cold-start problem in collaborative recommender systems: Efficient methods based on ask-to-rate technique. *J. Comput. Inf. Technol.* **2014**, *22*, 105–113.
11. He, C.; Parra, D.; Verbert, K. Interactive recommender systems: A survey of the state of the art and future research challenges and opportunities. *Expert Syst. Appl.* **2016**, *56*, 9–27.
12. Natarajan, S.; Vairavasundaram, S.; Natarajan, S.; Gandomi, A.H. Resolving data sparsity and cold start problem in collaborative filtering recommender system using linked open data. *Expert Syst. Appl.* **2020**, *149*, 113248.
13. Camacho, L.A.G.; Alves-Souza, S.N. Social network data to alleviate cold-start in recommender system: A systematic review. *Inf. Process. Manag.* **2018**, *54*, 529–544.
14. Son, L.H. Dealing with the new user cold-start problem in recommender systems: A comparative review. *Inf. Syst.* **2016**, *58*, 87–104.
15. Idrissi, N.; Zellou, A. A systematic literature review of sparsity issues in recommender systems. *Soc. Netw. Anal. Min.* **2020**, *10*, 1–23.
16. Khan, M.M.; Ibrahim, R.; Ghani, I. Cross domain recommender systems: A systematic literature review. *Acm Comput. Surv. (Csur)* **2017**, *50*, 1–34.
17. Zhang, S.; Yao, L.; Sun, A.; Tay, Y. Deep learning based recommender system: A survey and new perspectives. *Acm Comput. Surv. (Csur)* **2019**, *52*, 1–38.
18. Haruna, K.; Akmar Ismail, M.; Suhendroyono, S.; Damiasih, D.; Pierewan, A.C.; Chiroma, H.; Herawan, T. Context-aware recommender system: A review of recent developmental process and future research direction. *Appl. Sci.* **2017**, *7*, 1211.
19. Xu, B.; Bu, J.; Chen, C.; Cai, D. An exploration of improving collaborative recommender systems via user-item subgroups. In Proceedings of the 21st International Conference on World Wide Web, Lyon, France, 16–20 April 2012; pp. 21–30.
20. Koren, Y. Factorization meets the neighborhood: A multifaceted collaborative filtering model. In Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Las Vegas, NV, USA, 24–27 August 2008; pp. 426–434.
21. Kim, H.-N.; El-Saddik, A.; Jo, G.-S. Collaborative error-reflected models for cold-start recommender systems. *Decis. Support Syst.* **2011**, *51*, 519–531.
22. Bobadilla, J.; Ortega, F.; Hernando, A.; Bernal, J. A collaborative filtering approach to mitigate the new user cold start problem. *Knowl. -Based Syst.* **2012**, *26*, 225–238.
23. Zhang, M.; Tang, J.; Zhang, X.; Xue, X. Addressing cold start in recommender systems: A semi-supervised co-training algorithm. In Proceedings of the 37th International ACM SIGIR Conference on Research & Development in Information Retrieval, Gold Coast, QLD, Australia, 6–11 July 2014; pp. 73–82.
24. Pereira, A.L.V.; Hruschka, E.R. Simultaneous co-clustering and learning to address the cold start problem in recommender systems. *Knowl. -Based Syst.* **2015**, *82*, 11–19.
25. Guo, G.; Zhang, J.; Thalmann, D. Merging trust in collaborative filtering to alleviate data sparsity and cold start. *Knowl. -Based Syst.* **2014**, *57*, 57–68.
26. Wei, J.; He, J.; Chen, K.; Zhou, Y.; Tang, Z. Collaborative filtering and deep learning based recommendation system for cold start items. *Expert Syst. Appl.* **2017**, *69*, 29–39.
27. Duricic, T.; Lacic, E.; Kowald, D.; Lex, E. Trust-based collaborative filtering: Tackling the cold start problem using regular equivalence. In Proceedings of the 12th ACM Conference on Recommender Systems, Vancouver, BC, Canada, 2 October 2018; pp. 446–450.
28. Anwaar, F.; Iltaf, N.; Afzal, H.; Nawaz, R. HRS-CE: A hybrid framework to integrate content embeddings in recommender systems for cold start items. *J. Comput. Sci.* **2018**, *29*, 9–18.
29. Jiang, M.; Cui, P.; Chen, X.; Wang, F.; Zhu, W.; Yang, S. Social recommendation with cross-domain transferable knowledge. *Ieee Trans. Knowl. Data Eng.* **2015**, *27*, 3084–3097.
30. Yang, C.; Zhou, Y.; Chen, L.; Zhang, X.; Chiu, D.M. Social-group-based ranking algorithms for cold-start video recommendation. *Int. J. Data Sci. Anal.* **2016**, *1*, 165–175.

31. Hannech, A.; Adda, M.; Mcheick, H. Cold-start recommendation strategy based on social graphs. In Proceedings of the 2016 IEEE 7th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), Vancouver, BC, Canada, 13–15 October 2016; pp. 1–7.
32. Chen, Z.; Shen, L.; Li, F.; You, D. Your neighbors alleviate cold-start: On geographical neighborhood influence to collaborative web service QoS prediction. *Knowl. -Based Syst.* **2017**, *138*, 188–201.
33. Mu, R. A survey of recommender systems based on deep learning. *IEEE Access* **2018**, *6*, 69009–69022.
34. Kiran, R.; Kumar, P.; Bhasker, B. DNNRec: A novel deep learning based hybrid recommender system. *Expert Syst. Appl.* **2020**, *144*, 113054.
35. Bathla, G.; Aggarwal, H.; Rani, R. AutoTrustRec: Recommender system with social trust and deep learning using autoEncoder. *Multimed. Tools Appl.* **2020**, *79*, 20845–20860.
36. Bokde, D.; Girase, S.; Mukhopadhyay, D. Matrix factorization model in collaborative filtering algorithms: A survey. *Procedia Comput. Sci.* **2015**, *49*, 136–146.
37. Xue, H.-J.; Dai, X.; Zhang, J.; Huang, S.; Chen, J. Deep Matrix Factorization Models for Recommender Systems. In Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, Melbourne, Australia, 19–25 August 2017; pp. 3203–3209.
38. Fernández-Tobías, I.; Cantador, I.; Tomeo, P.; Anelli, V.W.; Di Noia, T. Addressing the user cold start with cross-domain collaborative filtering: Exploiting item metadata in matrix factorization. *User Modeling User-Adapt. Interact.* **2019**, *29*, 443–486.
39. Bi, Y.; Song, L.; Yao, M.; Wu, Z.; Wang, J.; Xiao, J. A Heterogeneous Information Network based Cross Domain Insurance Recommendation System for Cold Start Users. In Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval, Virtual Event, 25–30 July 2020; pp. 2211–2220.
40. Ocepek, U.; Rugelj, J.; Bosnić, Z. Improving matrix factorization recommendations for examples in cold start. *Expert Syst. Appl.* **2015**, *42*, 6784–6794.
41. Peng, F.; Lu, J.; Wang, Y.; Yi-Da Xu, R.; Ma, C.; Yang, J. N-dimensional Markov random field prior for cold-start recommendation. *Neurocomputing* **2016**, *191*, 187–199.
42. Zhang, Y.; Shi, Z.; Zuo, W.; Yue, L.; Liang, S.; Li, X. Joint Personalized Markov Chains with social network embedding for cold-start recommendation. *Neurocomputing* **2020**, *386*, 208–220.
43. Chou, S.-Y.; Yang, Y.-H.; Jang, J.-S.R.; Lin, Y.-C. Addressing cold start for next-song recommendation. In Proceedings of the 10th ACM Conference on Recommender Systems, Boston, MA, USA, 15–19 September 2016; pp. 115–118.
44. Zhao, W.X.; Li, S.; He, Y.; Chang, E.Y.; Wen, J.-R.; Li, X. Connecting social media to e-commerce: Cold-start product recommendation using microblogging information. *Ieee Trans. Knowl. Data Eng.* **2015**, *28*, 1147–1159.
45. Mirbakhsh, N.; Ling, C.X. Improving top-n recommendation for cold-start users via cross-domain information. *Acm Trans. Knowl. Discov. Data (TKDD)* **2015**, *9*, 1–19.
46. Li, C.-T.; Hsu, C.-T.; Shan, M.-K. A cross-domain recommendation mechanism for cold-start users based on partial least squares regression. *Acm Trans. Intell. Syst. Technol. (TIST)* **2018**, *9*, 1–26.
47. Zhu, F.; Wang, Y.; Chen, C.; Liu, G.; Orgun, M.; Wu, J. A deep framework for cross-domain and cross-system recommendations. *arXiv* **2020**, arXiv:2009.06215.
48. Barjasteh, I.; Forsati, R.; Masrour, F.; Esfahanian, A.-H.; Radha, H. Cold-start item and user recommendation with decoupled completion and transduction. In Proceedings of the 9th ACM Conference on Recommender Systems, Vienna, Austria, 16–20 September 2015; pp. 91–98.
49. Rana, C.A.Q.; Salima, H.; Usama, F.; Hammam, C. From a “cold” to a “warm” start in recommender systems. In Proceedings of the 2014 IEEE 23rd International WETICE Conference, Parma, Italy, 23–25 June 2014; pp. 290–292.
50. Zhu, J.; Zhang, J.; Zhang, C.; Wu, Q.; Jia, Y.; Zhou, B.; Philip, S.Y. CHRS: Cold start recommendation across multiple heterogeneous information networks. *IEEE Access* **2017**, *5*, 15283–15299.
51. Kumbhar, N.; Belerao, K. Microblogging Reviews Based Cross-Lingual Sentimental Classification for Cold-Start Product Recommendation. In Proceedings of the 2017 International Conference on Computing, Communication, Control and Automation (ICCUBEA), Pune, India, 1–18 August 2017; pp. 1–4.
52. Zhang, J.-D.; Chow, C.-Y.; Xu, J. Enabling kernel-based attribute-aware matrix factorization for rating prediction. *IEEE Trans. Knowl. Data Eng.* **2016**, *29*, 798–812.
53. Roy, S.; Guntuku, S.C. Latent factor representations for cold-start video recommendation. In Proceedings of the 10th ACM Conference on Recommender Systems, Boston, MA, USA, 15–19 September 2016; pp. 99–106.
54. Ji, K.; Shen, H. Jointly modeling content, social network and ratings for explainable and cold-start recommendation. *Neurocomputing* **2016**, *218*, 1–12.
55. Xu, X.; Yuan, D. A novel matrix factorization recommendation algorithm fusing social trust and behaviors in micro-blogs. In Proceedings of the 2017 IEEE 2nd International Conference on Cloud Computing and Big Data Analysis (ICCCBDA), Chengdu, China, 28–30 April 2017; pp. 283–287.
56. Tomeo, P.; Fernández-Tobías, I.; Di Noia, T.; Cantador, I. Exploiting linked open data in cold-start recommendations with positive-only feedback. In Proceedings of the 4th Spanish Conference on Information Retrieval, Granada, Spain, 14–16 June 2016; pp. 1–8.
57. Ostuni, V.C.; Di Noia, T.; Di Sciascio, E.; Mirizzi, R. Top-n recommendations from implicit feedback leveraging linked open data. In Proceedings of the 7th ACM Conference on Recommender Systems, Hong Kong, China, 12–16 October 2013; pp. 85–92.

58. Meymandpour, R.; Davis, J.G. Enhancing Recommender Systems Using Linked Open Data-Based Semantic Analysis of Items. In Proceedings of the 3rd Australasian Web Conference (AWC 2015), Sydney, Australia, 27–30 January 2015; pp. 11–17.
59. Nazari, Z.; Charbuillet, C.; Pages, J.; Laurent, M.; Charrier, D.; Vecchione, B.; Carterette, B. Recommending Podcasts for Cold-Start Users Based on Music Listening and Taste. In Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval, Sydney, Australia, 27–30 January 2015; pp. 1041–1050.
60. Kim, B.S.; Kim, H.; Lee, J.; Lee, J.-H. Improving a recommender system by collective matrix factorization with tag information. In Proceedings of the 2014 Joint 7th International Conference on Soft Computing and Intelligent Systems (SCIS) and 15th International Symposium on Advanced Intelligent Systems (ISIS), Kitakyushu, Japan, 3–6 December 2014; pp. 980–984.
61. Aggarwal, C.C. Neighborhood-based collaborative filtering. In *Recommender Systems*; Springer: Berlin/Heidelberg, Germany, 2016; pp. 29–70.
62. Zhu, T.; Ren, Y.; Zhou, W.; Rong, J.; Xiong, P. An effective privacy preserving algorithm for neighborhood-based collaborative filtering. *Future Gener. Comput. Syst.* **2014**, *36*, 142–155.
63. Rosli, A.N.; You, T.; Ha, I.; Chung, K.-Y.; Jo, G.-S. Alleviating the cold-start problem by incorporating movies facebook pages. *Clust. Comput.* **2015**, *18*, 187–197.
64. Belém, F.M.; Heringer, A.G.; Almeida, J.M.; Gonçalves, M.A. Exploiting syntactic and neighbourhood attributes to address cold start in tag recommendation. *Inf. Process. Manag.* **2019**, *56*, 771–790.
65. Elahi, M.; Hosseini, R.; Rimaz, M.H.; Moghaddam, F.B.; Trattner, C. Visually-Aware Video Recommendation in the Cold Start. In Proceedings of the 31st ACM Conference on Hypertext and Social Media, Virtual Event, 13–15 July 2020; pp. 225–229.
66. Baeza-Yates, R.; Jiang, D.; Silvestri, F.; Harrison, B. Predicting the next app that you are going to use. In Proceedings of the eighth ACM International Conference on Web Search and Data Mining, Shanghai, China, 2–6 February 2015; pp. 285–294.
67. Liu, J.; Shi, J.; Cai, W.; Liu, B.; Pan, W.; Yang, Q.; Ming, Z. Transfer Learning from APP Domain to News Domain for Dual Cold-Start Recommendation. In Proceedings of the RecSysKTL, Como Italy, 27 August 2017; pp. 38–41.
68. Han, D.; Li, J.; Li, W.; Liu, R.; Chen, H. An app usage recommender system: Improving prediction accuracy for both warm and cold start users. *Multimed. Syst.* **2019**, *25*, 603–616.
69. Tu, Z.; Fan, Y.; Li, Y.; Chen, X.; Su, L.; Jin, D. From fingerprint to footprint: Cold-start location recommendation by learning user interest from app data. *Proc. AcM Interact. Mob. Wearable Ubiquitous Technol.* **2019**, *3*, 1–22.
70. Sassi, I.B.; Mellouli, S.; Yahia, S.B. Context-aware recommender systems in mobile environment: On the road of future research. *Inf. Syst.* **2017**, *72*, 27–61.
71. Woerndl, W.; Huebner, J.; Bader, R.; Gallego-Vico, D. A model for proactivity in mobile, context-aware recommender systems. In Proceedings of the Fifth ACM Conference on Recommender Systems, Chicago, IL, USA, 23–27 October 2011; pp. 273–276.
72. Woerndl, W.; Brocco, M.; Eigner, R. Context-aware recommender systems in mobile scenarios. *Int. J. Inf. Technol. Web Eng. (Ijitve)* **2009**, *4*, 67–85.
73. Bouneffouf, D.; Bouzeghoub, A.; Gançarski, A.L. A contextual-bandit algorithm for mobile context-aware recommender system. In Proceedings of the International Conference on Neural Information Processing, Doha, Qatar, 12–15 November 2012; pp. 324–331.
74. Bouneffouf, D.; Bouzeghoub, A.; Gançarski, A.L. Hybrid- ϵ -greedy for mobile context-aware recommender system. In Proceedings of the Pacific-Asia Conference on Knowledge Discovery and Data Mining, Kuala Lumpur, Malaysia, 29 May 2012; pp. 468–479.
75. Del Carmen Rodríguez-Hernández, M.; Ilarri, S. AI-based mobile context-aware recommender systems from an information management perspective: Progress and directions. *Knowl. -Based Syst.* **2021**, *215*, 106740.
76. Khan, A.; Ahmad, A.; Rahman, A.U.; Alkhalil, A. A mobile cloud framework for context-aware and portable recommender system for smart markets. In *Smart Infrastructure and Applications*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 283–309.
77. Lin, J.; Sugiyama, K.; Kan, M.-Y.; Chua, T.-S. New and improved: Modeling versions to improve app recommendation. In Proceedings of the 37th International ACM SIGIR Conference on Research & Development in Information Retrieval, Gold Coast, QLD, Australia, 6–11 July 2014; pp. 647–656.
78. Liu, B.; Kong, D.; Cen, L.; Gong, N.Z.; Jin, H.; Xiong, H. Personalized mobile app recommendation: Reconciling app functionality and user privacy preference. In Proceedings of the Eighth ACM International Conference on Web Search and Data Mining, Shanghai, China, 2–6 February 2015; pp. 315–324.
79. Cao, D.; He, X.; Nie, L.; Wei, X.; Hu, X.; Wu, S.; Chua, T.-S. Cross-platform app recommendation by jointly modeling ratings and texts. *AcM Trans. Inf. Syst. (Tois)* **2017**, *35*, 1–27.
80. Lin, J.; Sugiyama, K.; Kan, M.-Y.; Chua, T.-S. Addressing cold-start in app recommendation: Latent user models constructed from twitter followers. In Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval, Dublin, Ireland, 28 July–1 August 2013; pp. 283–292.
81. Liu, H.; Jing, L.; Wen, J.; Xu, P.; Yu, J.; Ng, M.K. Bayesian Additive Matrix Approximation for Social Recommendation. *AcM Trans. Knowl. Discov. Data (TKDD)* **2021**, *16*, 1–34.
82. Zhao, G.; Lei, X.; Qian, X.; Mei, T. Exploring users' internal influence from reviews for social recommendation. *IEEE Trans. Multimed.* **2018**, *21*, 771–781.
83. Tang, J.; Hu, X.; Liu, H. Social recommendation: A review. *Soc. Netw. Anal. Min.* **2013**, *3*, 1113–1133.
84. Victor, P.; De Cock, M.; Cornelis, C. Trust and recommendations. In *Recommender Systems Handbook*; Springer: Berlin/Heidelberg, Germany, 2011; pp. 645–675.

85. Wang, X.; Lu, W.; Ester, M.; Wang, C.; Chen, C. Social recommendation with strong and weak ties. In Proceedings of the 25th ACM International on Conference on Information and Knowledge Management, Indianapolis, IN, USA, 24–28 October 2016; pp. 5–14.
86. Herce-Zelaya, J.; Porcel, C.; Bernabé-Moreno, J.; Tejeda-Lorente, A.; Herrera-Viedma, E. New technique to alleviate the cold start problem in recommender systems using information from social media and random decision forests. *Inf. Sci.* **2020**, *536*, 156–170.
87. Gao, H.; Tang, J.; Liu, H. Addressing the cold-start problem in location recommendation using geo-social correlations. *Data Min. Knowl. Discov.* **2015**, *29*, 299–323.
88. Fernández-Tobías, I.; Braunhofer, M.; Elahi, M.; Ricci, F.; Cantador, I. Alleviating the new user problem in collaborative filtering by exploiting personality information. *User Modeling User-Adapt. Interact.* **2016**, *26*, 221–255.
89. Nguyen, V.-D.; Sriboonchitta, S.; Huynh, V.-N. Using community preference for overcoming sparsity and cold-start problems in collaborative filtering system offering soft ratings. *Electron. Commer. Res. Appl.* **2017**, *26*, 101–108.
90. Shapira, B.; Rokach, L.; Freilikhman, S. Facebook single and cross domain data for recommendation systems. *User Modeling User-Adapt. Interact.* **2013**, *23*, 211–247.
91. Pliakos, K.; Joo, S.-H.; Park, J.Y.; Cornillie, F.; Vens, C.; Van den Noortgate, W. Integrating machine learning into item response theory for addressing the cold start problem in adaptive learning systems. *Comput. Educ.* **2019**, *137*, 91–103.
92. Salem, Y.; Hong, J.; Liu, W. CSFinder: A cold-start friend finder in large-scale social networks. In Proceedings of 2015 IEEE International Conference on Big Data (Big Data), Santa Clara, CA, USA, 29 October–1 November 2015; pp. 687–696.
93. Ebesu, T.; Fang, Y. Neural semantic personalized ranking for item cold-start recommendation. *Inf. Retr. J.* **2017**, *20*, 109–131.
94. Hernando, A.; Bobadilla, J.; Ortega, F.; Gutiérrez, A. A probabilistic model for recommending to new cold-start non-registered users. *Inf. Sci.* **2017**, *376*, 216–232.
95. Viktoratos, I.; Tsadiras, A.; Bassiliades, N. Combining community-based knowledge with association rule mining to alleviate the cold start problem in context-aware recommender systems. *Expert Syst. Appl.* **2018**, *101*, 78–90.
96. Hong, D.-G.; Lee, Y.-C.; Lee, J.; Kim, S.-W. CrowdStart: Warming up cold-start items using crowdsourcing. *Expert Syst. Appl.* **2019**, *138*, 112813.
97. Zhang, X.; Yuan, X.; Li, Y.; Zhang, Y. Cold-Start Representation Learning: A Recommendation Approach with Bert4Movie and Movie2Vec. In Proceedings of the 27th ACM International Conference on Multimedia, Nice, France, 21–25 October 2019; pp. 2612–2616.
98. Kulathilake, K.; Abdullah, N.A.; Bandara, A.; Lai, K.W. InNetGAN: Inception Network-Based Generative Adversarial Network for Denoising Low-Dose Computed Tomography. *J. Healthc. Eng.* **2021**, *2021*, 9975762.
99. Kulathilake, K.S.H.; Abdullah, N.A.; Sabri, A.Q.M.; Lai, K.W. A review on Deep Learning approaches for low-dose Computed Tomography restoration. *Complex Intell. Syst.* **2021**, 1–33; <https://doi.org/10.1007/s40747-021-00405-x>.
100. Villegas, N.M.; Sánchez, C.; Díaz-Cely, J.; Tamura, G. Characterizing context-aware recommender systems: A systematic literature review. *Knowl. -Based Syst.* **2018**, *140*, 173–200.
101. Panniello, U.; Tuzhilin, A.; Gorgoglione, M. Comparing context-aware recommender systems in terms of accuracy and diversity. *User Modeling User-Adapt. Interact.* **2014**, *24*, 35–65.
102. Champiri, Z.D.; Shahamiri, S.R.; Salim, S.S.B. A systematic review of scholar context-aware recommender systems. *Expert Syst. Appl.* **2015**, *42*, 1743–1758.
103. Zheng, Y.; Mobasher, B.; Burke, R. Emotions in context-aware recommender systems. In *Emotions and Personality in Personalized Services*; Springer: Berlin/Heidelberg, Germany, 2016; pp. 311–326.
104. Jäschke, R.; Hotho, A.; Mitzlaff, F.; Stumme, G. Challenges in tag recommendations for collaborative tagging systems. In *Recommender Systems for the Social Web*; Springer: Berlin/Heidelberg, Germany, 2012; pp. 65–87.
105. Martins, E.F.; Belém, F.M.; Almeida, J.M.; Gonçalves, M.A. On cold start for associative tag recommendation. *J. Assoc. Inf. Sci. Technol.* **2016**, *67*, 83–105.
106. Belém, F.M.; Almeida, J.M.; Gonçalves, M.A. A survey on tag recommendation methods. *J. Assoc. Inf. Sci. Technol.* **2017**, *68*, 830–844.
107. Ji, K.; Shen, H. Addressing cold-start: Scalable recommendation with tags and keywords. *Knowl. -Based Syst.* **2015**, *83*, 42–50.
108. Abdi, S.; Khosravi, H.; Sadiq, S.; Gasevic, D. Complementing educational recommender systems with open learner models. In Proceedings of the Tenth International Conference on Learning Analytics & Knowledge, Frankfurt, Germany, 23–27 March 2020; pp. 360–365.
109. Erdt, M.; Fernandez, A.; Rensing, C. Evaluating recommender systems for technology enhanced learning: A quantitative survey. *IEEE Trans. Learn. Technol.* **2015**, *8*, 326–344.
110. Ghauth, K.I.; Abdullah, N.A. Learning materials recommendation using good learners' ratings and content-based filtering. *Educ. Technol. Res. Dev.* **2010**, *58*, 711–727.
111. Ghauth, K.I.; Abdullah, N.A. The effect of incorporating good learners' ratings in e-Learning content-based recommender System. *J. Educ. Technol. Soc.* **2011**, *14*, 248–257.
112. Albatayneh, N.A.; Ghauth, K.I.; Chua, F.-F. Utilizing learners' negative ratings in semantic content-based recommender system for e-learning forum. *J. Educ. Technol. Soc.* **2018**, *21*, 112–125.
113. Wan, H.; Yu, S. A recommendation system based on an adaptive learning cognitive map model and its effects. *Interact. Learn. Environ.* **2020**, 1–19; <https://doi.org/10.1080/10494820.2020.1858115>.

114. Rahman, M.M.; Abdullah, N.A. A personalized group-based recommendation approach for Web search in E-learning. *IEEE Access* **2018**, *6*, 34166–34178.
115. Rasheed, R.A.; Kamsin, A.; Abdullah, N.A. An Approach for Scaffolding Students Peer-Learning Self-Regulation Strategy in the Online Component of Blended Learning. *IEEE Access* **2021**, *9*, 30721–30738.
116. Rasheed, R.A.; Abdullah, N.A.; Kamsin, A.; Ahmed, M.A.; Yahaya, A.S.; Umar, K. A framework for designing students peer learning self-regulation strategy system for blended courses. In Proceedings of the 2021 1st International Conference on Emerging Smart Technologies and Applications (eSmarTA), Sana'a, Yemen, 10–12 August 2021; pp. 1–5.
117. Rasheed, R.A.; Kamsin, A.; Abdullah, N.A. Challenges in the online component of blended learning: A systematic review. *Comput. Educ.* **2020**, *144*, 103701.
118. Haruna, K.; Akmar Ismail, M.; Damiasih, D.; Sutopo, J.; Herawan, T. A collaborative approach for research paper recommender system. *PLoS ONE* **2017**, *12*, e0184516.
119. Champiri, Z.D.; Asemi, A.; Binti, S.S.S. Meta-analysis of evaluation methods and metrics used in context-aware scholarly recommender systems. *Knowl. Inf. Syst.* **2019**, *61*, 1147–1178.
120. Zhao, G.; Liu, Z.; Chao, Y.; Qian, X. CAPER: Context-aware personalized emoji recommendation. *IEEE Trans. Knowl. Data Eng.* **2020**, *33*, 3160–3172.
121. Wu, Y.; Li, K.; Zhao, G.; Xueming, Q. Personalized long-and short-term preference learning for next POI recommendation. *IEEE Trans. Knowl. Data Eng.* **2020**, doi: 10.1109/TKDE.2020.3002531.