Vol 5 No 1 (2023) *e*-ISSN: 2682-8383

Journal of Engineering Technology and Applied Physics

Hybrid-based Recommender System for Online Shopping: A Review

Ying Fei Lim, Su Cheng Haw*, Kok Why Ng and Elham Abdulwahab Anaam
Faculty of Computing and Informatics, Multimedia University, Jalan Multimedia, Cyberjaya, Malaysia.
**Corresponding author:* sucheng@mmu.edu.my, *ORCiD:* 0000-0002-7190-0837
https://doi.org/10.33093/jetap.2023.5.1.3
Manuscript Received: 8 February 2023, Accepted: 21 February 2023, Published: 15 March 2023

***Abstract* - In the era of the digital revolution, online shopping has developed into a remarkably simple and economical option for consumers to make purchases securely and conveniently from their homes. In order for the online merchant to optimize their profit, the online shopping platform must always display a list of potential products that customers may purchase. The recommender system kicks in at this point to assist in finding products that customers would like and recommend a list of product recommendations that match the customer's preferences. This paper reviews the recommender system technology in detail by reviewing the classification technique. Other than that, the related works will be reviewed to understand how each technique works, the strengths and limitations, the datasets and evaluation metrics employed.**

Keywords - Recommender System, Online shopping, Semantic, Hybrid-based, Descriptive analytics

I. INTRODUCTION

The expansion of and the tendency towards electronic commerce due to the technological advancements in Internet-delivered devices and services has an impact on most individuals nowadays. The numerous and varied products available on online shopping platforms can occasionally overwhelm customers and make it challenging to choose the right or perfect item. Due to these circumstances, there is more competition among international commercial websites, which raises the demand for productive work to boost financial earnings. Moreover, the development of a recommendation system improves the performance of online shopping systems by assisting customers in identifying the right products based on their interests. The primary goal of any recommendation engine is to boost demand and encourage customer interaction. Recommender engines are primarily a component of a personalization approach for online shopping, enhancing customer experience by dynamically loading various products to emails, applications, or websites. These diverse and cross-channel recommendations are created from various data sources, including customer interests, historical transaction records, attributes, or contextual information.

Besides that, recommender systems are machine learning systems that guide customers to find new things. Recommender systems can also be thought of as information filtering systems that assist in prioritizing and personalizing information based on the preferences of customers, hence decreasing information overload. Furthermore, enormous volumes of dynamically generated data are the source of data that recommender systems will use to prioritize, filter, and efficiently distribute pertinent information while lessening the human cognitive burden. Moreover, recommender systems can improve revenue and profit. It helps a business stand out from rivals when the recommender systems are effectively developed and implemented. In essence, a product recommendation engine is a technology piece that allows marketers to provide pertinent product recommendations to customers in real-time. In addition, recommender systems, which function as effective data filtering tools, utilize data analysis techniques and machine learning algorithms to suggest the most suitable products to a particular customer. It can be done by learning data (such as past customer behaviours, preferences, and interests) and anticipating a customer's present. The techniques applied to recommend products have to be considered judiciously to produce recommendations relatable for customers as the techniques can greatly influence the performance and accuracy of the recommendation systems.

This paper explores data filtering techniques widely utilized in the recommender systems, such as the collaborative filtering technique, content-based filtering technique, hybrid-based filtering technique, semantic-based filtering technique, ontology-based filtering technique, and graph-based filtering technique. These techniques will be discussed briefly to comprehend their concepts and how each technique works by reviewing the related previous research papers. This paper will dive deeper into various semantic-based recommendation techniques focusing on using ontology for modeling semantic data and relationships. In addition, a review of the appropriate evaluation metrics to examine and assess performance and accuracy is also presented.

II. RECOMMENDER SYSTEM

A. *Overview of Recommender System*

Online shopping has grown by leaps and bounds over the past few years. With the development of online shopping, the total amount of user data generated is increasing daily. The enormous development in the number of online users and the volume of digital information available presents the possible issue of information overload, obstructing seasonable access to online products of preference. Besides that, in the era of information overload, the significance of recommender systems that provide prioritized and personalized information for online products and services is increasing rapidly. Hence, to extract items corresponding to users' interests and preferences, a well-developed recommendation system emerges to improve the user's online shopping experience.

A recommendation system can be described as an information filtering system that manages information overload problems by removing essential pieces of information from a huge amount of dynamically generated information in accordance with users' preferences, interests, or observed item characteristic. Next, a recommender system is an instrument that utilizes an array of algorithms, data analysis, and artificial intelligence to search for similar users and similar items based on their behavior and to generate online recommendations of items (products or services) that specific customers ought to like. The recommender system generally allocates users personalized service support by learning users' previous behaviors and estimating users' current interests for specific items [1].

With the fast expansion of online shopping, recommendation systems have introduced the requirement for information filtering techniques, facilitating users by filtering out the information that matches their preferences and tastes. Furthermore, Hu *et al.* [2] stated that the recommender system had attracted great attention from huge e-commerce companies, including Tmail.com and Amazon.com. In fact, many companies have discovered that recommender systems not only recommend items that are suitable for customers, but also play a massive role in turning viewers into purchasers, improving cross-selling, and cultivating purchaser loyalty. Therefore, recommender systems are typically used in commercial applications, especially e-commerce.

B. *Phases in the Recommendation Process*

The recommendation process has three main phases: the information gathering phase, the learning phase, and the recommendation or prediction phase. Figure 1 shows the overview of the phases included in the recommendation process.

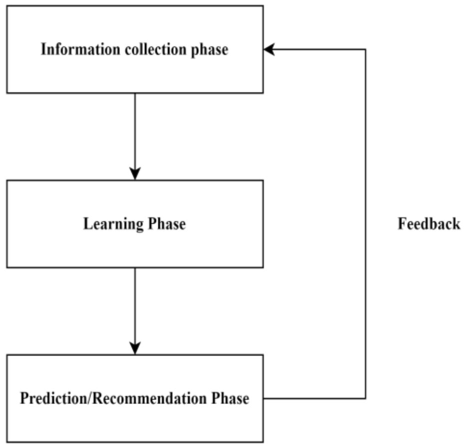

Fig. 1. Phases in the Recommendation Process.

During the information collecting phase, relevant customer information is gathered to create user profiles or models for prediction tasks. The information collected includes the users' behaviors, attributes, or the content of users' access to resources. The recommender system needs to collect many user information to suggest the best reasonable recommendations. Besides, the recommendation system relies on a variety of inputs such as high-quality explicit feedback that includes explicit input where the interest in items with respect to the users, and implicit feedback that indirectly deduces user preferences by monitoring user behavior. In addition, by combining implicit and explicit feedback allows for the collection of hybrid feedback.

In the explicit feedback, the recommender system asks for feedback from users to provide recommendations. The recommendation system's efficiency and quality rely on users' ratings. Furthermore, the drawback of this approach is that it requires users' efforts, and users occasionally lack the readiness to give adequate information. However, it is still viewed as offering more trustworthy information because it does not involve the extraction of interests from operations. It also offers perspicuity in the recommendation process, leading to marginally higher accuracy for the perceptual recommendation. On the other hand, in the implicit feedback, the recommender system automatically determines users' interests by analyzing various operations of users, such as the user's purchases transactions, links the user clicks, browsing history, email content, time spent on the web pages, and more. Besides, although this feedback does not require user action, it is less accurate as this approach automatically generates recommendations by examining the aforementioned contents. The feedback can be merged as hybrid feedback. With that, the drawbacks of both implicit and explicit feedback are eliminated, and their advantages are integrated to constitute hybrid feedback. Hybrid feedback can be obtained by using the implicit data as a recommendation attribute while permitting users to submit explicit feedback and ratings.

In the learning phase, the user data collected from the feedback in the information collection phase is filtered and utilized by applying learning algorithms. The learning algorithms are methods that aid in drawing patterns that are appropriate for application in particular circumstances.

In this prediction/recommendation phase, recommendations for the given data are made by analysing the patterns gathered from the learning phase. The trained data gathered throughout the learning phase provides particular patterns, which are then constrained to envision the user's behavioral trajectory or future interests.

III. RECOMMENDER SYSTEM TECHNIQUES

A recommender system is a piece of software for discovering and predicting the most relevant items to users that they are interested in and providing personalized recommendations in real-time. Various recommender system techniques are applied to manage data overload and suggest products of interest to users in accordance with dynamically generated data. Besides that, it is crucial to adopt accurate and efficient techniques for a system to strengthen the recommendation system's effectiveness and provide useful and reliable recommendations to individual users. This can help companies to acquire and retain customers by offering them personalized deals. Hence, this emphasizes the importance of understanding the characteristics, potentials, advantages, limitations, and differences among the recommender system techniques.

Next, recommender systems generally are categorized into three data filtering techniques, namely, hybrid-based filtering technique, collaborative filtering technique, and content-based filtering technique. Besides that, other personalized recommender system techniques exist, such as semantic-based, ontology-based, and graph-based filtering techniques. Figure 2 illustrates the classification of recommender system filtering techniques.

A. Content-based Recommender System

Content-based (CB) filtering technique generates recommendations of products to users that are deemed to be similar to products that are known to have been liked by users. The resemblance of products is computed according to the characteristics associated with the compared products [3]. Furthermore, Shruthi and Gripsy [4] demonstrated CB filtering technique depends on information about product contents and ratings of products by users. CB filtering technique integrates these ratings with a profile of user preferences according to the characteristics of the products being reviewed, and the recommendation engines can then discover products that users have previously preferred. Besides, recommendations in CB filtering techniques are based on personal information while disregarding user efforts.

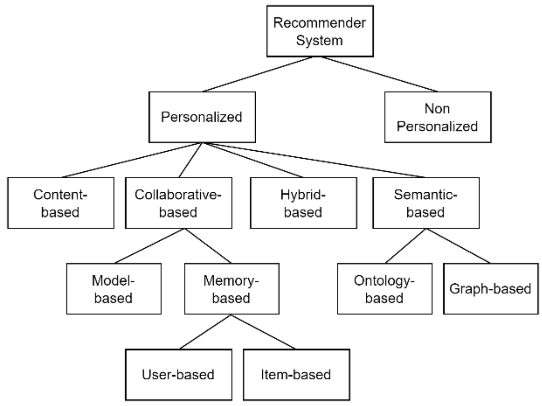

Fig. 2. Classification of Recommender System.

Besides that, Kaur and Bathla [5] stated that the CB filtering technique recommends items based on user profiles and users' item profiles. The user profiles are set up when users start the system. The recommender system gathers users' interests and recommends products after investigating the characteristics of items and users. Moreover, the recommended products are the same as those that the users have favored before, and the recommended items also match the user attributes. Apart from this, the CB filtering technique only performs effectively when attributes are presented appropriately and clearly. In CB filtering techniques, the descriptions of items and user profiles are of great importance.

Furthermore, the CB filtering technique applies various models to discover similarities among documents and give useful recommendations. According to Isinkaye *et al.* [6], this technique applies Probabilistic models, namely Neural Networks, Naïve Bayes Classifier or Decision Trees or Vector Space Model, namely Term Frequency Inverse Document Frequency (TF-IDF) to model the relationships between distinct documents in the collected works. These techniques generate recommendations by learning potential models using machine learning or statistical analysis techniques. Other than that, the CB filtering technique does not require other user profiles, as the user profiles do not influence recommendations. In addition, if user profiles alter, the CB filtering technique has the capacity to regulate recommendations in a timely manner. The main weakness of this technique is that it requires a detailed comprehension and elaboration of the characteristics of the profile's items.

B. Collaborative Filtering Technique

Collaborative filtering (CF) technique is the most commonly implemented technique applied in recommender systems [7]. Besides that, CF techniques depend on user history, in the form of user ratings for items as the source of information. This technique can be achieved by establishing relationships between items or between users [4]. Moreover, a significant feature of CF techniques is that the recommender system does not recognize the properties of users and items, and only knows the interactions between them [8]. This feature makes this technique a good selection for recommending complex items without clear descriptive keywords.

Besides that, the CF technique uses the implicit knowledge of users' community of applied products to determine the relationship of these products to other users who have not noticed or used these products [9]. Furthermore, CF technique operates by constructing a user × items matrix representing user interests for items. Then, it compares users with corresponding interests by computing the similarity between user profiles to generate recommendations [10]. In addition, Li *et al.* demonstrated that implementing a CF technique is simple when dealing with low data dependency and further provides accurate recommendations [11].

Furthermore, the technique of CF can be further branched into two categories, namely model-based CF technique and memory-based CF technique.

Model-based CF technique leverages existing ratings to learn a model to enhance the effectiveness of CF techniques. This technique populates the matrix by predicting items the users have not seen before. According to Patel *et al.* [12], data mining or machine learning techniques can be employed to complete the model-building process and forecast ratings for items that are unrated. Furthermore, with regard to this technique, instead of using the dataset every time, the model for recommendation is generated according to the information collected from the repository [13].

Moreover, the model-based CF technique can quickly recommend a series of products as it uses precomputed models, and it has been shown to generate similar neighbourhood-based recommendation results. Besides that, scalability and speed are factors that enhanced using this technique. It also improves the algorithm's prediction accuracy value [14]. There are several examples of these techniques: Clustering, Matrix Completion Technique, Dimensionality Reduction, Regression, Latent Semantic methods, and Singular Value Decomposition (SVD). In addition, matrix factorization is the most commonly used model-based CF technique [15].

Memory-based CF technique is a heuristic algorithm that produces the rating of a product according to the ratings of other users [16]. In this technique, the similarity among users or products is computed to give recommendations of products, and the users with the same preferences will have higher similarity values [11]. Besides that, the memory-based CF technique finds customers with the similar interests as the targeted user customer and then forecasts the interests of the targeted user customer for new products. Each user with similar interests and preferences is grouped together. Next, memory-based CF techniques can be classified into two types, such as item-based CF techniques and user-based CF techniques.

User-based CF technique assesses the resemblance between users in accordance with their ratings of specific items [17]. Besides that, the major concept of the user-based CF technique is to choose a user neighborhood that is similar to users, analyze their preferences, and then provide users with recommendations for items that are liked by the neighborhood users [18]. Recommending items to users depending on the ratings given by similar users is the idea of the user-based CF technique. Lagerstedt and Olsson [19] stated that the main conjecture made is that users who previously shared the same preferences and interests will continue to do so in the future.

Furthermore, the user-based CF technique is transformed into an item-based CF technique, which produces predictions based on the similarities of items [20]. In the item-based CF technique, the idea is to recommend products to customers that are similar to products the customers have previously rated. Besides that, to make predictions, this technique computes the similarity among items [17]. There are two phases in the item-based CF technique. The first phase is to apply similarity measures to assess the commonality between two objects. The ratings of unknown objects are then predicted using the similarity values.

C. Hybrid-based Filtering Technique

The aim of hybrid-based (HB) filtering technique is to increase the efficiency and accuracy of the recommender systems. It attains better performance by integrating the implementation of two or more recommendation algorithms or components in a single recommender system, thereby mitigating the weaknesses of the CB filtering technique and CF technique and benefiting from their strengths [21]. HB filtering technique combines two or more recommender system techniques to improve system optimization while reducing the shortcomings and limitations of any single recommendation technique. The idea of an HB filtering technique is that utilizing two or more algorithms in combination will result in more efficient and accurate recommendations than just one algorithm since the weaknesses of an algorithm can be made up for by the strengths of some algorithm [22].

In general, there are seven different approaches identified in the HB filtering technique.

Weighted hybridization: Outcomes of multiple recommender systems are combined through weighted hybridization, integrating the scores of each technique used through a linear formula to generate a prediction or a list of recommendations. Furthermore, weighted hybridization is an approach of gradually adjusting the weight according to the degree of agreement between the evaluation of a product by user and the evaluation forecasted by the recommender systems. The weighted hybridization combined all the advantages of the recommender systems straightforwardly throughout the recommendation process.

Switching hybridization: Switching hybridization is an approach to changing the recommendation model used according to the situation and can avoid issues specific to one approach. Moreover, this strategy has the advantage that the recommender system is aware of the weaknesses and strengths of its particular recommenders [6]. On the other hand, switching hybridization has the disadvantage that it ordinarily making the recommendation process more complex, as switching criteria must be determined which usually grows in the total amount of parameters of the recommender systems.

Cascade hybridization: An iterative refinement procedure is involved in cascade hybridization for building an order of interest among various products. Recommendations for first technique are improved by another recommender system technique. In other words, the first recommender system technique generates a general list of recommendation, which is then improved by the subsequent recommender system technique. Due to the coarse-to-fine nature of the iterations, the hybridization technique is extremely noise-tolerant and efficient [6]. After producing a candidate set that closely matches the user preferences by applying one of the recommender system models, cascade hybridization integrates the previously employed recommender system model with a different model to rank the candidate set according to which products that best fit the user preferences.

Mixed hybridization: Mixed hybridization combines the recommendation results of the recommender system technique simultaneously, instead of just one recommendation for one product. Every product has numerous recommendations connected to it from various recommender system techniques. Individual performance in mixed hybridization does not necessarily influence the overall performance of a particular area [6]. In mixed hybridization, recommendations from various recommender systems are simultaneously integrated.

Feature combination: The features derived by a particular recommender system technique are joined into those from another in feature combination hybridization. The benefit of feature combination is that it allows this technique to occasionally not solely depend on collaborative data.

Feature Augmentation: Feature augmentation hybridization utilizes the ratings and other details generated by the preceding recommender system and also needs extra features of the recommender system. The input of the second recommendation system will use the result from the first recommender system. In addition, feature augmentation hybridization is better than feature combination hybridization because this technique adds a few new features to the main recommendation system.

Meta-level: An approach to applying the whole model of one recommender system as input data in another recommender system model is known as meta-level hybridization. In the comparison of using raw rating data as single input data, the operation of the Collaborative Mechanism is easier to use because user preferences are compacted and conveyed using meta-level hybridization. According to Isinkaye *et al.* [6], in meta-level hybridization, an internal model produced by a recommender system technique is applied as input for another recommender system technique. There is always more information in the resulting model than just a single rating.

D. Semantic-Based Filtering Technique

Semantic-based (SB) filtering techniques can perform better than traditional filtering techniques by adding meta-knowledge representing the semantic characteristics or attributes of the items to be recommended [23]. This technique guarantees that the features of recommended items fit the preferences and interests of users. Moreover, information about users, such as user preferences, is typically stored in a personal data structure called profiles. This information is collected explicitly by requesting users to fill in particular attributes about items, or it can be gathered by implicit inference. Besides that, user interactions with the system, such as reviews, ratings, and transaction history, are collected and analyzed to infer user preferences. The most appropriate approach to building relationships between user profiles and item features is to apply domain ontology (reference ontology) to represent items. In addition, user profiles are stored in an ontology to be associated with the reference ontology.

Furthermore, there are four phases in the proposed SB filtering technique. In the first phase, the reference ontology of the items is constructed using information from different knowledge sources. An ontology user profile is produced for each user so that it is associated with the reference ontology through the interest score attribute. Next, the second phase shows the update and refinement of user profiles using the propagation activation technique. The third phase searches for the targeted user's nearest neighbors. With the help of nearest neighbor information, the rating value of the target item is predicted. The last phase aims to evaluate the performance of the recommendation system using Mean Absolute Error (MAE) along with other evaluation metrics such as Recall, F-Measure, and Precision. Moreover, in the SB filtering technique, two methods can be applied: the ontology and graph database.

Ontology-Based Filtering Technique: Ontology-based (OB) filtering technique leverages hierarchical organizations of items and users to improve profile building, recommendation, and browsing [24]. The concept of an ontology in computer science is initially described by Gruber [25] as a conceptualization's explicit specification. Ontologies are applied to represent knowledge domains that formally describe lists of terms. Moreover, ontologies usually consist of relationships between concepts and a vocabulary. Ontologies provide formal semantics that can be applied to process and combine a range of information on the internet. One of the main objectives of using an OB filtering technique is to model information at the semantic level. Besides that, ontologies provide equipment for the formal modelling of a system's structure in accordance with the relationships that generate from its observations. In addition, in recommendation systems, the semantic information of products comprises attributes, relationships between products, and the relationship between products and meta information.

Furthermore, ontology user profiles allow the use of inference, allowing the discovery of interests that are not explicitly seen in user behavior [26]. By restricting user-interested examples to a single ontology, all users can share instances of ontology classes, thereby expanding the size of the classifier training set. However, binary-class classification is fundamentally more accurate than multi-class classification, which lowers the accuracy of classification. Besides that, profiles can interact with other ontologies that have the same concepts after being represented by ontologies. This permit using an external knowledge base to bootstrapping the recommendation system and lower the cold-start issue affecting all recommendation systems.

Graph-Based Filtering Technique: Graph-based (GB) filtering technique represents the relationship between items and users as a bipartite graph, where there are unweighted or weighted links between the user and each product rated by the user [27]. In the GB method, the data is modelled in the form of graphs where the edges represent the similarities between items and users, and the nodes represent users or items, or both [28]. GB filtering technique mainly focuses on the construction of graphs. Next, according to Kamta and Verma [28], when constructing recommender system tools, researchers encountered many issues: data-sparse, scalability, startup, lack of time and resources, and information overload. These issues can lower the prediction accuracy of recommender systems. Hence, the researchers modelled the rating data as graphs in order to get around these issues. The transitive associations captured by graphs are very useful for recommendations of items because they are able to deal with scarcity and limited coverage. However, this technique is basically designed for a rating or binary feedback and suffers from severe deficiencies for ranking-oriented categories of neighbor-based CF. Moreover, the current GB filtering technique fails to capture the user preferences order. Besides that, the weakness of the current GB filtering technique is raised for binary implicit feedback, this technique unable to capture the user pairwise preferences produced by different implicit feedbacks [29].

IV. SUMMARY OF RECOMMENDER SYSTEM TECHNIQUES

In previous sections, several data filtering techniques of recommender systems have been discussed. Each technique operates differently from the other. Hence, it is worth taking the effort and time to understand the features, potentials, and how each technique works to deploy the best data filtering technique for the recommender system for your purposes. In this section, we will compare the differences between the data filtering techniques of recommender systems, especially their pros and cons. Table I illustrates the advantages and limitations of each recommender system data filtering technique.

Overall, each recommender system data filtering technique has its advantages and limitations. CB filtering technique does not require other user profiles as the user profiles do not influence recommendations. If the user profiles alter, this technique can potentially regulate recommendations in a timely manner. Kaur and Bathla [5] stated that this technique also provides user independence based on proprietary ratings used to build user profiles. Furthermore, in this technique, there is a great deal of transparency for users to understand how the recommender system processes. In addition, this technique recommends items that have not been used by any users, which indirectly benefits current users.

On the other hand, the CB filtering technique only works well when user attributes are presented in an appropriate and precise manner. This technique requires a thorough comprehension and elaboration of the characteristics of items in the profile. Besides that, it is hard to create attributes of products in some specific areas. This technique also suffers from over-specialization because the items it recommends are all of similar types. In this technique, it is unable to check whether the recommendations are correct. The reason is that in this technique, user feedback are not gathered as users do not provide ratings for the products. In addition, this technique only applied user statistics not user interaction data [5].

Next, in the CF technique, the recommendation systems do not recognize the attributes of items and users and only knows the interactions between them, and this feature makes this technique a good selection for recommending complex items without precise descriptive keywords. According to Kaur and Bathla [5], implementing a memory-based CF technique simplifies the recommendation process. Furthermore, applying a model-based CF technique can improve prediction performance. Besides, in the memory-based CF technique, it is easy to include new data incrementally. On the contrary, in the CF technique, the recommendation systems have insufficient information about the items or users to make corresponding predictions (cold start problem). Besides that, the CF technique has a scalability problem, the recommender system technique that is effective with a limited number of datasets may not be able to generate a sufficient quantity of recommendations when the number of datasets increases.

Table I. Advantages and limitations of Data Filtering Technique of the Recommender System.

Advantages	Limitations
Content-based (CB):	
-Does not require other user profiles as it do not influence recommendations. -Can potentially regulate recommendations in a timely manner -Transparency for users to understand how the recommender system processes. -Recommends items that have not been used by any users, which indirectly benefits current users.	-Only works well when user attributes are presented in a precise manner. -Requires a thorough comprehension and elaboration of the characteristics of items in the profile. -Hard to create attributes of products in some specific areas. -Suffers from over-specialization. -Only applies user statistics, not user interaction data.
Collaborative Filtering (CF):	
-Good selection for recommending complex items without precise descriptive keywords. -Memory-based CF technique makes the recommendation process simpler. -Model-based CF technique can improve prediction performance (easy to include new data incrementally).	-Cold start problem: The recommendation system has insufficient information about the items or users for making corresponding predictions. -Scalability: The recommender system technique that is effective with a limited number of datasets possibly unable to generate a sufficient quantity of recommendations when the number of datasets increases. -Data sparsity problem: The recommender system undergoes this problem when there are a small amounts of products available in the database that have user ratings.
Hybrid-based:	
-Benefits from the strengths of multiple recommendation techniques and mitigates the drawbacks of any single recommendation technique. -Can achieve better system optimization by generating more accurate and reliable recommendations.	-More computational power is required. -More complex in terms of space and time. -Implementation of this technique is expensive.
Semantic-based (SB):	
-Performs better than traditional filtering techniques.	-More complex technique.
Ontology-based (OB):	
-Provide a rich conceptualization of an organization's work domain. -Ontology user profiles allow the use of inference, which allows the discovery of interests that are not specifically seen in user behavior. -By restricting user-interested examples to a single ontology, all users can share instances of ontology classes, thereby expanding the size of the classifier training set. -Ontology user profiles can interact with other ontologies that have the same concepts after being represented by ontologies, and this permits the use of an external knowledge base to facilitate bootstrapping the recommendation system and lower the cold-start issue that affects all recommendation systems.	-Multi-class classification in this technique is fundamentally less accurate than binary-class classification, which also reduces the accuracy of classification.
Graph-based (GB):	
-Transitive associations captured by graphs are very useful for recommendations of items because they are able to deal with scarcity and limited coverage.	-Basically, designed for a rating or binary feedback and suffers from a severe deficiency for ranking-oriented categories of neighbor-based CF. -Fails to capture the user preferences order. -For binary implicit feedback, this technique unable to capture the user pairwise preferences produced by different implicit feedbacks. -Relies on contextual information which is not present or unavailable to the system in all applications and can be costly to gather.

Moreover, this technique also suffers from data sparsity problems. The recommender system undergoes this problem when a small number of products are available in the database with user ratings [6]. Furthermore, the hybrid-based filtering technique benefits from the strengths of multiple recommendation techniques and mitigates the drawbacks of any single recommendation technique. This technique can achieve better system optimization by generating more accurate and reliable recommendations. On the other hand, more computational power is required for this technique. Besides that, this technique is more complex in terms of time and space because two or more different data filtering techniques are combined to work as one recommender system [30]. In addition, the implementation of a HB filtering technique is expensive [18].

Moreover, the SB filtering technique performs better than traditional filtering techniques because the meta-knowledge representing the semantic characteristics or attributes of the items to be recommended is added. However, the SB filtering technique is more complex [31]. Besides that, the OB filtering technique can provide a rich conceptualization of an organization's work domain, representing the major concepts and relationships of work activities. Apart from this, Middleton *et al.* [26] demonstrated that ontology user profiles allow the use of inference, which enables the discovery of interests that are not explicitly seen in user behavior. By restricting user-interested examples to a single ontology, all users can share instances of ontology classes, thereby expanding the size of the classifier training set. Furthermore, ontology user profiles can interact with other ontologies that have the same concepts after being represented by ontologies, and this permits the use of an external knowledge base to facilitate bootstrapping the recommendation system and lower the cold-start issue that affects all recommendation systems. However, in this technique, multi-class classification is fundamentally less accurate than binary-class classification, which lowers the accuracy of classification [26]. Moreover, in the GB filtering technique, transitive associations captured by graphs are very useful for recommendations of items because they are able to deal with scarcity and limited coverage [28]. On the contrary, the GB filtering technique is basically designed for a rating or binary feedback and suffers from a severe deficiency for ranking-oriented categories of neighbor-based CF. This technique also fails to capture the user preferences order. For binary implicit feedback, this technique is unable to capture the user pairwise preferences produced by different implicit feedbacks. In addition, most of these techniques rely on contextual information which is not present or unavailable to the system in all applications and can be costly to gather [29].

V. RELATED WORKS

This section will further elaborate on the related research works on HB, SB, and OB recommender systems In 2010, Elgohary *et al.* [31] proposed a SB recommender system with the application of Wikipedia as an ontology. The authors applied Wikipedia as an ontology to address the issues of applying conventional ontologies for text analysis in text-based recommender system. Hence, a complete system model is proposed to combine SB analysis with collaboration through a content recommendation system. Figure 3 demonstrates the major components of the proposed system and their embeddings.

The authors stated that the Wikipedia annotator serves as the recommender system's semantic analyzer. Every written document is marked up with Wikipedia terminology and saved in the repository.

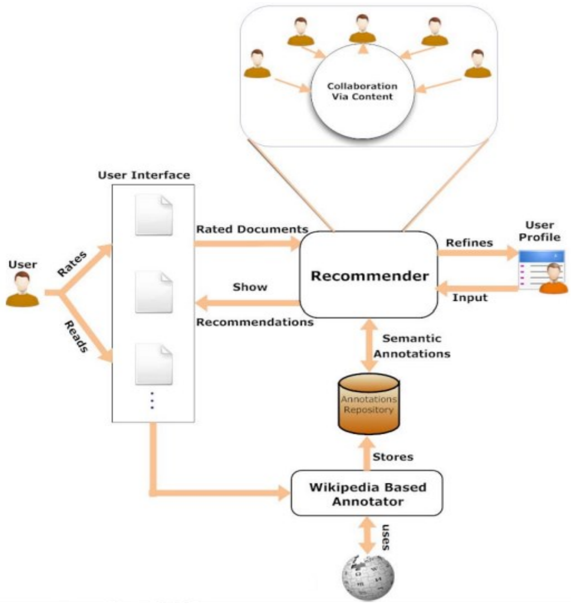

Fig. 3. Proposed model by Elgohary *et al.* [31].

Furthermore, each user must also keep a profile up to date with the information that represents their preferred topics. The moment a user reviews contents, semantic annotations for the contents are pulled from the document repository and applied to refine the profile of user. Besides, to execute the algorithm of recommendation, the recommender components leverage user profiles to discover users with the same preferences. Then, the outcome of the recommendation is a list of documents in which the user is most preferable to be interested in. Moreover, a Wikipedia-Based Semantic analyzer is proposed in this research. Wikipedia articles are utilized as concepts by the authors to annotate the documents. The reasons for using this method instead of applying the traditional ontologies are the extensive coverage of many concepts that makes Wikipedia an appropriate broad domain ontology. Furthermore, articles on current events are frequently added to Wikipedia. Besides that, new relationships between various concepts are only implicitly defined by mentioning the category of each new article. As mentioned for these two factors, Wikipedia is the current state-of-the-art ontology and can generate more accurate outcomes, especially when used in online text mining applications. Apart from this, Wikipedia is a multilingual ontology since it is accessible in a large total amount of different languages. In addition, the concepts of Wikipedia are clearly explained in quite lengthy text chunks, removing any semantic ambiguity in the notions. The suggested semantic analysis model is essentially the foundation of the Explicit Semantic Analysis (ESA) model. The lexical similarity between the documents and the concepts is represented by the weighted vector of Wikipedia conceptions used to annotate the documents. The authors suggested that a text should only contain a small number of notions (those that best match lexically), thus minimizing the inclusion of noisy concepts. Moreover, more related notions are found and added to the concept vector beginning from these notions without introducing more noisy concepts. For this purpose, the authors use one of the associative retrieval techniques, which is the spreading activation technique. An approach was presented to use the spreading activation. Furthermore, spreading activation reduces noisy concepts by propagating from the first N nodes and decrementing the weights of each pulse. This results in a more accurate semantic representation of the content. Therefore, spreading activation is preferable depending on the foundational notions applied in the ESA model. Spreading activation also improves annotation with additional relevant notions. In order to enhance the quality of semantic annotation, the authors applied another method, which the authors called a concept hierarchy-based method in this study. The Wikipedia category graph structure in spreading activation is applied to identify parent-child relationships. As an alternative, it can be applied to reweight notions based on the hierarchical structure. The higher-level (more general) are allocated less weight than the lower-level notions (more specific ones). Besides that, the authors utilized the Vector Space Model as representation of user profiles. The profile is shown as a vector of weighted notions that have been weighted according to the model from before. Every user profile consists of two notion vectors, the NEGATIVE concept vector that model user-unfavorable notions and the weights (degree of aversion), and the POSITIVE notion vector that model user-attracting notions and the weights (degree of attractiveness). In order to understand the user profiles, these vectors are regularly adjusted based on the user feedback regarding the appropriateness of the suggested items.

Rocchio's algorithm of relevance feedback is employed to learn the user profiles. Information retrieval research served as the Rocchio's algorithm. Initially, relevance feedback was applied to enhance search outcomes by gathering user opinions on the usefulness of the documents that were retrieved. Each user profile is modelled as a two-category document classifier. From the notion vectors of user-rated documents, user profiles are learned. The process of learning is accomplished by integrating document vectors into a prototype vector cj for each class Cj. Firstly, a summary of the normalized document vectors for positive examples of a class and the normalized document vectors for negative examples of a class is provided. Then, the prototype vector is computed as a weighted difference. Equation (1) shows the formula to calculate the prototype vector.

$$\overrightarrow{C_j} = a\frac{1}{|c_j|}\sum_{\vec{d}\in\overrightarrow{C_j}}\frac{\vec{d}}{\|\vec{d}\|} - \beta\frac{1}{|D-C_j|}\sum_{\vec{d}\in D-C_j}\frac{\vec{d}}{\|\vec{d}\|} \quad (1)$$

Where:

- a and β are parameters that alter the relative influence of negative and positive examples of training.
- C_j is the group of training documents given to class j.
- $\|\vec{d}\|$ indicates the Euclidian length of a vector d.

With the establishment of the user profile model, clarifying the recommendation model is necessary. The collaborative recommendation model looks for similarities between users to provide recommendations. Individual users' rating patterns are generally applied to identify user similarities. This association is most significant when users have numerous products rated similarly. The authors expected the number of common item ratings among users to be lower in some practical cases that the collaborative method would be anticipated to fail. These similarities are applied as weighting factors for collaborative recommendations given by neighboring users in a collaborative recommendation architecture. Furthermore, the evaluation process is split into two sections: an evaluation of the recommendation technique and an evaluation of the semantic annotation of documents. A benchmark of 50 documents with a "human-judged" inter-document similarity matrix is applied to assess the semantic annotation component. Besides that, this benchmark is applied to evaluate the ESA model. The semantic similarity of documents exhibits the semantic annotation's accuracy. In order to assess the commonalities, the authors applied the cosine similarity between document interpretation vectors. The model's accuracy is demonstrated by the correlation between the outcoming similarity matrix of the suggested model and the "human-judged" similarity matrix. Besides that, the authors chose to set the quantity of articles N at 200 in order to prevent the pitfalls of the ESA model by removing the influence that the quantity of articles (N) applied had on the quality of the annotations. In the experimental results, the authors observed that the suggested model attains a greater correlation than the greatest value obtained by the ESA model (0.72) at different N (category) values. Besides that, the authors also observed that weighting concepts in accordance with their position in the category graph, as suggested, attains better outcomes than the spreading activation with no reweighting. In addition to this, the authors stated that the correlation obtained by the suggested model (concepts hierarchy-based method) did not vary considerably with the quantity of categories. This indicates that the concept-hierarchy-based method is more reliable than earlier methods. Moreover, the authors conducted two experiments to ensure the suggested recommendation system is performing properly and to demonstrate how the system analysis component affects the resolution of the recommendation issues. In the experiments, the authors gathered 70 blog posts in various categories, such as lifestyle, technology, sports, and politics. There are 20 users who participated in these experiments. Table II illustrates the accuracy outcomes for the quantity of categories of 100 and the number of pulses of 2.

Table II. Accuracy outcomes of Recommendation System [31].

Evaluation metrics	Normal circumstance	Cold start condition
F-measure	0.85714	0.842
Recall	0.9	0.842
Precision	0.8181	0.842
RMSE	1.93028	1.865

Based on Table II, the authors presented that the recommendation system's accuracy under cold-start conditions is near enough to normal circumstances. This demonstrates the advantage of enhancing the concept of user profiles with the proposed improved semantic annotation model.

In 2015, Martinez-Cruz *et al.* [32] proposed a recommendation system by employing ontologies to enhance the user profiles representation. The architecture of the recommender system consists of four elements:

a. On2Trust (Trust network ontology) - An ontology that simulates user trust. The authors explicitly defined fuzzy characteristics and classes to manage fuzzy data in this ontology.
b. Domain ontology - An ontology for the semantic classification of system items. This ontology permits for establishment of relationships between items and users.
c. Database - A generic database for storing ontologies and data. Ontologies represent the semantic structure of the system, while data are kept in databases to maintain the system's effectiveness.
d. Recommendation engine - It depicts the computer-based information classification in an ontology and recommendation-making process.

Users are retained in the system by On2Trust, an ontology, according to their level of trust. Ontology models this concept as the property of an object whose presence symbolizes the trustworthiness between two users. Besides, the reliability level of one user is relative to another specific user, not the other way around, hence this characteristic has an asymmetric value. Next, the system employs a domain ontology to maintain the items' semantic organization. The authors stated that this ontology can be a ‘‘lightweight’’ ontology since they only need a simple classification of items. Moreover, the relationship between the items and users is built using both the On2Trust and Domain ontologies to convey the level of pleasure and projected level of correlation. In addition, the authors created an ontology called On2Trust to describe user trust, and they also applied fuzzy linguistic modelling to help in the description of various concepts. Therefore, in the recommendation generation process, the authors considered individuals who each user may trust rather than those with comparable rating histories.

Furthermore, by introducing domain ontologies into the system, the proposed approach can illustrate the relationships between individuals and their interests in particular items. Besides that, the authors presented their ontology and developed an approach for aggregating trust information obtained in the trust ontology and updating user profiles in response to comments. Apart from this, the fundamental concept of the proposed recommendation method is to prioritize trustworthy users, or users that each user may particularly trust, rather than users with the same rating records when generating recommendations. To attain this, the authors suggested an approach to assess the trust score between two users. This approach discovers every route between a pair of users, investigating On2Trust. Finally, the trust information expressed in the most pertinent routes discovered between the two users is then aggregated. Furthermore, to design the experiments, the authors divided the dataset into two sets: a validation set containing the 20% of the data and a training set including the remaining 80% of the data. Therefore, the authors used ontology in both situations, first conducting the interpretation on the training set before testing against the validation set. Besides, the process is repeated numerous times but with alternative partitions to reduce the bias introduced by the way these testing and training subsets are selected. The original set is partitioned into k parts, and the process is repeated k times so that each partition is utilized as the test set at least once. This approach is called as k-fold cross validation. The objective of performing this cross-validation is to obtain a value that can be used to analyze the accuracy of the various algorithms that are used to predict ratings. Moreover, the authors applied MAE and coverage metrics to assess the suggested method. The predictability of predictions made by the recommendation system increases with decreasing MAE. Equation (2) shows the formula of MAE.

$$\text{MAE} = \frac{\Sigma_{i=1}^{n} |P_i - r_i|}{n} \tag{2}$$

Where:

- r denotes the group of ratings created by the users.
- p the predicted ratings by the system.

In the field of recommendation systems, the evaluation of various methods is performed based on both online and offline tests. Online tests need the system to be fully operational to collect large amounts of data, making this validation costly and time-consuming. Hence, most recommendation systems in the research field are validated by offline tests, where a predetermined set of data is applied. This guides the authors to employ the offline testing method. Specifically, the authors contrasted various system parameter setups as well as the new recommendation method by using the collaborative recommendation method. In the experimental results, the authors presented that the results obtained show an improvement over the previous proposals.

Moreover, Badriyah *et al.* [33] developed a hybrid recommender system on the basis of user profile and item profile. Figure 4 demonstrates the main representation of the recommender system, showing the process of the recommendation system generating a list of item recommendations to display to users.

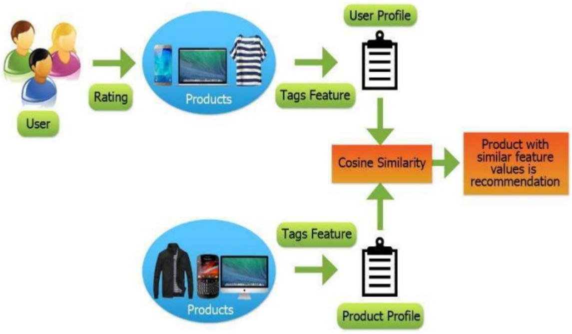

Fig. 4. Recommender System proposed by Badriyah [33].

Based on Fig. 4, several components and processes are involved in the recommendation process. These are searching for product characteristics, rating products, forming user profiles, forming product profiles, and then the matching process until a product listing with similar characteristics (product recommendation list) is finally generated. Following the rating of a product by user, a user profile can be established. Each product in the recommendation process contains all the tags and scores of all other items, resulting in the product profile. Next, a matching procedure that determines the degree of separation between the two profiles can be carried out using the item profile and user profile. The formula of cosine distance is used to compute the similarity between the matrixed profiles. Besides that, there is a certain distance between each product and the user, and then the product recommendations are exhibited to the user in decreasing order of distance from highest to lowest. In their research, the authors applied the TF-IDF approach to automatically generate labels from item descriptions. The TF-IDF technique generally aims to understand the importance of the quantity of associated words between documents. Equation (3) shows the formula for calculating TF-IDF.

$$\mathrm{TFIDF}_{d,t} = \mathrm{FREQ}_{d,t}(1 + \log(\frac{N}{DFREQ_T})) \quad (3)$$

Where:

- FREQd,t = total amount of term t in document d.
- N = total amount of documents applied.
- $DFREQ_T$ = total amount of documents where term t seems.

It is a component of the recommender system's CB filtering technique. The cosine similarity approach is then used to integrate the product profile, which takes the product tags' form, with the user profile. Furthermore, two items are viewed as two vectors in the user space dimension m in the cosine similarity approach. The similarity between items is determined by computing the cosine of the angles between the two vectors. According to the mathematical principle, two vectors are considered to be equal if their cosine is equal to one or their angle is zero degrees. The formal term for the similarity between items A and B is sim (A, B). Equation (4) illustrates the formula of sim(A, B).

$$sim(A,B) = \cos\theta = \frac{A \cdot B}{\|A\|\|B\|} = \frac{\Sigma_{i=1}^{n} A_i B_i}{\sqrt{\Sigma_{i=1}^{n} A_i^2}\sqrt{\Sigma_{i=1}^{n} B_i^2}} \quad (4)$$

Where:

- $sim(A,B)$ is a similarity metric between vectors A and B, $A_i B_i$ is the components of vectors A and B respectively.

The cosine similarity approach is commonly applied in CF techniques in recommender systems. Therefore, the authors combined two approaches, namely CB filtering technique and the CF technique in the proposed system. Next, the advantage of identifying product profiles by automatically creating tags is that the recommendation process becomes more dynamic and efficient. It is efficient as the tag of the product does not require manual input by the administrator, and it is more dynamic as the collected tag results are adjusted to the product description's content. The labels will also alter automatically if the description of the item is changed. Besides that, using product characteristics as labels is more descriptive than identifying items on the basis of the item specification, year of manufacture, and category, as is commonly used in CB filtering techniques. It is more descriptive, as the possibility that one product or other products, and the use of manufacturing year. There must be so many items produced in the same year, so, finding recommended items may be less relevant. Besides that, the computation of the Recall and Precision values as well as the computation of the operation time are utilized to gauge the performance analysis of the recommendation. Equation (5) and Equation (6) are the formulas applied to compute the values of Recall and Precision respectively.

$$\text{Precision} = \frac{|\{Relevant\} \cap \{Retrieved\}|}{|\{Retrieved\}|} \quad (5)$$

$$\text{Recall} = \frac{|\{Relevant\} \cap \{Retrieved\}|}{|\{Relevant\}|} \quad (6)$$

The experimental results demonstrated that the recommendations have commonalities with the product description. Additionally, the user profile preference has an average Recall value of 71.47% and Precision value of 67.5% on a scale from 0% to 100%. From these values, it can be concluded that the Recall value is higher than the Precision value based on the product user rating tags. The magnitude of Recall Precision and Precision also greatly relies on how to determine the document's relevance and what "relevant item" really means. Furthermore, achieving an ideal Recall Precision level is difficult because they are based on dynamic and flexible correlation measures. Next, the time required to execute the program during generating recommendations and creating product characteristics in the form of tags is calculated. Moreover, the web browser's page load time application is a necessary supplementary application for the computation of the execution time. The experimental results revealed that the average execution time needed to create the tag features is 7.7 seconds. This high execution time is due to reading the product description content, which relies on the quantity of characters in the product description information. Due to the algorithm's placement in the system administration part, the users are not inconvenienced by the lengthy execution time.

Furthermore, Shruthi and Gripsy [4] proposed a product recommender for e-commerce websites by employing the hybrid recommender technique. The suggested system implemented a new recommender system with two recommendation techniques, namely, CF technique and demographic analysis, to execute effective product recommendations on e-commerce applications and improve customer satisfaction. The item-based CF technique discovers neighborhoods of items that are the same as the user-selected item. This neighbourhood is constructed by discovering some commonalities between each item and other objects in the current system. Moreover, the item-item model addresses these issues in systems with more users than items. Due to each item typically having more ratings than each user and because there are more users than items, the average rating of an item usually does not alter rapidly. As a result, the model's rating distribution becomes more stable, reducing frequent model rebuild requirements. When users rate items, users get similar items after evolving their profiles and datasets of features of items and other users. Next, due to personalization, especially for users with a larger number of rating records in the field of online shopping, users usually select items by themselves and are rarely influenced by others. In order to provide product recommendations without affecting the personality of seasoned users, the authors suggested an ideal tailored recommendation system. Apart from this, Re-Order Point (ROP), a level of inventory, initiates the replenishment of a particular inventory stock. The replenishment lead time and forecasted usage during the safety stock period are typically utilized for this computation. Assumedly, ordering and purchasing materials in the Economic Order Quantity (EOQ) happen simultaneously, and there is no time delay. Reorder levels for replenishment are triggered when inventory levels drop to zero. The inventory levels rebounded due to immediate deliveries from suppliers. By using the ROP technique, the authors can only decide when to order, and they cannot work out what quantity to order when placing an order. Equation (7) illustrates the formula of reorder level.

$$\text{Reorder Level} =$$

$$\text{Lead Time in Days} \times \mathit{Daily\ Avearge} \tag{7}$$

The ROP may differ for each item in stock because each item in stock may have various usage rates and may take a varied amount of time to obtain replenishment delivery from suppliers. Besides, the authors applied the demographic analysis technique to understand the population's gender, age, and ethnic makeup as well as how it has altered over time. Any type of dynamic living population can be analyzed using this technique, such as a population that alters through space or time. In addition to this, this technique provides measurable characteristics of a provided population, as well as the fundamental demographic processes of death, birth, and the computation of similarity value between users in a similar category using inter-personal, intra-personal, and product-based methods. These are vital tasks for the recommender system. The system presented the effectiveness of the proposed model, taking into account reordering and demographic methods as well as intrapersonal and interpersonal heuristics. The system also takes into account the user preferences' independence in the field of online shopping. This implies that it can, to a certain extent, make product recommendations according to the user interests, which also takes advantage of the user associations with products, especially for the migration of current and more anticipated users.

Next, Nilashi *et al.* [34] proposed a recommender system based on a CF technique that applies ontology with the help of clustering and dimensionality reduction techniques. They addressed two major deficiencies of recommendation systems in CF, such as scalability and sparsity. In the CF technique, the authors also applied a dimensionality reduction technique, namely SVD, to discover the most similar users and items in each user and item cluster. This can significantly enhance the recommendation approach scalability. Figure 5 illustrates the proposed recommender system.

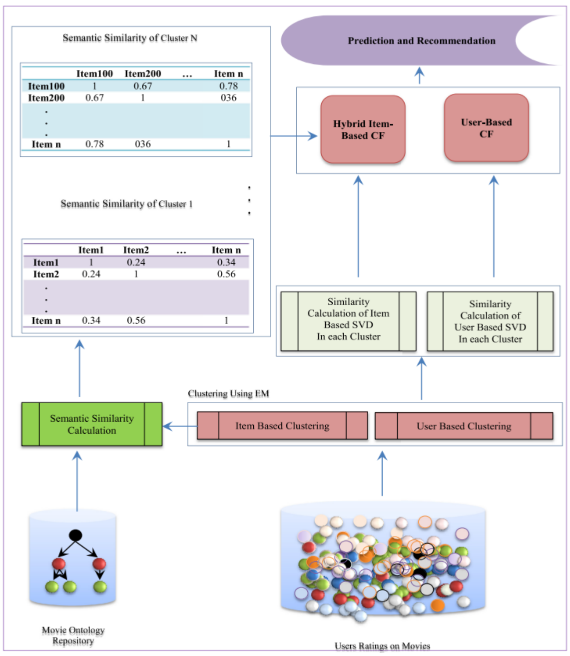

Fig. 5. Framework of proposed System by Nilashi *et al.* [34].

The suggested approach seeks to generate scalable and accurate recommendations, and the proposed approach considers two main phases. The recommendation model is built in the first phase. Several tasks are performed in this phase, including dimensionality reduction using SVD, clustering ratings, and generating similarity matrices of the users and items. In the initial step, the authors applied the Expectation Maximization (EM) algorithm to cluster the user ratings of movies. After that, the authors gave each cluster the semantic similarity computation matrices from the movie ontology repository. At the same time, on each cluster, the authors performed SVD to get the decomposition metrics. Based on Fig. 5, the authors developed the SVD models for items and users. Therefore, each matrix can be subject to an effective similarity computation following the process of decomposing the matrices. After executing the initial training of the models in the offline phase, the prediction and associated recommendation roles are carried out for a target user in the second phase. In effect, a sorted list of recommended products is produced by the recommendation system to the target user. For this purpose, in the first phase, the target user is allocated to one of the clusters identified. After that, the SVD computation is carried out in accordance with previous ratings to discover the similarities between the target user and other users by discovering the target user's neighbors. The authors also utilized a similar process for the items for item-based recommendations. Then, the authors finally applied a weighted method to combine the item- and user-based predictions. Moreover, two real-world datasets in the area of movie suggestions supplied by Yahoo! Webscope R4 and MovieLens are used to assess the complexity of time (scalability) and prediction accuracy of the suggested approach. The gathered data is applied to build and complete the item ontology. Besides that, the dataset is split into two sets: the testing set and the training set to test the model. The testing set was chosen at random 20% of the date, and the training set was selected from the remaining 80%. The experimental results indicated that the throughput of the approaches using dimensionality reduction and clustering techniques is significantly higher than other approaches. Furthermore, for all datasets, the SVD, ontology, and EM approach has higher throughput than approaches that depend only on the nearest neighbor algorithm. This is because by using clustering, the recommendation algorithms use a small subset of the neighbors. Moreover, the authors observed that the throughput of the approaches increased as the clustering size increased, however, the nearest neighbour algorithm's throughput is unaffected by the quantity of clusters because it must scan every neighbour. Next, the MAE between the forecasted rating and actual rating is then determined through statistical metrics. In contrast, the recommended products are compared with the related products by decision support metrics, such that by calculating the overlap. Equation (8) is the formula of MAE applied.

$$MAE(pred, act) = \sum_{i=1}^{N} \frac{pred_{u,i} - act_{u,i}}{N} \quad (8)$$

Where:

- N is the quantity of items on which user u has stated a comment.

The prediction accuracy of the suggested approach was assessed by applying MAE, and it was compared to Pearson's nearest neighbour algorithm (an item-based prediction approach with ontology, clustering, and SVD) as well as item-based, user-based, and prediction approaches with ontology, SVD, and clustering but without ontology. According to the authors' research, the proposed approach, which ontology and SVD facilitate, significantly enhances MAE's prediction accuracy compared to Pearson's nearest neighbour algorithm in the considered neighbor sizes. The user-based and item-based + EM + SVD + ontology approach has a better performance than the item-based + EM + SVD + ontology approach, but there are subtle differences between the two approaches. Besides, the fact that this prediction approach applies an ontology for the item-based CF technique serves as further evidence of the superiority of user-based and item-based + EM + SVD + ontology. In addition, the authors stated that from its passing results, the throughput of item-based and user-based + EM + SVD is higher than the proposed approach since the proposed approach does not apply SVD, despite the fact that item-based + EM + SVD + ontology has somewhat better prediction accuracy than item-based and user-based + EM + SVD. Regarding accuracy measures, especially decision support metrics will perform a vital role in multi-criteria recommender evaluations. Precision measures the fraction of the item associated with the received result. Instead, Recall measures the fraction of the related retrieved items. Due to Recall increases as the total amount of items retrieved increases and Precision generally decreases with bigger sizes of results, these two metrics should be applied together. Equation (9) is the formula for Precision, and Equation (10) is the formula for Recall.

$$Precision = \frac{TR}{TR+FR} \quad (9)$$

$$Recall = \frac{TR}{TR+FN} \quad (10)$$

Where:

- FN is the total amount of false not related predictions.
- TR is the total amount of true related predictions.
- FR is the total amount of false related predictions.

Apart from this, the F-measure, which computes the mean of Recall and Precision, is the metric that considers these two values. The impact of one of the two can be given a weight by applying the β, where β is greater than 1 elevates Precision in importance and β is smaller than 1 elevates Recall in importance. Hence, a balanced F-measure is defined as β is equal to 1. Equation (11) is the formula of F-measure.

$$F1 = \frac{(1+\beta^2)\cdot precison\cdot recall}{\beta^2\cdot precision+recall} \quad (11)$$

The F1 and Precision are computed on various Top-N numbers to assess the suggested approach with decision support accuracy metrics. In the experimental results, the authors found that the Precision collected by the new approach is comparatively high compared to the nearest neighbor algorithm. Next, the authors stated that the proposed approach outperforms the nearest neighbor algorithm across all datasets. Besides that, for F1, the authors discovered the ontology-based approach outperforms the nearest neighbour algorithm in the Top-N under consideration. The effectiveness of the proposed approach can be illustrated by the fact that, in the proposed method, the authors applied an ontology in the item-based CF technique. These findings are adequate to substantiate the authors' claim that the approach is more accurate and scalable than the nearest neighbour algorithm. Other than that, the authors also evaluated the approach with various sparsity levels and computed the average MAE. The MovieLens dataset has a sparsity level of 93.7%. Furthermore, the Yahoo! Webscope R4 dataset has a sparsity level of 99.8%. Therefore, for the Yahoo! Webscope R4 and MovieLens datasets, the authors created six datasets with various levels of sparsity. Besides, the authors also applied the approach to datasets with these levels of sparsity and compared the outcomes with those of other recommendation algorithms. In the experimental results, they observed that for all sparsity levels of the dataset, the values of MAE for two approaches, namely item-based + EM + SVD + ontology and item-based and user-based + EM + SVD + ontology, are lower than item-based and user-based + EM + SVD and the nearest neighbor. Moreover, in comparison with the other approaches, the authors presented that the increase rate of the MAE of the nearest neighbor is quite high. Besides that, the outcomes demonstrated that for more sparse datasets, the approaches using ontology have superior prediction accuracy when compared to other approaches. The reason is that approaches using ontologies are more efficient at addressing sparsity problems and thus more accurate. In addition, Recall, Precision, F1, and MAE results demonstrated that CF recommender systems can effectively enhance sparsity and scalability difficulties by employing ontology in conjunction with dimensionality reduction and clustering techniques.

In 2018, El-Deen *et al.* [35] presented a suggested framework of a personalized recommendation system to online shopping by employing semantic web technology and data mining. The best classifier for categorizing individuals based on characteristics, preferences, and personal data is discovered using the data mining technique, thereby providing users with accurate recommendations based on ontology base knowledge. Besides, the dataset was subjected to the data mining phase, which involved the application of various classification algorithms for extracting and building user access sequences. This research seeks to enhance conventional recommendation systems by creating a users' information ontology for tailored recommendations and integrating user information from social networks. Hence, the authors proposed a user profile ontology and items ontology-based semantic recommender system framework. The suggested model calls for the use of data mining techniques to understand which classification algorithm is suitable for the study of the user data, which will next be applied when creating the ontology of user profile. After that, a semantic application is built and the system is integrated to check the recommendation model accuracy after validating all techniques developed. Moreover, comparative research is created to demonstrate the best classifier algorithm applied for the dataset by assessing its performance parameters in order to obtain the best classifier for the attitude and behavior of online shoppers on the basis of the collected dataset. In the experimental results, the analysis of the findings indicated that the decision table classifier provides the maximum accuracy. Clustering and a simple shopping cart gave the lowest accuracy. Apart from this, to facilitate online users to find products, the best classifier (decision table) analysis will be aided and applied in constructing an ontology model. Furthermore, the authors applied the data mining technique of the suggested model to the experimental results of the first stage. The analysis of user data using various classification algorithms revealed that the decision table algorithm provided the greatest True Positive Rate (TPR) (0.871), making it the best classifier algorithm available for developing a user profile recommender model.

In 2019, Guia *et al.* [36] suggested a hybrid method that combines the simplicity of the most popular algorithms in CF, the K-Nearest Neighbor algorithm (KNN), with the effectiveness of ontology-based recommendation systems in e-commerce. User, Person, Product, and Neighbour are the four major classes that the authors fundamentally developed. The knowledge and specifics about each active user and the relationship between them are represented and modeled using these classes. Moreover, to suggest products to the active users, the authors initially utilized the KNN algorithm to discover the nearest neighbors. Then the authors re-applied the KNN algorithm to identify the nearest products to suggest to active users. Furthermore, in addition to users with the same interests as the active users, the proposed recommender system also acquires knowledge about users, their neighbors, the items they purchase, and the relationships among them. Besides, the authors mentioned that their proposal would increase the number of products from categories active users are recommended but have not yet purchased. They also mentioned that the suggested system is scalable, which implies that even as the demand grows, it keeps up a good performance. Figure 6 shows an overview of the suggested hybrid ontology-based recommender system. It begins with the creation of a user profile that includes, among other attributes, the categories and corresponding products that the user has purchased.

Next, finding other users (neighbors) who have purchased at least one similar product and choosing only users (neighbours) who have also purchased other products while respecting the aforementioned requirements is necessary to generate product recommendations to a user. These are the products that are likely to be recommended to users. Furthermore, the authors utilized the KNN algorithm to discover users' the nearest neighbors after creating the profiles of the users and their neighbors. This process allows to extraction of products that the customer has not bought. The final stage is to utilize the KNN algorithm to locate the nearest items based on the parameters like the overall rating provided by specific neighbors, and the price of the item to determine whether it falls within the range of the smallest and greatest value purchased by the active user. Besides that, the authors also included textual reviews for each product as it is a great resource for the extraction of important information.

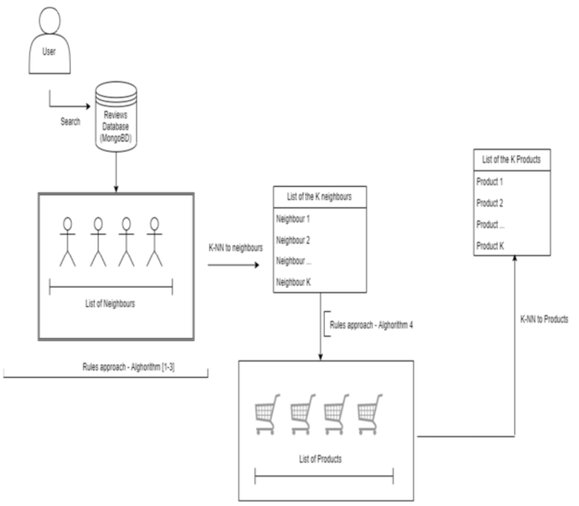

Fig. 6. Hybrid Ontology-based Recommender System by Guia *et al.* [36].

In the experimental assessment, the authors assessed the effectiveness of the suggested hybrid ontology-based recommender system using two evaluation approaches, which include the KNN algorithm's performance time when applied to neighbours and subsequently to items and the quality of the items suggested to the active users. In the experimental results, the authors demonstrated that the suggested hybrid ontology-based recommendation system produces higher quality recommendations than CF. The major improvements are validated by the outcomes of products that, despite falling under categories that users are unaware of, still fit their interests and are suggested. Moreover, the proposed approach chooses the active user's k-nearest neighbours, which leads to a limited quantity of products to which KNN will be utilized. In the comparison with the CF version, which offers a vast array of items to which KNN will be employed, it does, however, satisfy the preferences of active users because most of the time, those products fall into categories that the active user is already familiar with. Apart from this, the authors also stated that the experimental assessment illustrates that the proposed hybrid ontology-based recommendation system, in contrast to the CF version, can suggest products to the active user that have obtained greater average overall ratings.

In 2020, Garcia-Sanchez *et al.* [9] proposed an OB advertisement recommendation system that utilizes user-generated data from social networking sites. This method is validated by sharing an ontology model that can describe advertisement content and user profiles. Besides that, vectors created by Natural Language Processing (NLP) techniques that glean ontology entities from the textual content represent both advertisements and users. Moreover, the framework presented in their paper utilizes ontologies to model user preferences and the primary characteristics of advertisements semantically. Furthermore, the discovery of appropriate matches between advertisements and users who may discover those advertisements interesting is greatly aided by the application of this shared model. The framework proposed leverages NLP tools to automatically process textual content related to advertisements and users, and then produce vectors representing advertisement and user profiles based on domain ontology. Then, user profiles are dynamically updated when users click on displayed advertisements, post comments on social networking sites, and create new friendship links with other users in the network. In order to rank the matches of advertisements to respective users, a similarity measure that analyzes advertising vectors and users is finally used, and diversity is advocated by recommending hitherto less advocated advertisements. Figure 7 illustrates the architecture of the proposed system.

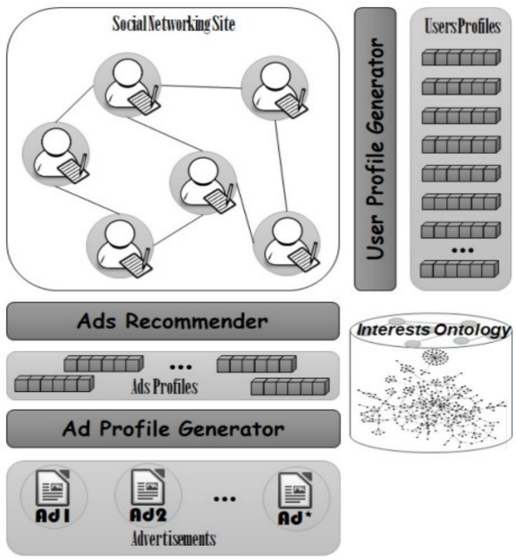

Fig. 7. Proposed framework by Garcia-Sanchez [9].

Based on Fig.7, they demonstrated that the user profile generator, the advertisement profile generator, the interest ontology, and the advertisement recommender are the four primary components of the proposed system. The system's inputs are a group of interconnected users interacting on the social media network platform and a collection of advertisements. Its outputs are the most appropriate user advertisement recommendation. The system performs, in essence, as follows. NLP tools examine the textual descriptions of advertisements, emphasizing preference ontology elements that match in the texts. Then, the advertisement profile generator represents each advertisement as a vector. Each dimension represents a different domain ontology concept and its value is determined by how frequently it appears in the advertisement's descriptive text. Similarly, the user profile generator leverages the interested ontology and produces a vector for each user with the same dimensions as the notions in the domain ontology. Users' registration information is utilized to generate the initial version of the vector in this instance, and the vector is updated as users interact with the social networking site or react to the recommended advertisements. Last but not least, the advertisement recommender generates recommendations based on the similarity found between the vectors representing the user and the vector connected to the advertisements, as well as the quantity of times each advertisement has been displayed. Furthermore, the metrics commonly applied to assess the quality of recommendation system are mainly divided into two categories: performance and predictive accuracy. In order to confirm that the recommender system correctly forecasts that users will click on the recommended advertisements, the authors emphasise the first category of metrics, especially those that can quantify click prediction accuracy. Table III shows the four possible outcomes of recommended advertisements to users.

Table III. Possible results of Recommended Advertisements to users [9].

	Recommended	Not recommended
Interested	True Positive (tp)	False Negative (fn)
Not interested	False Positive (fp)	True Negative (tn)

The authors claimed that the F-measure, Recall, and Precision metrics are ideally suited for the quality evaluation of predictions and that these values significantly help in the computation of these metrics. Precision and Recall are also known as Positive Predictive Value (PPV) and TPR or Sensitivity, respectively, in this context. The percentage of suggested products that are actually related to the users is known as Precision. Equation (12) shows the formula to get the Precision value.

$$Precision = \frac{Correctly\ recommended\ ads}{Total\ recommended\ ads} = \frac{tp}{tp+fp} \quad (12)$$

Where:

- tp donates the advertisements correctly suggested to a target user.
- $tp + fp$ consists of all the advertisements recommended to the target user.

Next, the Recall gauges the capacity of the recommendation system to recommend relevant products to users, and Recall is computed as the percentage of related advertisements that the system truly recommends. Equation (13) shows the formula to get the Recall value.

$$Recall = \frac{Correctly\ recommended\ ads}{Total\ relevant\ ads} = \frac{tp}{tp+fn} \quad (13)$$

Where:

- $tp + fn$ consists of all the advertisements of interest to a target user.

Besides, the harmonic mean of Recall and Precision is known as F-measure. Equation (14) shows the formula to compute the F1 value.

$$F - measure = 2 * \frac{Precision*Recall}{Precision+Recall} \quad (14)$$

Moreover, in this research, the authors applied the Mean Average Precision (MAP), which is especially suitable for evaluating recommendation engines while considering recommendations as a ranking problem. The MAP value measures the ability of the recommendation system to rank products. A high MAP score indicates that the system will offer related recommendations in the first few suggested advertisements to grab the attention of users. The authors stated that before computing the MAP value, it is required to establish the "Precision at cut off k" (P(k)), which is just the Precision computed by taking into account only a portion of the sorted list of suggestions from rank 1 to k. Equation (15) shows that formula to get the Average Precision (AP@k). In addition, Equation (16) is the formula to compute the MAP@k.

$$AP@k = \frac{1}{\min(m,k)} * \sum_{i=1}^{k} (P(i) * rel(i)) \quad (15)$$

Where:

- k is the total amount of advertisements the system is requested to suggest.
- m is the total amount of related advertisements in the advertising space.
- rel (i) = 1 if the ith advertisements are ralated and rel (i) = 0 if the ith advertisements is not relevant.

$$MAP@k = \frac{1}{|U|} * \sum_{u=1}^{|U|} (AP@k)_u$$

$$= \frac{1}{|U|} * \sum_{u=1}^{|U|} \left(\frac{1}{m} * \sum_{i=1}^{k} \left(P_u(i) * rel_u(i)\right)\right) \quad (16)$$

Where:

- $|U|$ indicates all the users participated in the experiment.

In the experimental result, extensive validation testing in a simulated environment, the advertisement framework achieved a Mean Average Precision at 3 (MAP@3) of 85.6%, a Recall value of 81%, a Precision value of 77.5%, and an aggregated F-measure of 79.2%.

In 2020, Nasir and Ezeife [37] proposed the semantics embedded sequential recommendation for e-commerce products (SEMSRec). The proposed SEMSRec system integrates the semantic information of e-commerce products and sequential information extracted from buying records of users into various phases of the recommendation process, which include preprocessing, pattern mining, and recommendation, to calculate similarities of items for personalized recommendations without applying purchase frequency or item ratings. This is achieved by using the prod2vec model to learn the semantic commonalities between items from the users' purchase records, exploiting this information to mine sequential purchase patterns that are rich in semantics, and, prior to applying item-based CF, adding semantic and sequential information about product purchases to the matrix of items. Therefore, SEMSRec can deliver Top-K tailored suggestions according to the semantic similarities between items with no user ratings of items. Furthermore, the proposed model is assessed against various metrics, such as Recall@K, Normalized Discounted Cumulative Gain (NDCG@K), Precision@K, MRR, and Hitrate@K. In the experimental results, the authors presented that the publicity available e-commerce dataset shows that SEMSRec gives more related suggestions than other existing recommendation approaches and exhibits enhanced performance in generating personalized recommendations.

Besides that, Kartheek and Sajeev [38] developed a SB recommender system through link prediction in knowledge graphs. The authors applied graph embedding techniques to extract semantics for explainable suggestions. The suggested approach is proven to be effective by constructing a knowledge graph with the MovieLens dataset. Besides that, the knowledge graph's missing links are predicted by the recommendation engine. The authors generated a knowledge base to implement effective link prediction from knowledge graphs via graph embedding. The knowledge base is applied for information retrieval, text comprehension, and query resolution. A knowledge base is a collection of triples (h, r, and t), where r relation of (h, t) denotes the representation of multi-relational data, h denotes the head entity, and t denotes the tail entity. Besides, relational data are mostly found in features and user-item interaction data. The formation of triples is aided by building a knowledge base schema from relational data. In addition, they stated that feature information such as user meta-information (gender, age, location, and occupation) can create triplets by taking r relation (gender, age, location, and occupation), t tail as the value of the feature information, and h as the user identity. In addition, a scoring function in the form of triples is used to represent the facts in the knowledge base. The scoring function provides a triplet (h, r, t) a correctness score value, and numerous triplets create a knowledge base, where each triplet is a fact. Furthermore, a loss function and a negative generator are applied to increase the correctness and optimisation score. Furthermore, by ranking the scores, the better the true triplets are ranked, the better the model is trained. Hence, to assess the performance of the suggested models, Mean Rank (MR), Hits@N, and Mean Reciprocal Rank (MRR) are applied. MR represents the arithmetic mean of all sorted score ranks, the results are better when the MR value is lower. MR value ranges between 0 and ∞ is sensitive to outliers. Equation (17) shows the formula to compute the MR value.

$$MR = \frac{1}{|N|} \sum_{i=1}^{N} R_i \tag{17}$$

Next, MRR, which is the reciprocal of the harmonic mean of the rank, is the mean of the reciprocal score rank. The MRR value ranges between 0 and 1, the results are better when the MRR value is higher. Equation (18) shows the formula to compute the MRR value.

$$MRR = \frac{1}{|N|} \sum_{i=1}^{N} \frac{1}{R_i} \tag{18}$$

Moreover, the percentage of predictions with ranks that are either equal to or below a given threshold value is known as hits@N. Hits@N value ranges between 0 and 1, the results are better when the value is nearer to 1. Equation (19) shows the formula to calculate the Hits@N value.

$$Hits@N = \sum_{i=1}^{N} = 1 \; if \; R_i \leq N \tag{19}$$

In the experimental results, the authors observed that the factorization-based embedding approaches, namely HolE and DistMult generate better semantic suggestions on the basis of Hits@N. The recommendation engine suggested is suitable for multi-directional and alleviates the multi-relation problems. In addition to this, it is able to give recommendations for users and items.

According to Hanafi [39], research is carried out to improve the rating prediction for e-commerce recommender systems. He proposed combining a Stack Denoising Auto Encoder (SDAE), an attention mechanism designed to improve the comprehension representation of product review documents, and a Probabilistic Matric Factorization (PMF)-based matrix factorization to generate rating predictions. Moreover, in previous work on most models utilizing Convolutional Neural Network (CNN) and Long Short Term Memory (LSTM), the attention mechanism is responsible for improving the presentation of product documents. The attention mechanism considers implementing the seq2seq feature. On the other hand, the seq2seq feature is in charge of improving the understanding of product documents from a contextual perspective to assist PMF in providing rating predictions. Furthermore, the SDAE and PNF combined attention mechanism model is used in ML.

1M, (MovieLens dataset). In accordance with the experiment reports and comparisons, the attention mechanism successfully produces rating predictions with excellent results. Based on the RMSE evaluation metrics, attention obtained a better performance of 1.6% than DDL-PMF on average, 3% higher than PHD_MF on average, and 8% higher than traditional PMF on average. Besides that, the efficiency of this model is significantly influenced by how engaging item document enhancement using an attention mechanism works. Next, the second experiment illustrated how to use the attention mechanism in large datasets (ML.10M) containing 10 million ratings, successfully improving the accuracy of rating prediction by 2.5% on average compared to the prior best operation using DDL-PMF, and achieving an average of 8% higher than the PMF model. In addition, in order to accomplish training convergence, the attention mechanism model also obtains low repetition. The authors also argued that item document representation improvement according to user information representation and attention becomes an important impact on performance outcomes.

VI. DISCUSSION AND WAY FORWARD

This section will summarise the related works that have been reviewed in the previous section. Table IV demonstrates the high-level summary of related research works, including the finding, dataset employed and evaluation metrics used in the respective system. From the review, the hybrid-based recommender system has a high potential to be an optimal recommender system since it is able to overcome one system weaknesses with other system strength.

Table IV. Summary of prior related research works.

References & Titles	Findings & Datasets	Evaluation Metrics
Elgohary *et al.* [31] Wiki-Rec: A Semantic-Based Recommendation System Using Wikipedia as an Ontology	Finding: In this research, a SB recommender system with the application of Wikipedia as an ontology is proposed. In order to address the issues of applying conventional ontologies for text analysis in text-based recommender system, the authors applied Wikipedia as an ontology. Hence, a complete system model is presented that combines SB analysis with collaborative through a content recommender system. Dataset: Not available.	A benchmark of 50 documents with a "human-judged" inter-document similarity matrix is applied to assess the semantic annotation component and to assess the ESA model. The accuracy of the model is demonstrated by the correlation between the outcoming similarity matrix of the suggested model and the "human-judged" similarity matrix. In the experimental results, the authors observed that the suggested model attains a greater correlation than the greatest value obtained by the ESA model. Besides, the authors also observed that weighting concepts in accordance with their position in the category graph, as suggested, attains better outcomes than the spreading activation with no reweighting.
Martinez-Cruz *et al.* [32] A Model to Represent Users Trust in Recommender Systems Using Ontologies and Fuzzy Linguistic Modeling	Finding: In this research, a recommender system by employing ontologies to enhance the user profiles representation is developed. The authors took into account individuals who each user may trust rather than those with comparable rating histories. Besides that, the authors presented their ontology and developed an approach for aggregating trust information obtained in the trust ontology and updating the profiles of user in response to comments. Dataset: Epinions dataset. Source: http://www.trustlet.org/epinions.html	Cross-validation is performed for the purpose of getting a value that can be used to analyze the accuracy of the various algorithms that are used to predict ratings, and thus there is a metric to contrast their performance. The authors applied MAE and coverage metrics to evaluate the proposed approach. Next, in the field of recommendation systems, the assessments of various methods is performed based on both online and offline tests.
Badriyah *et al.* [33] A Hybrid Recommendation System for E-Commerce based on Product Description and User Profile	Finding: In this research, a hybrid recommender system for e-commerce that applies the CB filtering technique and CF technique, which calculates the similarities of user-profiles and product descriptions is developed. The authors applied the Text TF-IDF approach, which is part of the CB filtering techniques, to automatically generate tags from item descriptions. The authors also applied the cosine similarity approach, which is commonly employed in CF techniques, to integrate the product profile, which takes the product tags' form, with the user profile. Dataset: MovieLens dataset. Source: https://grouplens.org/datasets/movielens/	The performance analysis of the suggestions is measured by computing the Recall and Precision values as well as the computation of the operation time. In the experimental results, it demonstrates that the recommendations have commonalities with the item description, and the user profile preference has an average Recall value of 71.47% and Precision value of 67.5%. In addition, the average execution time needed to create the tag features is 7.7 seconds, and this high execution time is due to the process of reading the product description information, which relies on the quantity of characters in the product description content.

Shruthi & Gripsy [4] An Effective Product Recommendation System for e-Commerce Website Using Hybrid Recommendation Systems	Finding: In this research, a product recommender system is suggested by combining two recommendation techniques, namely CF technique and demographic analysis, to execute effective product recommendations on e-commerce websites and improve customer satisfaction. Besides, the authors applied demographic analysis technique to understand the gender, age, and ethnic makeup of the population and how it has altered over time. Dataset: Not available.	Not available.
Nilashi *et al.* [34] A Recommender System based on Collaborative Filtering Using Ontology and Dimensionality Reduction Techniques	Finding: In this research, a recommendation system based on a CF technique that applies ontology with the help of clustering and dimensionality reduction techniques are proposed. In the CF technique, a dimensionality reduction technique, namely SVD is applied to find the most similar items and users in each item and user cluster, which can significantly enhance the recommendation approach scalability. Dataset: MovieLens dataset. Source: https://movielens.org/ Yahoo! Webscope R4 dataset. Source: https://webscope.sandbox.yahoo.com/	The time complexity (scalability) and prediction accuracy of the suggested approach is evaluated by using the two real-world movie datasets supplied by Yahoo! Webscope R4 and MovieLens. The outcomes from Precision, Recall, F1, and MAE demonstrated that CF recommender systems can effectively enhance sparsity and scalability difficulties by employing ontology in conjunction with dimensionality reduction and clustering techniques.
El-Deen *et al.* [35] Using Semantic Web Technology and Data Mining for Personalized Recommender System to Online Shopping	Finding: In this research, a personalized recommendation system to online shopping by employing semantic web technology and data mining technique is proposed. The suggested model calls for the implementation of data mining techniques to understand which classification algorithm is suitable for the study of the user data, which will next be applied when creating the ontology of user profile. Furthermore, the authors applied the data mining technique of the suggested model to the experimental results of the first stage. Dataset: Egypt dataset. Source: https://data.world/datasets/egypt	Not available.
Guia *et al.* [36] A Hybrid Ontology-Based Recommendation System in e-Commerce	Finding: In this research, a hybrid method that combines the simplicity of the most popular algorithms in CF, the KNN, with the effectiveness of ontology-based recommender systems in e-commerce is proposed. User, Person, Product, and Neighbour are the four major classes that the authors fundamentally developed. The knowledge and specifics about each active user and the relationship between them are represented and is modeled using these classes. Furthermore, in addition to users with the same interests as the active users, the proposed recommender system also acquires knowledge about users, their neighbors, the items they purchase, and the relationships among them. Dataset: Amazon dataset. Source: http://jmcauley.ucsd.edu/data/amazon/	The effectiveness of the suggested hybrid ontology-based recommender system is evaluated using two evaluation metrics, which include the KNN algorithm`s performance time when applied to neighbours and subsequently to items, as well as the quality of the items suggested to the active users. In the comparison with the CF version, which offers a vast array of items to which KNN will be employed, it does, however, satisfy the preferences of active users because most of the time, those products fall into categories that the active user is already familiar with. Apart from this, the authors also stated that the experimental assessment illustrates that the suggested hybrid ontology-based recommender system, in contrast to the CF version, can suggest products to the active user that have obtained greater average overall ratings.
García-Sanchez *et al.* [9] A Social-Semantic Recommender System for Advertisements	Finding: The framework presented in the paper utilizes ontologies to model user preferences and the primary characteristics of advertisements semantically. The framework currently leverages NLP tools to automatically process textual content related to advertisements and users, and then produce vectors representing advertisement and user profiles based on domain ontology. A similarity measure that analyzes advertising vectors and users to be finally used, and diversity is advocated by recommending hitherto less	The metrics commonly applied to assess the quality of recommendation system are mainly divided into two categories, such as performance metrics and predictive accuracy metrics. In order to confirm that the recommendation system correctly forecasts that users will click on the recommended advertisements, the authors emphasise the first category of metrics, especially those that can quantify click prediction accuracy. The authors applied the F-measure, Recall, and Precision metrics for the quality assessment of

	advocated advertisements. Dataset: Not available.	predictions.
Nasir & Ezeife [37] Semantics Embedded Sequential Recommendation for E-commerce Products (SEMSRec)	Finding: The proposed SEMSRec system integrates the semantic information of e-commerce products and sequential information extracted from buying records of users into various phases of the recommendation process to calculate similarities of items for personalized recommendations without applying purchase frequency or item ratings. Therefore, SEMSRec can deliver Top-K tailored suggestions according to the semantic similarities between items with no user ratings of items. Dataset: Online Retail dataset. Source: https://archive.ics.uci.edu/ml/datasets/online+retail Amazon dataset. Source: http://jmcauley.ucsd.edu/data/amazon/	The effectiveness of the suggested model is assessed against various metrics, such as Recall@K, Precision@K, Normalized Discounted Cumulative Gain (NDCG@K), MRR, and Hitrate@K. In the experimental results, the authors presented that the publicity available e-commerce dataset shows that SEMSRec gives more related suggestions than other existing recommendation approaches and exhibits enhanced performance in generating personalized recommendations.
Kartheek & Sajeev [38] Building Semantic Based Recommender System Using Knowledge Graph Embedding	Finding: In this research, a SB recommender system through link prediction in knowledge graphs is developed. The authors applied graph embedding techniques to extract semantics for explainable recommendations and the suggested approach is proven to be effective by constructing a knowledge graph with the MovieLens dataset. In order to increase the correctness and optimization score, a loss function and a negative generator are applied. Dataset: Movielens dataset. Source: https://movielens.org/	The performance of the suggested models is assessed by applying the MR, MRR, and Hits@N. In the experimental results, the authors observed that the factorization-based embedding approaches, namely HolE and DistMult generate better semantic recommendations based on Hits@N.
Hanafi [39] Enhance Rating Prediction for E-commerce Recommender System Using Hybridization of SDAE, Attention Mechanism and Probabilistic Matrix Factorization	Finding: The aim of this research is to enhance the rating prediction for e-commerce recommendation systems. The authors considered combining a SDAE, an attention mechanism designed to improve the comprehension representation of product review document and a PMF-based matrix factorization to generate rating predictions. The attention mechanism considers implementing the seq2seq feature. On the other hand, the seq2seq feature is in charge of improving the understanding of product documents from a contextual perspective to assist PMF in providing rating predictions. In accordance with the experiment reports and comparisons, the attention mechanism successfully produces rating predictions with excellent results. Next, the second experiment illustrated how to use the attention mechanism in large datasets (ML.10M) containing 10 million ratings, successfully improving the accuracy of rating prediction. Dataset: Movielens dataset. (Rating representation) Source: https://movielens.org/ Amazon dataset. (Product review document representation) Source: http://jmcauley.ucsd.edu/data/amazon/	The evaluation metric applied in this research is RMSE. Based on the RMSE evaluation metrics, attention obtained a better performance of 1.6% than DDL-PMF on average, 3% higher than PHD_MF on average, and 8% higher than traditional PMF on overage. The efficiency of this model is significantly influenced by how engaging item document enhancement using an attention mechanism works.

First of all, according to Elgohary *et al.* [31], a SB recommender system with the application of Wikipedia as an ontology is proposed. The accuracy of the proposed recommendation system under cold-start conditions are near enough to normal circumstances. This demonstrates the advantage of enhancing the concept of user profiles with the proposed improved semantic annotation model, which can achieve more accurate analysis than earlier models. A hybrid text-based recommendation system incorporates this model. On a benchmark data set, the improved semantic analysis model was demonstrated to be effective at addressing some of the shortcomings of the recommendation systems. Additionally, the accuracy of the recommendations is provided, and certain prior limitations are removed.

Next, based on Martinez-Cruz *et al.* [32], a recommender system by employing ontologies to enhance the user profiles representation is developed. Data are kept in database to maintain the effectiveness of the system. Besides that, in the process of recommendation generation, the authors took into account individuals who each user may trust rather than those with comparable rating histories. Moreover, the offline testing method is employed to validate the proposed recommender system instead of online testing, which need the system to be fully operational to gather large amounts of data, making the validation process costly and time-consuming.

Apart from this, according to Badriyah *et al.* [33], a hybrid recommender system for e-commerce is developed. The advantage of identifying product profiles by automatically creating tags is that the recommendation process becomes more dynamic and efficient. It is efficient as the tag of the product does not require manual input by the administrator, and it is more dynamic as the collected tag results are adjusted to the product description's content. The labels will also alter automatically if the description of the item is changed. Besides that, using product characteristics as labels is more descriptive than identifying items on the basis of the item specification, year of manufacture, and category, as is commonly used in CB filtering techniques. Furthermore, the limitation of the proposed recommender system is the high execution time of program during the process of generating recommendations and creating product characteristics in the form of tags, and the users are not inconvenienced by the lengthy execution time.

Moreover, based on Shruthi and Gripsy [4], a product recommender system is suggested by combining two recommendation techniques, namely CF technique and demographic analysis. The suggested system presented the effectiveness of the proposed model, taking into account reordering and demographic methods as well as intrapersonal and interpersonal heuristics. The system also takes into account the user preferences' independence in the field of online shopping. This implies that it can, to a certain extent, make product recommendations according to the user interests, which also takes advantage of the user associations with products, especially for the migration of current and more anticipated users. Furthermore, customers can be guaranteed greater satisfaction with the proposed work since related products are suggested as soon as a customer chooses a product to buy. This is because the recommendation algorithm uses a variety of techniques to locate related products. Additionally, the availability problem has been fixed, which boosts consumer satisfaction because there will never again be "out of stock" problem.

Besides that, according to Nilashi *et al.* [34], a recommendation system based on a CF technique that applies ontology with the help of clustering and dimensionality reduction techniques are proposed. The proposed recommender system based on a CF technique that applies ontology with the help of clustering and dimensionality reduction techniques can address two major deficiencies of recommendation systems in CF, such as scalability and sparsity. The outcomes also demonstrated that for more sparse datasets, the approaches using ontology have superior prediction accuracy when comparing to other approaches. The reason is that approaches using ontologies are more efficient at addressing sparsity problems and thus more accurate. Moreover, based on El-Deen *et al.* [35], a personalized recommendation system to online shopping by employing semantic web technology and data mining technique is proposed. The outcomes demonstrated that for more sparse datasets, the approaches using ontology have superior prediction accuracy when compared to other approaches. The reason is that approaches using ontologies are more efficient at addressing sparsity problems and thus more accurate.

Besides that, according to Guia *et al.* [36], a hybrid method that combines the simplicity of the most popular algorithms in CF, the KNN, with the effectiveness of ontology-based recommender systems in e-commerce is proposed. The suggested hybrid ontology-based recommender system increases the quantity of products from categories that active users are recommended but have not yet purchased. Moreover, the suggested system is scalable, which implies that even as the demand grows, it keeps up a good performance. Besides that, the suggested system produces recommendations of higher quality compared to CF. Apart from this, the authors also stated that the experimental assessment illustrates that the proposed hybrid ontology-based recommendation system, in contrast to the CF version, can suggest products to the active user that have obtained greater average overall ratings. Furthermore, the suggested hybrid ontology-based recommender system costumes too long time to use the KNN algorithm to discover the k-nearest products.

In addition, according to García-Sánchez *et al.* [9], an OB advertisement recommendation system utilizes user-generated data from social networking sites is suggested. The proposed semantic approach for advertisement recommendation can deal with the massive and diverse amount of data. Besides that, the suggested social recommendation system is built on a hybrid technique that combines CB, CF and knowledge-based recommender system filtering techniques. This method avoids the conventional issues that degrade recommender system performance, such as the diversity, sparsity, and cold-start problem. Besides that, the limitation of the proposed semantic approach for advertisement recommendation is that some information items that are naturally accessible on social media platforms are not yet fully utilised in the proposal, which could aid increase the reliability and accuracy of the proposed framework.

Moreover, based on Nasir and Ezeife [37], a SEMSRec system is proposed. In the experimental results, the authors presented that the publicity available e-commerce dataset shows that SEMSRec gives more related suggestions than other existing recommendation approaches and exhibits enhanced performance in generating personalized recommendations. According to Kartheek and Sajeev [38], a SB recommender system through link prediction in knowledge graphs is developed. The recommendation engine suggested is suitable for multi-directional and alleviates the multi-relation problems. In addition to this, it is able to give recommendations for users and items. Furthermore, based on Hanafi [39], the aim of this research is to enhance the rating prediction for e-commerce recommendation systems. In accordance with the experiment reports and comparisons, the attention mechanism successfully produces rating predictions with excellent results. Besides that, in order to accomplish training convergence, the attention mechanism model also obtains low repetition.

Some insights that we can gain is that HB recommender system has great potential compared to other recommender systems that employ only one recommender system filtering technique. HB filtering technique is a technique that utilizes two or more recommendation algorithms or techniques in a single recommender system to benefit from the strengths of multiple recommendation techniques and mitigates the drawbacks of any single recommendation technique, such as cold start problem, data sparsity problem, scalability, or more. Hence, HB filtering technique can achieve better system optimization than other recommender system filtering technique by generating more accurate and reliable recommendations to users.

VII. CONCLUSION AND FUTURE WORK

In this paper, the overview of the recommender system has been covered. Next, the phases in the recommendation process have been discussed, which include the information collection phase, the learning phase, and the prediction or recommendation process. Besides that, the classification of the recommender system filtering technique has been explored in detail, including the CB filtering technique, CF technique, HB filtering technique, and SB filtering technique. Other than that, related research works in this field have been reviewed. From the review, we observed that HB recommender system with the combination of two or more recommendation algorithms, has great potential. We will conduct a systematic review focusing on the HB recommender system in our future work. In addition, we will implement the HB recommender system in an e-commerce dataset.

REFERENCES

[1] Q. Zhang, J. Lu and Y. Jin, "Artificial Intelligence in Recommender Systems," *Complex & Intellig. Sys.*, vol. 7(1), pp. 439–457, 2021.

[2] X. Hu, W. Zhu and Q. Li, "HCRS: A Hybrid Clothes Recommender System based on User Ratings and Product Features," *Int. Conf. on Management of e-Commerce and e-Government*, pp. 270-274, 2014.

[3] C. Obeid, L. Lahoud, H. el Khoury and P.A. Champin, "Ontology-based Recommender System in Higher Education", *Companion of the The Web Conf.*, pp. 1031–1034, 2018.

[4] S. Shruthi and V. J. Gripsy, "An Effective Product Recommendation System for E-Commerce Website Using Hybrid Recommendation Systems," *Int. J. of Comp. Sci. & Comm.*, vol. 8(2), pp. 81-88, 2017.

[5] H. Kaur and G. Bathla, "Techniques of Recommender System," *Int. J. of Innov. Techn. and Explor. Eng.*, vol. 8(9S), pp. 373–379, 2019.

[6] F. O. Isinkaye, Y. O. Folajimi and B. A. Ojokoh, "Recommendation Systems: Principles, Methods and Evaluation," *Egyptian Inform. J.*, vol. 16(3), pp. 261–273, 2015.

[7] C. Yang, X. Yu, Y. Liu, Y. Nie and Y. Wang, "Collaborative Filtering with Weighted Opinion Aspects", *Neurocomputing*, vol. 210, pp. 185–196, 2019.

[8] F. Ricci, L. Rokach, B. Shapira and P. B. Kantor, "Recommender Systems Handbook," *Springer US*, 2019.

[9] F. García-Sánchez, R. Colomo-Palacios and R. Valencia-García, "A Social-semantic Recommender System for Advertisements," *Inform. Process. & Manage.,* vol. 57(2), 102153, 2020.

[10] J. L. Herlocker, J. A. Konstan, L. G. Terveen and J. T. Riedl, "Evaluating Collaborative Filtering Recommender Systems," *ACM Trans. on Inform. Sys.,* vol. 22(1), pp. 5–53, 2004.

[11] J. Li, K. Zhang, X. Yang, P. Wei, J. Wang, K. Mitra and R. Ranjan, "Category Preferred Canopy–K-means Based Collaborative Filtering Algorithm," *Future Generation Comp. Sys.*, vol. 93, pp. 1046–1054, 2019.

[12] A. Patel, A. Thakkar, N. Bhatt and P. Prajapati, "Survey and Evolution Study Focusing Comparative Analysis and Future Research Direction in the Field of Recommendation System Specific to Collaborative Filtering Approach," *In Satapathy, S., Joshi, A. (eds) Inform. and Comm. Techn. for Intellig. Sys.. Smart Innov., Sys. and Techn.*, vol 106. Springer, 2019.

[13] N. Mustafa, A.O. Ibrahim, A. Ahmed and A. Abdullah, "Collaborative Filtering: Techniques and Applications," *Int. Conf. on Comm., Control, Comput. and Electronics Eng.,* pp. 1–6, 2019.

[14] P. B. Thorat, R. M. Goudar and S. Barve, "Survey on Collaborative Filtering, Content-based Filtering and Hybrid Recommendation System," *Int. J. of Comput. Appl.*, vol. 110(4), pp. 31–36, 2019.

[15] F. Ortega, R. Hurtado, J. Bobadilla and R. Bojorque, "Recommendation to Groups of Users Using the Singularities Concept," *IEEE Access*, vol. 6, pp. 39745–39761, 2018.

[16] S. M. Al-Ghuribi and S. A. M Noah, "A Comprehensive Overview of Recommender System and Sentiment Analysis," Preprint, 2021. https://doi.org/10.48550/arXiv.2109.08794

[17] H. P. Ambulgekar, M. K. Pathak and M. B. Kokare, "A Survey on Collaborative Filtering: Tasks, Approaches and Applications," *Int. Ethic. Hacking Conf. Advanc. in Intellig. Sys. and Comp.*, vol 811, 2018.

[18] M. V. Kumar and P. N. V. S. Pavan Kumar, "A Study on Different Phases and Various Recommendation System Techniques," *Int. J. of Recent Techn. and Eng.*, vol. 7(5C), pp. 38-41, 2019

[19] M. Lagerstedt and M. Olsson, "A Hybrid Recommender System for Usage Within e-commerce," Master Thesis, Chalmers University of Technology, Sweden, 2017.

[20] F. T. A Hussien, A. M. S. Rahma and H. B. A. Wahab, "Recommendation Systems for E-commerce Systems An

Overview," *J. of Phys.: Conf. Series*, vol. 1897(1), 012024, pp. 1-14, 2021.
[21] F. Hdioud, B. Frikh and B. Ouhbi, "Multi-Criteria Recommender Systems Based on Multi-Attribute Decision Making," *Int. Conf. on Inform. Integr. and Web-Based Appl. & Serv.*, pp. 203–210, 2013.
[22] J. B. Schafer, D. Frankowski, J. Herlocker and S. Sen, "Collaborative Filtering Recommender Systems," *Lecture Notes in Comp. Sci.*, pp. 291–324, 2007.
[23] G. Ali and A. Elkorany, "Semantic-based Collaborative Filtering for Enhancing Recommendation," *Int. Conf. on Knowledge Eng. and Ontol. Develop.*, pp. 176-185, 2014.
[24] P. Sheridan, M. Onsjo, M., C. Becerra, S. Jimenez and G. Genas, "An Ontology-Based Recommender System With An Application to The Star Trek Television Franchise," *Future Internet*, vol. 11(9), pp. 182, 2019.
[25] T. R. Gruber, "A Translation Approach to Portable Ontology Specifications," *Knowledge Acquis.*, vol. 5(2), pp. 199–220, 1993.
[26] S. E. Middleton, D. D. Roure and N. R. Shadbolt, "Ontology-Based Recommender Systems," *Handbook on Ontologies*, pp. 779–796, Springer, 2009.
[27] X. Li and H. Chen, "Recommendation As Link Prediction in Bipartite Graphs: A Graph Kernel-based Machine Learning Approach," *Decis. Supp. Sys.*, vol. 54(2), pp. 880–890, 2013.
[28] S. Kamta and V. Verma, "A Survey on Graph-Based Collaborative Filtering Techniques in Recommender Systems," *Int. J. of Knowledge Based Comp. Sys.*, vol. 7(2), 2019.
[29] B. Shams and S. Haratizadeh, "Graph-based Collaborative Ranking," *Expert Sys. with Appl.*, vol. 67, pp. 59–70, 2017.
B. B. Sinha and R. Dhanalakshmi, "Evolution of Recommender Paradigm Optimization Over Time," *J. of King Saud Univ. – Comp. and Inform. Sci.*, vol. 34(4), pp. 1047–1059, 2022.
[30] A. Adala, N. Tabbane and S. Tabbane, "A Framework for Automatic Web Service Discovery Based on Semantics and NLP Techniques," *Advanc. in Multimedia*, vol. 2011, pp. 1-8, 2011.
[31] A. Elgohary, H. Nomir, I. Sabek, M. Samir, M. Badawy and N.A. Yoursi, "Wiki-rec: A Semantic-based Recommendation System Using Wikipedia As An Ontology," *Int. Conf. on Intellig. Sys. Design and Appl.*, pp. 1465–1470, 2010.
[32] C. Martinez-Cruz, C. Porcel, J. Bernabe-Moreno and E. Herrera-Viedma, "A Model to Represent Users Trust in Recommender Systems Using Ontologies and Fuzzy Linguistic Modeling," *Inform. Sci.*, vol. 311, pp. 102–118, 2015.
[33] T. Badriyah, E. T. Wijayanto, I. Syarif and P. Kristalina, "A Hybrid Recommendation System for E-commerce Based on Product Description and User Profile," *Int. Conf. on Innov. Comput. Techn.*, pp. 95–100, 2017.
[34] M. Nilashi, O. Ibrahim and K. Bagherifard, "A Recommender System Based on Collaborative Filtering Using Ontology and Dimensionality Reduction Techniques," *Expert Sys. with Appl.*, vol. 92, pp. 507–520, 2018.
[35] R. A. El-Deen, S. Morsi and N. Magdi, "Using Semantic Web Technology and Data Mining for Personalized Recommender System to Online Shopping," *Int. Conf. on Comp. and Appl.*, pp. 358–363, 2018.
[36] M. Guia, R. R. Silva and J. Bernardino, "A Hybrid Ontology-Based Recommendation System in e-Commerce," *Algorithms*, vol. 12(11), pp. 239, 2019.
[37] M. Nasir and C. I. Ezeife, "Semantics Embedded Sequential Recommendation for E-Commerce Products (SEMSRec)," *Int. Conf. on Advanc. in Social Netw. Analy. and Mining*, pp. 270–274, 2020.
[38] M. Kartheek and G. P. Sajeev, "Building Semantic Based Recommender System Using Knowledge Graph Embedding," *Int. Conf. on Image Informa. Process.,* pp. 25–29, 2020.
[39] H. Hanafi, "Enhance Rating Prediction for E-commerce Recommender System Using Hybridization of SDAE, Attention Mechanism and Probabilistic Matrix Factorization," *Int. J. of Intellig. Eng. and Sys.*, vol. *15*(5), pp. 427–438, 2022.