# Recommendation Systems: Classification, Open Issues and Recent Developments

**Deepak Vats, Avinash Sharma**

***Abstract*: When internet is overwhelmed with infinite options then you need intelligent program like recommendation system to help you dug out and prioritize relevant facts. Everyone is confronted with information flood phenomena and the recommendation engine alleviate Information flood on internet. Personalizing information, to each individual is the solution of information flood, is performed by these intelligent systems through searching web. Texts here trace out diverse characteristics concerned with recommendation system and highlight possible recommendation-methodologies capacity in this arena.**

***Index Terms*: Challenges, Classification, Collaborative Filtering, Content-Based Filtering, Evaluation Metrics, Recommendation Systems.**

## I. INTRODUCTION

The availability of vast digital information and enormous amount of visitors are the basic reason behind the problem of information overload on internet. Due to this problem, user is not able to access items (product or service) as per its interest on web. The problem has been address/resolved partially by search engines or information retrieval system like Google. The growing demand of recommender system is due to the non-existence in providing personalized information to the web user by these systems. Recommendation system is a type of web application that takes care of information overload problem [1] and present personalized items to user by filtering vast information accessible on internet as per the behaviour of user [2]. A profile of user is created which is used by recommender system to forecast whether an item is liked or preferred by user or not. These types of systems are valuable for both user and marketer. The benefit provided to the user include: finding preferred service with great efficiency, to make a purchase confidently and the ability to discover new things. The marketer benefits include: increase in users probability to purchase item or service recommended by the system, more frequent visit to the website and suggest website to other people in a friend list [3]. The transaction price to find and select a service in online environment is reduced with the help of these systems [4]. These systems are evolved with the objective to help users and increase revenues in online environment [3]. They have enhanced the quality of decisions taken by the retailer during online interaction with the website and also improve the overall process [5]. The users can search online libraries and stores beyond the catalog with the support of these systems. Therefore, there is a scope to incorporate more advanced recommendation methods within existing systems so that they can provide more relevant, concise, accurate recommendation for online retailers in e-commerce and we cannot ignore their existence. So a system that provides an effective decision making technique to online retailer or user of e-commerce website within multifaceted environment of information is called recommender or personalized system [6]. The subsequent part of the paper is arranged in a manner as: Part 2 discussed two conventionally established methodologies and their salient merits and demerits with possible hybridization are summarized, Part 3 put light on numerous challenges and issues, Part 4 summarize evaluation-metrics established by earlier researchers, Part 5 elaborates recent developments in context of techniques incorporated for creating such systems and lastly conclusion is presented.

Revised Manuscript Received on July 06, 2019

**Deepak Vats**, Research Scholar, Department of Computer Science & Enginering, Maharishi Markandeshwar (Deemed to be University), Mullana, Ambala, India. Also with
Assistant Professor, Department of Computer Applications, University Institute of Computing, Chandigarh University, Gharuan, Mohali, Punjab, India.
**Avinash Sharma**, Professor, Department of Computer Science & Enginering, Maharishi Markandeshwar (Deemed to be University), Mullana, Ambala, India.

## II. CLASSIFICATION OF RECOMMENDER SYSTEMS

The adaption of relevant and perfect recommendation methodology has great importance in a recommender system that present personalized items to user by filtering vast information accessible on internet as per the behaviour of user. Therefore, it necessitates people to study the advantages and disadvantages of various recommendation strategies which evolved in the last two decades. Anatomy of various strategies evolved in the last two decades for recommendation is depicted in Fig. 1.

### *A. Content-Based Filtering*

It is a domain specific algorithm in which emphasis is given to the study of various attributes related to items for generating relevant predictions. It is a famous technique which is incorporated into the recommender system for recommending news, movies, product and web-pages to the online visitor of a website. In this technique, the basis for recommendation is profile of the visitor on a site which is created by analyzing the items selected by him in recent past [7, 8]. All Items which are closely correlated to items within a profile are favourable for recommendation to the current visitor.


*Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication*

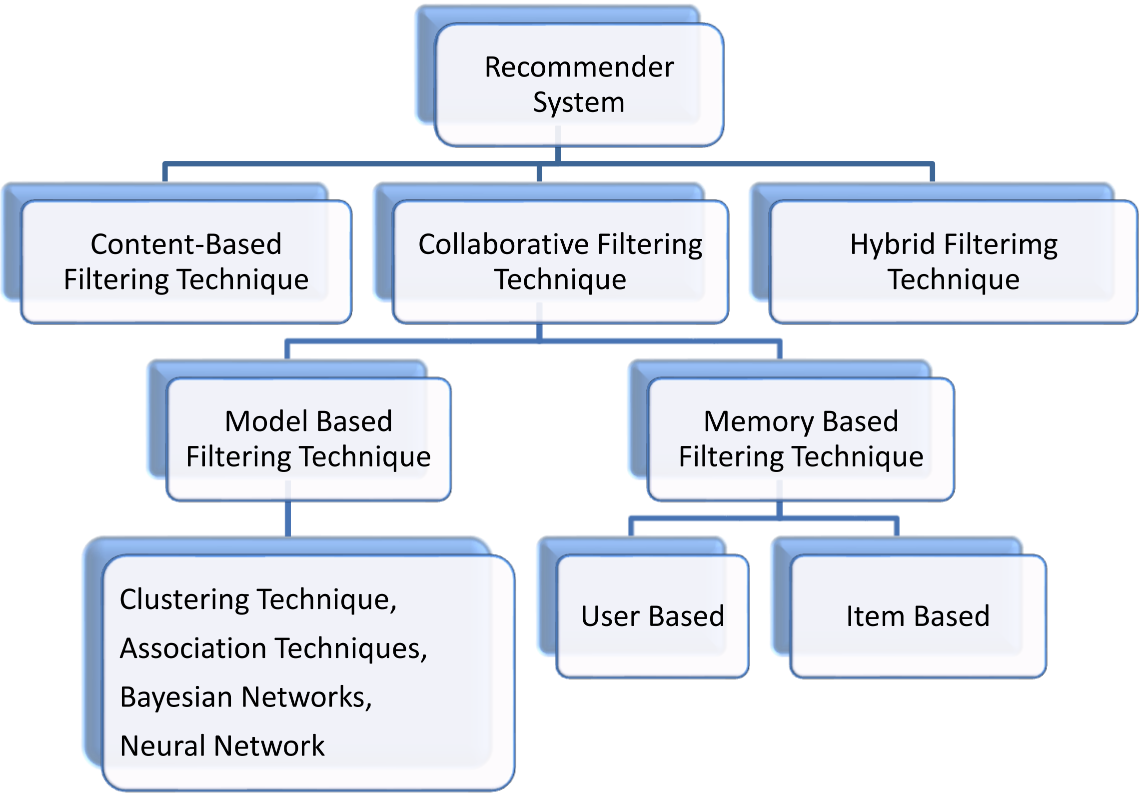


**Fig. 1 Recommendation-Technique [44]**

The recommendations are done by many models based on similarity calculation of different documents in this technique. One of the first models is TF/IDF which is a vector-space model and other used models used can include: decision tree [9], Naive Bayes Classifier [10] which is a probability based model and Neural Network [11], to represent similarity relationship among documents. The Model used for the purpose of recommendation in Content-Based Technique is built using the techniques based on Machine Leaning and Statistical Method. The recommendations in Content-Based Technique are not hampered by the other user profile because they are not evaluated on them. Also, Content-Based Technique is adjustable to the dynamic changes in the profile of user and the complete knowledge of items attributes within a profile is a major shortcoming of the method.

**i)Advantages and disadvantages of Content-Based Filtering Strategy**

The challenge pose by CF (Collaborative Filtering) is efficiently handled by Content-Based Filtering. The system based on this technique is capable to recommend items even if user rating is not available. It means accuracy of recommendation system is not affected by the fact whether the user rating were stored in the database. In case user changes its rating, these methods have the capability to accommodate change in quick time span. Recommendations can be possible in case user is not ready to share its profile and user can maintain its privacy by doing this [12]. These techniques can explain about how to suggest recommendations to the users. These techniques have shortcoming as provided by [13]. Content Based Filtering methodology has its dependency on items description. Before recommending something to the user, there is a need of metadata related to items and a properly organized summary of user profile. It is known as Limited Content Analysis. Therefore, effectiveness of such techniques is dependent on the amount of descriptive information available. Overspecialization [14] of content is other shortcoming of CBF strategy. Recommendation received by user is constrained to items contained in its profile. There are number of recommendation systems developed based on CBF like News Dude [15] to recommend news tales and LIBRA [16] to recommend books.

### *B. Collaborative Filtering*

A domain independent method used for predicting content which is not defined using metadata is known as Collaborative Filtering. The technique start with constructing the database of items preferred by users called matrix of user item and in a next step it perform match making between users having common preferences and taste by finding similarity within their profiles for the purpose of making recommendation[17]. The group of these users is known as neighbour-hood. The items rated positive by users in its neighbourhood are recommended and these recommendations include those items which they have never rated in the past. Recommendation based on model CF is referred as prediction or recommendation. Here prediction represents a numeric value ( $R_{ij}$ ) and recommendation represent top-N list for items that will be liked by the active user (see Fig. 2). Memory-Based and Model-Based are the categories of Collaborative Filtering Strategy [18, 19]. In Memory-Based method, items ranked by user in the past play an important role in identifying a neighbour who is having same taste that of user [20, 21]. Once you find user neighbour then many algorithms are there to join interests of neighbours to give recommendation. These techniques are very effective and therefore there are many applications where these techniques can be used successfully.

| | ITEM$_1$ | ITEM$_2$ | ......... | ITEM$_J$ | ITEM$_N$ |
|---|---|---|---|---|---|
| USER$_1$ | | | | | |
| USER$_2$ | | | | | |
| . | | | | | |
| USER$_I$ | | | | | |
| USER$_M$ | | | | | |

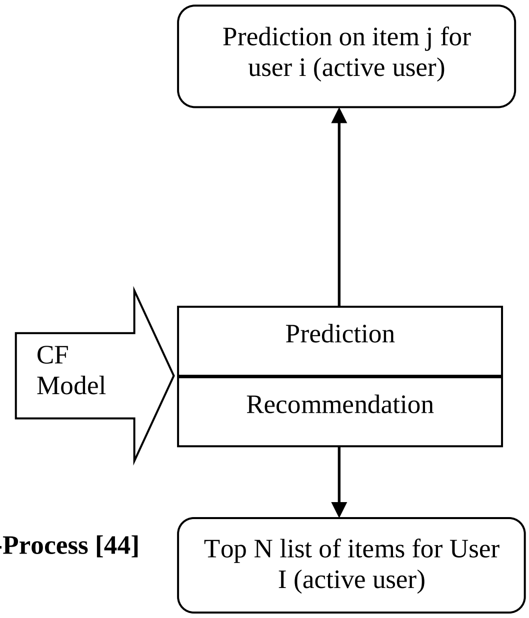


**Fig. 2 Collaborative-Filtering-Process [44]**

**i)Advantages and disadvantages of Content-Based Filtering Strategy**

CFS (Collaborative Filtering Strategy) outperforms CBF methods in a way where contents are not related to the items and also contents are hard to analyze by computer (e.g. analyzing opinion). This method is capable of making grievous recommendation. In case contents are not being there in the profile of user, still relevant items can be recommended [22]. The use of these techniques on large scale has identified some of the major drawbacks as follows:

1) The situation where system has not sufficient information related to profile of user or an item for the task of generating appropriate predictions is identified as Cold-Start Problem [23]. It is responsible for degrading the performance of recommendation system. There exists an empty profile for user or item because he never provides rating to items. Therefore, system is not aware of the taste of user.
2) This drawback is associated with the issue of information deficiency. It occurs when users give rating to only a small subset of database items. As a result there is sparsity within information related to rating of database items by different users [24, 25]. Due to this, it is difficult to find neighbours which leads to generate weak predictions.
3) Another issue concerned with these algorithms is scalability. An algorithm that forecast efficient recommendation for a limited dataset will be unable to give satisfactory result for larger dataset. Thus, it is important to embed Recommendation Algorithms that are competent of scaling with the increased dimension of dataset.

### *C. Hybrid Recommendation Filtering Techniques*

These techniques apply diverse recommendation algorithms to overcome limitations of earlier filtering algorithms and to optimize the system which use single algorithm for prediction [26, 27]. The general idea is to incorporate more than one algorithm to enhance the quality of recommendation. An enhanced model which is a combination of two or more algorithm overshadows the weakness of individual algorithm [22]. The different approaches can be merged as follows: implement each algorithm individually and then merges their result, introducing one CB (Content Based) algorithmic approach into collaborative algorithmic approach, introducing one collaborative algorithmic approach into CB algorithmic approach, creating an integrated approach where both individual approaches exist jointly.

**i) Weighted Hybrid Method**

Here the score calculation for each item is done by individual recommender mixed in hybrid model and then their results are added linearly in a formula. P Tango [28] is one such example of this type of system. It provides a weighted hybrid model that is a mixture of collaborative with content approach. Initially, equal weights are given to each algorithm and weights are readjusted as recommendations for items are approved or disapproved by user. This algorithm takes advantage of all the algorithms used in a process of recommendation.

**ii) Switched Hybrid Method**

In the recommendation process of this system, a well defined heuristic is used for swapping between different recommender involved in the model. The decision of selecting one of the available recommender is based on profile under consideration. The problems related to one type of recommender are easily handled by such systems by shifting to other recommender. These systems are highly sensitive to advantages and disadvantages of their component recommenders. The major drawback of switching hybrid is due to the complexity of recommendation algorithm that involves more number of parameters used for switching purpose [7]. DailyLearner [29] is an


*Published By:*
*Blue Eyes Intelligence Engineering*
*& Sciences Publication*

example of switched hybrid recommender where component recommenders are based on content and collaborative approaches respectively. The process of recommendation in these systems begins by picking any one of the available component recommender based the criteria of switching in the present scenario.

**iii) Cascade Hybrid Method**

The recommendation is done in stages one after the other. Firstly, one of the recommender is applied to generate a candidate list of items and then it is refined by the other recommender of the system. This hybridization is highly effective to noise. EntreeC [7] recommender system is used for providing restaurant recommendation to its users which is based on Knowledge-Based and collaborative algorithms.

**iv) Mixed Hybrid Method**

The outcome of this approach is the combination of all results generated by various component recommender used in Mixed-Hybrid instead of presenting outcome of any one recommender. For each item, several recommendations are associated by each component recommender. Here, the final result is not affected by the performance of individual algorithm. PTV [30] system is one that belongs to the category of mixed hybrids. It is a mix of two approaches: content based and collaborative. It is used to predict the list of viewing TV programs by user based on the content based and collaborative recommendation strategy. Other example of such system include: Profinder [31] and PickAFlick [32].

**v) Feature-Combination Method**

The feature generated by one particular recommender act as input for the next level recommender. Rating alike users is a characteristic of collaborative algorithm and this characteristic is applied to Case-Based Reasoner to calculate resemblance among items. A movie recommender, Ripper [33], is example of this method. It applies the rating produced by collaborative filter within Content Based system for movie prediction. The biggest advantage of these algorithms is that, they are not dependent on data of collaborative nature.

**vi) Feature Augmentation Method**

The FA (Feature Augmentation) model of hybridization is preferred over FC (Feature Combination) model because it is difficult or sometimes not feasible to define FC model using different combination of hybrids. It means FC model is more rigid than FA model. These hybrid systems are similar to Feature Combination approach based system. The method consumes rating and feature generated by previous system for recommendation task and need extra working capacity from recommendation framework. One example in this category is Libra [34] framework which provides Content Based book suggestion using data available on the amazon website with the help of Naive bayes Text Classifier. FA hybrid is more efficient than FC hybrid because they contribute fewer features in key recommender.

**vii) Meta-Level Method**

Here, model produced by one of the component recommender engine act as input to other component recommender engine. The final received model is highly informative compared to other single rating model. The solution to the problem of data sparsity is provided by ML (Meta-Level) [35] model where learned model from one algorithm is completely fed as input into second algorithm. LaboUr [36] system is based on ML hybridization.

## III. OPEN CHALLENGES AND ISSUES

This segment looks into different identified issues and challenges found in the literature of recommendation engine and narrate different techniques evolved to cope up with these issues. The issues may include: Cold Start Problem, Synonymy, Privacy, Sparsity, Shilling Attacks and Scalability.

**i)Cold-Start Issue**

In REs (Recommendation Engine), the problem jointly relates with New-User and New-Item scenario. Initially RE show rigid behaviour toward items having zero rating. It means RE can't recommend item having zero rating to its users. Therefore RE can't be able to gather rating information of New-item. It is hard to cope with New-user problem by RE. The reason may include: its inability to identify similar user or to define User-profile without knowing users past preferences. The solution to this problem includes: taking user rating for subset of items at the start, know user taste explicitly by asking questions or exploit user demographic data to understand his taste and recommend items [22].

**ii)Sparsity Issue**

The maximum users interacting with e-commerce website are not interested to confirm item rating which is responsible for sparse rating matrix. Sparse rating matrix originates this problem and decreases the opportunity of identifying a user group having alike rating. It is a famous shortcoming of the CF methodology. This matter could be vanquished by consuming extra domain related information.

**iii)Scalability Issue**

These days adaptability of same recommendation methodology with tiny and bulky dataset is a fundamental issue in REs which is called scalability issue. The task of dealing a bulky data set created through interaction of user with an e-commerce website is challenging. Some methodology is there in literature which give excellent recommendation when apply on tiny data set and give worst recommendation for bulky data set. Hence sophisticated prediction methods are requisite to defeat scalability.

**iv)Privacy Issue**

It is necessity for user or client to insert its private information within recommendation system so that they do quality personalization for users. As the RS (Recommendation System) is filled with private data that belongs to individual client of system, it gives rise to the issue of privacy in the mind of client because system has maximum knowledge about its clients. Therefore sophisticated RS are mandatory to secure this information from malware or malevolent visitors.

v)**Synonymy Issue**

Representing the same item with diverse names having same interpretation is known synonymy issue. Therefore system can't differentiate whether entries stand for alike item or diverse items. Few techniques in recommendation treat drama film and drama movie differently. The exorbitant usage of synonymy phrases degrade performance of RS belonging to Collaborative Filtering. Synonymy handling techniques include: LSI, ontology and SVD.

**vi)Shilling Attack Issue**

Sometimes malevolent competitors start feeding bogus rating for item either to enhance product popularity or degrade the product popularity in RSs. The belief of user is shattered by such activities towards diverse recommendation provided by RSs and diminishes system performance. The CF methodology is highly vulnerable or exposed to this threat. Different attack models include: bandwagon model, average model, reverse bandwagon model and random model and attacks categorization are based on attack size, motive of attack and quantity of facts vital to initiates attack.

## IV. EVALUATION OF RECOMMENDER SYSTEM

The metrics [36] evolved to gauge effectiveness of prediction engine in RS are classified in following principal classes: CAM (Classification-Accuracy Metrics), PAM (Predictive-Accuracy Metrics), NAM (Non-Accuracy Metrics) and RAM (Rank-Accuracy Metrics).

**i) Predictive-Accuracy-Metrics**

The answer to the question of closeness between rating produced by prediction engine of system and actual user rating is given by PAM. PAM is commonly preferred method of evaluation where rating is Non-Binary type. Such metrics are best suited in usage scenarios where importance of correct rating for all existing items is high. The movie predictor system, MovieLens [37], predict star rating given by user to each movie and display forecasted value in front of user. The different prime representative metrics in PAM include: Mean-Absolute-Error (MAE), Mean-Squared-Error (MSE), Root-Mean-Squared Error (RMSE) and Normalized-Mean Absolute-Error (NAME). The first representative metric signify average deviation among predicted and actual value. Its advantages includes: Firstly, easily understandable and simple calculation and secondly, nicely defined and proved properties in statistics that help in proving test of significance among difference amid MAE of different systems. The second and third metric does summation after squaring error. Here, the focus is on finding maximum errors. For Example, A single error can add one to the total of error and double error can add four to the total of error. The fourth metric is concerned with normalized MAE in relation to rating range.

**ii) Classification-Accuracy-Metrics**

They measure how frequently decisions taken by prediction engine regarding item goodness are right or wrong. Therefore they are preferable in those works whose focus is on finding good quality products where customers have shown twofold preferences. Applying them in offline test on synthetic records of data suffer from sparsity problem. Problem found in situations, when system under evaluation presents catalog containing highly ranked items. During evaluation of list, the recommended enlist can contain non-rated items. Now question is how to treat these items in evaluation process which can add bias. The different way to hold sparsity in data during evaluation phase includes: ignore non-rated items and assume somewhat fixed negative value for non-rated items. The various metrics are: Precision, Recall and ROC. Cleverdon introduced the first two metrics in 1968 to evaluate Information-retrieval system [38]. Many researchers used them to evaluate recommendation system in the year 1998 [39, 40] and 2000 [41, 42]. The first two can be applied where whole cluster of item.

**iii) Rank-Accuracy-Metrics**

These metrics estimate the capability of algorithm employed in recommendation process that how many times the recommended ordering match with the actual ordering given by user for the same item-set. Rank metric is highly suited where user get ascending recommendation list ordered by rank, in the domain where recommendations show Non-Binary User-Preferences These metrics are highly responsive in areas where person interacting to system is not interested in organization of objects beyond dual classes such as good or bad, relevant or Non-Relevant. This kind of metric can make use of relative organization or arrangement of user liking or preferred values that are independent of recommender approximated values. As an example, a system can continuously calculate lower rating for items than real preferences, but still be able to achieve the desired score in case of accurate ranking. In other words, rank metric may provide non-accurate result in view of the fact that chief item (Top ranked) is placed at last position or ranked last. Majority of the metrics like Kendall-Tau, Pearson and Spearmen rank correlation do comparison of two full ordering. Correlation defined by pearson formula give approximation about linear association within two values. It is calculated using the eq. (1) [36].

$$c = \frac{\sum(x-\bar{x})(y-\bar{y})}{n \times \mathrm{stdev}(x) \times \mathrm{stdev}(y)} \qquad (1)$$

Spearmen and Tau correlation are calculated using Eq. (2) and Eq. (3) [36]

$$\rho = \frac{\sum(u-\bar{u})(v-\bar{v})}{n \times \mathrm{stdev}(u) \times \mathrm{stdev}(v)} \qquad (2)$$

$$\mathrm{Tau} = \frac{C-D}{\mathrm{sqrt}((C+D+TR)(C+D+TP))} \quad (3)$$

The merits may include: they are known by mostly researcher and they present single numerical value to represent whole system. The demerits may include: In majority of the cases, rating is not provided for every item which means partial ranking and also recommendation check-list may have cluster of objects having equal numerical grade in random order.

In 2007 to overcome the dilemma of biasness in earlier metrics, like precision, F1-Measures and other, Powers [43] proposed: 1) informedness 2) markedness 3) Matthews metrics of correlation to deal with the phenomena of biasness and give following formulas:

$$\mathrm{Markedness} = \mathrm{precision} + \mathrm{inverseprecision} - 1$$

$$= \frac{t_p}{t_p+f_p} + \frac{t_n}{f_n+t_n} - 1 \qquad (4)$$

$$\mathrm{informedness} = \mathrm{recall} + \mathrm{inverserecall} - 1$$

$$= \frac{t_p}{t_p+f_n} + \frac{t_n}{f_p+t_n}\ 1 \qquad (5)$$

$$\text{Matthews Correlation}$$

$$= \frac{(t_p \,.\, t_n)\,(f_p \,.\, f_n)}{\sqrt{(t_p+f_n).(f_p+t_n).(t_p+f_p).(f_n+t_n)}}$$

$$= \pm\sqrt{\mathrm{informedness}\,.\,\mathrm{markedness}} \qquad (6)$$


*Published By:*
*Blue Eyes Intelligence Engineering*
*& Sciences Publication*

## V. RECENT DEVELOPMENTS

RSs are extension of earlier system used by industry to extract information from internet for commercial intention. Today, RS are heavily employed in commercial use for personal growth of business. Example includes: news, movie, document, real estate, education, restaurant, E-Commerce, tourism, hotel, books, goods RSs and many more. The list of famous companies which provide online services around the globe includes: amazon, jingdong, Alibaba, ebay, rakuten, B2W, zalando and groupon. All of them take help from engine which facilitates recommendation to realize the need of system visitors. The various construct like DL/ML (Deep/Machine Learning), WM (Web Mining) and SW (Semantic Web) mining are served as backbone in recommendation regime. The most famous RS in modern era are based on Deep/Machine Learning. Deep Learning acts as crucial weapon to promote earning and experience of many websites. Examples include: Among all visitors of Netflix, the success percentage of movies and Television programme seen by peoples using Netflix recommendation engine, is more than 80. The success rate of video recommendation proposed by the starting page of youtube is 60. The top-most companies around the globe are looking to increase quality of engine employed for recommendation by applying deep-learning methodology. Algorithms for video, google app and news recommendation are proposed in recent period by Covington, Cheng and Shumpei respectively. For detailed summarization of numerous methodologies in this category, reader can go through [45, 46]. The next era of recommendation engine belongs to semantic web. The interoperability provided by renowned standards at different stages is responsible for existence or realization of web. Web realization is based on the concept that data is consumed by humans not by machines. The first medium for exchanging data online between users is web. Semantic web is considered as new vision in relation to enlarge capability of existing internet or web which provides an immersive way to access data and its prime focus is on data not on documents. The difference lies in who will consume the available content on web: human, machine or both. The technology used to define it includes: 1) XML (Extensible-Markup- Language) 2) RDF (Resource-Description-Framework) 3) SPARQL and 4) OWL (Web-Ontology-Language). For detailed highlights of numerous methodologies, reader can go through [47]. The next series of RSs are based on web mining. The help of internet is necessary for every person whenever it looks for shopping, travelling, education, hospitality, watching movies and playing games online, buying/selling goods like books, product, home, vehicles and looking for information and knowledge. All users on internet are confronted with one and other kind of recommendation engine for the above said scenario. Now if they are installed and work correctly, high revenues are generated. Answer is to employ WM (web mining) which is closer to DM (Data Mining), a medium/channel used to derive knowledge from internet by directly communicating with recommendation engine. In WM, recommendation engines are basically classified into Collaborative, Content-Based and Hybrid recommenders. The first type of recommender can further include: U-U (USER-USER), I-I (ITEM-ITEM) and other Collaborative Filtering algorithms. The essential step to make important item recommendations for its customer is the ability to handle better conversions. And if you truly want to connect with the customers then only way is to do interaction with each individual and use recommendation engine involving advanced web mining procedure to serve this purpose. For detailed things to grasp knowledge on the topic of numerous methodologies, reader can go through [48, 49].

## VI. CONCLUSION

With the increase in number of online visitors over the past two decades, recommender systems with diverse recommendations are there to assist them in almost every aspect. Recommendation system is heavily used to provide commercial and personal assistance in both online/offline modes for users. They provide new paths to its user on internet for accessing related information to user itself and take care of the weakness of information overload in information provider engines. In this way, they assist the users to fetch desired service which is not accessible otherwise. Here, discussion on two conventionally established methodologies is made and their salient merits and demerits are summarized. A light on different flavor of hybridization to improve quality using these two conventional methodologies is presented. To create model of recommendation, How to use learning procedures of various kinds, is highlighted. Further, metrics established by earlier researchers to estimate quality quantatively are highlighted. The information highlighted will strengthen researchers to look forward in improving these systems.

## REFERENCES


1. J. A. Konstan, and J. Riedl, "Recommender systems: from algorithms to user experience," User Model User-Adapt Interact, 2012; vol. 22, pp. 101–23.
2. C. Pan, and W. Li, "Research paper recommendation with topic analysis," in Computer Design and Applications IEEE 2010, vol. 4, pp. V4-264.
3. P. Pu, L. Chen, R. Hu, "A user-centric evaluation framework for recommender systems," in Proceedings of the fifth ACM conference on Recommender Systems (RecSys'11), ACM, New York, NY, USA, 2011, pp. 57–164.
4. R. Hu, and P. Pu, "Potential acceptance issues of personality-ASED recommender systems," in Proceedings of ACM conference on recommender systems (RecSys'09), New York City, NY, USA, 2009.
5. B. Pathak, R. Garfinkel, R. Gopal, R. Venkatesan, and F. Yin, "Empirical analysis of the impact of recommender systems on sales," JManage Inform Syst, 2010, vol. 27, No. 2, pp. 159–188.
6. A. M. Rashid, I. Albert, D. Cosley, S. K. Lam, S. M. McNee, and J. A. Konstan, "Getting to know you: learning new user preferences in recommender systems," In Proceedings of the international conference on intelligent user interfaces, 2002, pp. 127–34.
7. R. Burke, "Hybrid recommender systems: survey and experiments," User Model User-adapted Interact, 2002, vol. 12, No. 4, pp. 331–70.
8. J. Bobadilla, F. Ortega, A. Hernando, and A. Gutie´rrez, "Recommender systems survey," Knowl-Based Syst 2013, vol. 46, pp. 109–32.
9. R. O. Duda, P. E. Hart, and D. G. Stork, "Pattern classification," John Wiley & Sons, 2012.
10. N. Friedman, D. Geiger, and M. Goldszmidt, "Bayesian network classifiers," Mach Learn 1997, vol. 29(2–3), pp. 131–63.
11. C. M. Bishop, "Pattern recognition and machine learning," vol. 4, no. 4, Springer, New York, 2006.
12. K. Shyong, D. Frankowski, and J. Riedl, "Do you trust your recommendations? An exploration of security and privacy issues in recommender systems," in Emerging trends in information and communication security," Berlin, Heidelberg: Springer, 2006, pp. 14–29.

13. G. Adomavicius, and A. Tuzhilin, "Toward the next generation of recommender system: A survey of the state-of-the-art and possible extensions," IEEE Trans Knowl Data Eng 2005, vol. 17, no. 6, pp. 734–49.
14. T. Zhang, and S. I. Vijay, "Recommender systems using linear classifiers," J Mach Learn Res 2002, vol. 2, pp. 313–34.
15. D. Billsus, and M. J. Pazzani, "User modeling for adaptive news access," User Model User-adapted Interact 2000, vol. 10, Issue 2–3, pp. 147–80.
16. R. J. Mooney, and L. Roy, "Content-based book recommending using learning for text categorization," in Proceedings of the fifth ACM conference on digital libraries ACM, 2000, pp. 195–204.
17. J. L. Herlocker, J. A. Konstan, Terveen, LG, and J. T. Riedl, "Evaluating collaborative filtering recommender systems," ACM Trans Inform Syst 2004, vol. 22, Issue 1, pp. 5–53
18. J. Bobadilla, F. Ortega, A. Hernando, and A. Gutie´rrez, "Recommender systems survey," Knowl-Based Syst 2013, vol. 46, pp. 109–32.
19. J. Breese, D. Heckerma, and C. Kadie, "Empirical analysis of predictive algorithms for collaborative filtering," In: Proceedings of the 14th conference on uncertainty in artificial intelligence (UAI-98),1998. pp. 43–52.
20. Z. D. Zhao, and M. S. Shan, "User-based collaborative filtering recommendation algorithms on Hadoop," In: Proceedings of 3rd international conference on knowledge discovering and data mining, (WKDD 2010), IEEE Computer Society, Washington DC, USA, 2010, pp. 478–81. doi: 10.1109/WKDD.2010.54.
21. X. Zhu, H. W. Ye, and S. Gong, "A personalized recommendation system combining case-based reasoning and user-based collaborative filtering," In: Control and decision conference (CCDC 2009), Chinese, 2009, pp. 4026–28.
22. J. B. Schafer, D. Frankowski, J. Herlocker, and S. Sen, "Collaborative filtering recommender systems," In: Brusilovsky P, Kobsa A, Nejdl W, editors. The Adaptive Web, LNCS 4321. Berlin Heidelberg (Germany): Springer, 2007, pp. 291–324. http://dx.doi.org/10.1007/978-3-540-72079-9_9.
23. R. Burke, "Web recommender systems," In: P. Brusilovsky, A. Kobsa, and W. Nejdl, editors, The Adaptive Web, LNCS 4321. Berlin Heidelberg (Germany): Springer, 2007, pp. 377–408. http://dx.doi.org/10.1007/978-3-540-72079-9_12.
24. R. Burke, "Hybrid recommender systems: survey and experiments," User Model User-adapted Interact 2002, vol. 12, Issue 4, pp. 331–70.
25. D. H. Park, H. K. Kim, I. Y. Choi, and J. K. Kim, "A literature review and classification of recommender systems research," Expert Syst Appl 2012, vol. 39, Issue 11, pp. 10059–72.
26. G Adomavicius, and J. Zhan, "Impact of data characteristics on recommender systems performance," ACM Trans Manage Inform Syst 2012, vol. 3, issue 1.
27. D. H. Stern, R. Herbrich, and T. Graepel, "Matchbox: large scale online bayesian recommendations," In: Proceedings of the 18th international conference on World Wide Web. ACM, New York, NY, USA, 2009, pp. 111–20.
28. M. Claypool, A. Gokhale, T. Miranda, P. Murnikov, D. Netes, and M. Sartin, "Combining content- based and collaborative filters in an online newspaper," In: Proceedings of ACM SIGIR workshop on recommender systems: algorithms and evaluation, Berkeley, California, 1999.
29. D. Billsus, and M. J. Pazzani, "A hybrid user model for news story classification," In: J. Kay, editor. In: Proceedings of the seventh international conference on user modeling, Banff, Canada. Springer-Verlag, New York; 1999, pp. 99–108.
30. B. Smyth, and P. Cotter, "A personalized TV listings service for the digital TV age," J Knowl-Based Syst 2000, vol. 13, Issue 2–3, p. 53–9.
31. A. M. Wasfi, "Collecting user access patterns for building user profiles and collaborative filtering," In: Proceedings of the 1999 international conference on intelligent user, Redondo Beach, CA; 1999, pp. 57–64.
32. R. Burke, K. Hammond, and B. Young, "The FindMe approach to assisted browsing," IEEE Expert, 1997, vol. 12, Issue 4, pp. 32–40.
33. C. Basu, H. Hirsh, and W. Cohen, "Recommendation as classification: using social and content-based information in recommendation," In: Proceedings of the 15th national conference on artificial intelligence, Madison, WI, 1998, pp. 714–20.
34. R. J. Mooney, and L. Roy. "Content-based book recommending using learning for text categorization," In: Proceedings of the fifth ACM conference on digital libraries, ACM, 2000, pp. 195–204.
35. I. Schwab, A. Kobsa, I. Koychev, "Learning user interests through positive examples using content analysis and collaborative filtering," Draft from Fraunhofer Institute for Applied Information Technology, Germany, 2001
36. J. L. Herlocker, J. A. Konstan, L. G. Terveen, and J. T, Riedl, "Evaluating collaborative filtering recommender systems," ACM Transaction on Information Systems, 2004, vol. 22, Issue 1, pp. 5-53,.
37. I. Schwab, K. Kobsa, and I. Koychev, "Learning user interests through positive examples using content analysis and collaborative filtering," Draft from Fraunhofer Institute for Applied Information Technology, Germany, 2001.
38. C. Cleverdon, and M. Kean, "Factors Determining the Performance of Indexing Systems, Aslib Cranfield Research Project, Cranfield, England, 1968.
39. D. Billsus, and M. J. Pazzani, "Learning collaborative information filters," In Proceedings of the 15th National Conference on Artificial Intelligence (AAAI-98), C. Rich, and J. Mostow, Eds. AAAI Press, Menlo Park, Calif., 1998, pp. 46–53.
40. C. Basu, H. Hirsh, and W. W. Cohen, "Recommendation as classification: using social and content-based information in recommendation," In Proceedings of the 15th National Conference on Artificial Intelligence (AAAI-98), C. Rich, and J. Mostow, Eds. AAAI Press, Menlo Park, Calif., 1998, pp. 714–720.
41. B. M. Sarwar, G. Karypis, J. A. Konstan, and J. Riedl, "Analysis of recommendation algorithms for E-commerce," In Proceedings of the 2nd ACM Conference on Electronic Commerce (EC'00), ACM, New York, 2000a, pp. 285–295.
42. B. M. Sarwar, G. Karypis, J. A. Konstan, and J. Riedl, "Application of dimensionality reduction in recommender system–A case study," In Proceedings of the ACM WebKDD 2000 Web Mining for E-Commerce Workshop, 2000b.
43. D. M. Powers, "Evaluation: From precision, recall and f-factor to roc, informedness, markedness and correlation," Technical report, School of Informatics and Engineering, Flinders University Adelaide, South Australia, 2007.
44. F. O. Isinkaye, Y. O. Folajimi, B. A. Ojokoh, "Recommendation Systems: Principles, methods and evaluation,", Egyptian Informatics Journal, Elsevier B.V. , 2015, pp. 261-273.
45. I. Portugal, P. Alencar, D. Cowan, "The use of machine learning algorithms in recommender systems: A systematic review, Expert Systems with Applications," 2018, vol.-97, pp. 205-227.
46. S. Zhang, Y. Lina, S. Aixin, and Y. Tay, "Deep Learning based Recommender System: A Survey and New Perspectives," ACM Comput. Surv.,July 2018.
47. Dou De, Wang H, Liu H., Semantic Data Mining: A survey of Ontology-based Approaches, IEEE 9th International Conference on Semantic Computing, 07-09 Feb, 2015.
48. Q. Zhang and R. S. Segall, "Web Mining: A Survey Of Current Research, Techniques and Software," International Journal of Information Technology & Decision Making, 2008, Vol. 7, No. 4, pp. 683–720
49. R. Kosala and H. Blockeel, "Web Mining Research: A Survey," SIGKDD Explorations, July 2000, vol.-2, Issue-01.

## AUTHORS PROFILE

**Deepak Vats** has received his M. Tech. in Computer Science & Engineering at Chaudhary Devi Lal University, Sirsa, Haryana, India in 2007. He is currently working as an Assistant Professor at Department of Computer Applications, University Institute of Computing, Chandigarh University, Gharuan, Mohali, Punjab. He served as Assistant Professor at the Department of Information Technology/Computer Science in Maharishi Markandeshwar (Deemed to be University), Mullana, Ambala, Haryana from 2007 to 2019. His research interests include data mining, algorithms, semantic web and machine learning.

He is currently working as Principal and Professor at Maharishi Markandeshwar Engineering College, Mullana, Ambala, Haryana. The Constituent institution of Maharishi Markandeshwar (Deemed to be University), Mullana, Ambala, Haryana is NAAC accredited 'A' Grade Deemed to be University & first private engineering college of Haryana established in 1995. He is Dean Faculty of Engineering and Technology & Member of Board of Studies & DRC Committee for Research. He works as Ex-Principal & Professor, Rajasthan College of Engineering for Women (Leading Women's Engineering College in the state of Rajasthan). Publications: International Journals: Published: 45 (Accepted: 61), International Conferences: Published: 90 (Accepted: 102) National Conferences & Workshops: 75 (Accepted: 110 & more) Text

Books/Edited: 06. Total Experience: 20 years (10 years PG) + 03 years Research International Conferences Organized: 10. He is an active member of Advisory/Technical program committee of reputed International/National conferences & reviewer of number of reputed Journals e.g. Springer, Elsevier Journal Computers & Electrical Engineering. Approximate 20 years of rich experience in Teaching, research and industry managing technical institution, serving in all capacity including Head of Department, Professor, Controller of Examination, Dean Academics Affairs, Principal etc. He has played leading role in accreditation of the institution and ISO 9001: 2000 certification (Including 05 years of industrial/research experience).