

# Survey of User Profiling in News Recommender Systems

Mahboobeh Harandi  
School of Information Studies  
Syracuse University  
Syracuse, USA  
mahboobh@stud.ntnu.no

Jon Atle Gulla  
Department of Computer and Information  
Science, NTNU  
Trondheim, Norway  
jag@idi.ntnu.no

## ABSTRACT

In order to personalize news articles in online news recommender systems, a number of user profiling techniques have been employed. Both long-term and short-term interests of the user are captured in this process and are important to the construction of user profiles. Due to the short life span and the unstructured format of news articles, the changing interests of the user and the lack of explicit feedback, user profiling is challenging in the news domain. Natural language processing and machine learning techniques are used, though there is no accepted best approach to user profiling. In this survey we discuss the most common user profiling techniques in news recommendation and show how they can be classified according to features used and challenges addressed.

## Categories and Subject Descriptors

H.4 [Information Systems Applications]: Miscellaneous

## Keywords

Recommender systems, user profiling, supervised and unsupervised learning techniques, natural language processing, content-based filtering, collaborative filtering, hybrid filtering, news recommendation

## 1. INTRODUCTION

Recommender systems have become popular over the last few years because they have the ability to provide relevant information to users based on their needs or interests. Their goal is to filter out information and present only the required and interesting pieces to the user. This type of application need to understand the information they manage, but also to understand the users' behavior and their underlying information needs.

With the abundance of information on the web, there is a problem that users get overwhelmed while looking for the desired information. The users cannot be asked to browse through hundreds of items to find the correct one, or to formulate unambiguous search queries that can pick out the correct document from millions of similar documents. Recommender systems try to analyze previous user behavior to understand what the user might find interesting in the future. There are today several popular recommender systems used in a wide variety of domains. In the movie domain, Netflix is a successful pioneer of the technology and can on-the-fly recommend movies to users on the basis of what s/he has watched and what s/he has rated. The company has more than 50 million subscribers all around the world and has been fairly successful in monetizing the technology and coming up with attractive personalized movie recommendations. Another successful application of recommender systems is found in Amazon, where different items are proposed to the user based on their shopping history.

News recommender systems is another popular application of recommendation technologies. These systems provide interesting and integrated news from thousands of news stories from media houses and news agencies. From the user perspective, it is helpful if the recommender system can propose interesting articles instead of forcing the reader to spend too much time looking for them. Times, Google News, Daily learner, News360 and NTNU SmartMedia are examples of both commercial and research-oriented news recommender systems.

In order to personalize news articles, item descriptions and user profiles need to be studied and taken into account. The latter one will be discussed in this paper.

The internal functionality of recommender systems is very similar to information retrieval in search engines, in which user profiles are interpreted as queries submitted to an underlying search index. To recommend news articles based on user's interest, the interaction logs are manipulated and stored as user profiles. To build a solid foundation for intelligent recommendations with only user's implicit feedback, different techniques of machine learning are normally applied. After the interaction of user with system, the history of user interactions can serve as training data that is a basis for prediction of new desired news article.

In this paper, we classify supervised and unsupervised machine learning techniques and discuss the features of the user profiles after applying each of them. The challenges of news recommender systems handled by these features are presented. The paper is

organized as follows: In section 2 news recommender systems and their details and challenges are described. In section 3 different dimensions of user profiles and machine learning techniques are explained. Features of the user profiles with respect each technique of learning are summarized. In section 4, applying the filtering techniques for content-based, collaborative and different kinds of hybrid system is discussed. The classification of machine learning techniques and their addressed problems are illustrated, before the conclusions are presented in Section 5.

## 2. News Recommender Systems

News recommender systems share many features with information retrieval systems and human computer interaction as well. Text mining techniques for large scale data sets are needed, and machine learning methods are employed when learning cycles can be built into the systems. In general there are three steps. First of all, data pre-processing such as sampling, dimension reduction, denoising with use of similarity functions are normally applied. Then the text is analyzed through supervised or unsupervised machine learning techniques depending on availability of training data sets. At the end the result is interpreted through for example the  $F_1$ -measure, ROC or MAE [1].

If we consider news recommender system as search engines, the user profiles can be regarded as long search queries. The system ranks the results on the basis of well the profile matches the descriptions of the news articles. Formally, the appropriateness of recommended news to the user can be described by the following utility function [1]:

$$u: C \times S \rightarrow R$$

This function assigns a score  $r$  for each combination of user  $c$  and news story  $s$ . Matrix  $C$  indicates the characteristics of the user and  $S$  shows the different specifications of available articles such as topic, location, news agency, date and other useful attributes. All different algorithms in recommender systems try to maximize the result matrix. Each entry of  $R$  could be any non negative internal between 0 and 1 or 0 and 100 based on the system definition. At the end, an article that maximizes the utility function will be recommended [1]:

$$c' = \underset{s \in S}{\operatorname{argmax}} u(c, s)$$

News recommender systems differ in the context of items structures from other recommenders. The structure of news articles is not following any specific format. There are many news articles in a day that have very short life spans while the system must scale to deal with huge volumes of data. Besides, the news recommender system must always recommend interesting articles to the user, though it should not make over-specialize for the target user. [2]

## 3. User Profiles

The desired user profiles need to have a changing essence and flexible content. These profiles show their preferences towards news articles by modeling the interesting articles. Besides, storing user interactions is a basis to know their favorite topics which last longer and which are only for a short period of time.

This model consists of meta-data such as time and location, which is changing according to the user behavior.

The content of the user profile for this kind of recommender system which has not very structured format is different from others. In order to have an exact and practical model of the user profiles, the system needs to know the behavior of the user including background, interest and goals. These features are changing over time, so considering the temporal parameters such as time and location is crucial [3].

There are three major presentations of terms in the user profile. The first approach is presenting terms as vectors in a vector space model. In order to weigh correctly every single word based on its frequency in every document and in the collection of documents, TF-IDF is often applied. This measure puts more emphasis on one word that appears frequently in one specific document and not in other ones. So it will gain more weight and appointed document, will be retrieved to a target user. But the problem of polysemy (multiple meaning for one word) and synonymy (multiple words for identical meaning) remain. The desired approach reflects cultural and linguistic knowledge of terms and also could use reasoning on their content. As a result, the presentation is more intelligent and is not a simple bag of words and could provide the knowledge about desired terms [1]. The second one is the analysis words in the format of entity. They have meanings and relations, but they suffer from generalization or specialization since there is no hierarchical relationships among the entities [3]. The third one is the semantic analysis that is ontology-based. It has hierarchical relationships between the semantic concepts modeling user interests. The terms that indicate the user interests including their interests that last longer or the ones that appear only for a short time could be enriched by semantic approaches. The advantage of providing ontologies for the user interests is that all the terms or entities are in hierarchical relationships which give more specific detail of user interests at the side of the general ones [3]. The semantic enrichment could benefit from encyclopedic knowledge beside the knowledge of applied documents. So the terms are semantic vectors in word space model [1]. Each of them are indexed by their weights but later will be interpreted semantically by using Wikipedia. It is called Explicit Semantic Analysis (ESA) [4].

The feedback of the user is the other approach of user modeling. In general s/he could communicate and provide their interest towards the news explicitly or implicitly. Explicit feedback is to provide their interest (disaffection) directly to the system. It could be actions such as rating, like or filling the survey through the interface of the application. Implicit feedback includes the interactions such as click on articles (touch in mobile device), scrolling articles using a mouse or a keyboard (swapping in the mobile device), printing or saving articles, copying and posting a part or all of articles, reading articles, forwarding or sharing the articles and providing the qualitative comments on the article. Recommender systems are highly dependent on user feedback. As long as the user interacts with the application, the accuracy of the system may gradually improve. Explicit feedback tends to produce more exact user profiles than what is possible with implicit feedback. Unfortunately, not all users are willing to spend time to provide such feedback, so the implicit signals of the users are normally the basis of the recommendation [5].

Specifying the type of user's interest could help the system to cover all domains of their attention. The long-term interest is more dependent on the user profession and the personal background than what will be traced by the log history. But the short-term interest is the one mostly related to the current trend of public that s/he has communication with. Although depending on the goals, the long-term interest will change gradually. Besides,

supervising the context of user's attention could provide good evidence to capture the short-term interest and update their long-term interest time by time. In [6] by defining running context over category and topic, the current user's interest is captured. The old user profile that is the indicator of their long-term interest is updated progressively if there is nothing in common with their current focus. Besides, there should be a balanced focus on the old and new user profile. While keeping the old user profile and over looking the context results in dissatisfaction, giving too much priority to the current context will not cover the news articles that are related to their background and are the basis of their interest. In addition, different time of day (morning, evening) and week (weekdays and weekend) could affect the user profile [7]. Considering the topic of the news articles, target users may have different desires at different times. As an example, s/he might have more interests in politics and economics in weekdays and focus more on lifestyle news in the weekend [8].

While personalizing the news is desirable, the importance of public trend is not negligible. In [9] based on the frequency of user clicks, public trend could provide the interesting news articles as well. If there are not enough clicks from the user side, then according to their location, public trend of that location is a good indicator to recommend the news. This dimension of the user profile that specifies the location has a key role in recommending news articles. Short-term interests of the user are highly dependent on their location. Location could capture public trend and find similar networks of users as well. Sometimes ignoring the user profile and focus on the context is helpful (in economical news, user profile is not very helpful but the context tracing is more informative), while other times it is better to count only on the user profile (for entertainment section user profile enrichment is much better than context) [10].

As the amount of data explodes, the importance of extracting models and predicting unseen data with machine learning techniques is increasing [11]. There are two major types of learning techniques, supervised and unsupervised. In the former one, an annotated training dataset is provided, whereas in the latter one, the machine explores the data to identify interesting patterns without training data. Below is the list of supervised learning techniques used in recommender systems:

- Decision Trees (C4.5 or KART) handle categorical-nominal and heterogeneous data. It is also able to cope with missing values. Through pre pruning, overfitting will be addressed. It tends to work well with small sized datasets, though the cost of decisions on continuous data streams is high [11, 12].
- Rule-based (RIPPER) can handle multi value features very well. It is decision tree-based and uses rules to categorize new items. It utilizes post pruning to find the best fit for the rule set [13].
- K Nearest Neighbor (KNN) can handle continuous data through Euclidean, Manhattan or Minkowski distance and cope with categorical data through Hamming distance. It is a lazy learner that works well with few instances [14, 15].
- Rocchio and Relevance Feedback: the user profile is regarded as a query [16] and based on the implicit feedback of user, the recommendation will be improved in time.
- Support Vector Machine (SVM): through SVM reduction of sensitivity to the noises and increasing generalization is done. For non linear problem if

features are more than instances, linear kernel is good enough to be applied [16, 17].

- Probabilistic methods and Naive Bayes: Bayesian Belief Network with conditional independency is the most applicable one. Multinomial (Bernoulli) and multivariate are two types of Naive Bayes. While in the Bernoulli model absence or presence of a model is checked, in multivariate one the number of occurrences of a term will be calculated [18, 19].
- Neural Network: Single layer perceptron and multi layer for non linear separable problems are the samples of applied neural network in the recommender systems [20].

Below is the list of unsupervised learning techniques:

- Probabilistic methods: If the structure of Bayesian network is not known then building the DAG Bayesian with scoring function, constraint based learning or Conditional Independency can be applied. The last one has more efficiency [21]. The other techniques such as Bayesian Hidden Score (pairwise learning) and graph-based learning have been applied in [22].
- Neural Network: Self Organizing Map (Kohonen) and Restricted Boltzmann Machine belong to the category of unsupervised learning [20].
- Clustering: flat clustering by k-means algorithm deals with the categorical data and the most frequent term will be the centroid. In the hierarchical clustering, the other type of clustering, divisive is more accurate than agglomerative. There are two approaches to label clusters. The first one is differential that through feature selection a label with a higher score will be chosen. The second one is inter clustering that the closest one to the title or the higher weight to the centroid of the cluster will be chosen as the label. The drawback of cluster-internal labeling is disability to distinguish between words which are frequent in the whole clusters and the ones that are frequent only in one specific cluster. Labeling in hierarchical clustering due to the dependent definitions of parent, child and sibling is more complicated [16].

Table 1 shows the applied machine learning techniques to build up a user profile.

#### 4. Applying User Profiles in Recommender Systems

There are different approaches to filter out the information. Content-based and collaborative filtering are the most applicable ones. In content-based filtering, the concept of news articles will be analyzed. Then according to the content of the user profile (i.e. characteristic of read articles), similar articles are predicted and presented to the user. In the content-based filtering, the utility function is:

$$u(c, s) = \text{score}(\text{contentbased user profile } (c), \text{item content } (s))$$

If each of the content of the user profile and item profile is represented by TF-IDF weight, then the scoring function could be calculated through cosine similarity of vectors of the weight. To achieve the accurate prediction, attributes of news articles that have been counted on, are important. Since the nature of news article is unstructured, extracting relevant and important features has a key role in content-based filtering. If the articles are

categorized with minimum misclassification error, then storing interesting news articles in the user profile is much easier and consequently, recommendations are of higher quality. Bayesian Networks can be utilized well for learning user profiles based on the articles that have been read. It can model profiles of the users through ignoring missing data and considering conditional dependency in one specific category of news articles. It can provide probabilities of each attribute of article by its nodes. The modeled domain includes continuous data. Then similarity of the user profile based on predicted attributes of article and available news articles is computed and the ones with the highest score will be recommended. If another technique such as Naive Bayes

(Bernoulli Model) is applied for modeling user behavior, the output is binary as it is considering absence or presence of terms regardless of their conditional independency [1]. It can suggest the new item to the target user by comparing the new item's characteristics to the terms in the user's profile. But if there is not enough attributes, content-based filtering is normally not the most efficient one. If the user is new to the system it cannot recommend anything as there is no content of their profile available. Besides, it causes lack of serendipity due to providing too many similar news articles to the user. Considering the collaborative approach for filtering information, there are two different models, memory-based and model-based. Memory-based utilizes the log

**Table 1. ML techniques and features of user profiles**

ML Techniques	User Profile Features
<i>Decision Tree (C4.5)</i>	Semantic enrichment can be handled at entity level, but in the beginning of building the user profile or for capturing short-term interest [13, 23].
<i>Rule-based (RIPPER)</i>	Semantic enrichment can be handled at entity level. More interesting categories of news may be predicated through rules [1].
<i>KNN</i>	Captures the short term interest of user and popularity of the item among a group of user.
<i>Rocchio and Relevance Feedback</i>	User profiles are regarded as queries, the system improves over time from relevance feedback of the user [16].
<i>Support Vector Machine</i>	It outperforms KNN, C4.5 and Rocchio [16] with the Reuters dataset
<i>Probabilistic methods and Naive Bayes</i>	Bernoulli works well with small sizes of data set and multinomial works well in large sizes of datasets. DAG captures the dependency of items in more detailed capturing interest, vigorous towards missing data and could disregard noisy data. BHS and graph-based capture online interest of the user [22]
<i>Neural Network</i>	It can represent details of the user's interest through deep learning of three layer perceptron [24].
<i>Clustering</i>	The content of the items are clustered and then item-based collaborative is implemented on the output. Fuzzy membership over the k-means. Similarity of the item-rating matrix, the group-rating matrix (MovieLense) Hierarchical clustering for the news groups (LDA for small dataset and PLSI for large dataset) [25]

history of all users and put top-N similar users who have the same taste about the news articles into one specific group.

Then to provide the latest and interesting news articles to the target user, it filters out users with the same interest and recommends the new articles that have been read by them. It is working with a matrix of user's profile and all the news articles. It is possible to apply K Nearest Neighbor (through neighborhood measurement) to find the closest users to the current active user. The other approach is applying similarity measurement like cosine similarity or Pearson correlation, which provide the new item for the target user if it has similarity with previous chosen items. It can help us find similar users or items regarding to the context of memory [23].

The other type of collaborative filtering is model-based. It is more scalable and much faster than memory based collaborative filtering. Through this type of filtering not all the dataset will be traced and investigated, but only some information will be modeled. As finding the similarity between users or news articles (users with the same interest in the specific news or two

similar news articles that are interesting to one specific user) is not feasible due to lack of labeled data in the training phase, clustering of news or users could be a practical solution. With the Google News dataset, clustering is done on the basis of users' clicks on different news article. Through clustering, latent factors (latent semantic analysis) can be revealed. Consequently, ignoring the hidden values will result in a very poor accuracy. It could be helpful to distinguish hidden variables through the clustering and provide more accurate prediction of news articles [23]. One the technique to implement this approach is building up the matrix of users and item as matrix factorization. The matrix of users and news articles is suffering from sparsity, since there are several positions that users do not provide any feedback. To find the hidden variables that affect the recommendation as well,  $UV$  decomposition (it is one instance of Singular Value Decomposition) is possible to be applied. If the utility matrix  $M$  is  $n \times m$  ( $n$  indicates the user and  $m$  indicates the news articles), then  $UV$  decomposes it multiplication of two different matrixes including  $n \times d$  and  $d \times m$ :

$$M \rightarrow U \times V$$

RMSE is a common tool to measure the accuracy of prediction blank entries in  $M$  considering the product  $UV$ .

Although it is working much faster than memory based, it is less exact than it. In spite of all the applicable different approaches of collaborative filtering, it cannot make the accurate prediction for the new user or the new item (cold-start problem). The core of all the algorithms is dependent on the group of users (or items) in order to find the proper match for the target user. Consequently it has nothing to present to the user with unique taste.

As each of these filtering techniques has its own problems and challenges in recommender systems, a hybrid system is often preferred. It takes into account both filtering in predefined step and could overcome drawback of each. Considering two techniques of filtering (content-based and collaborative), the order of combination of them might be important to build a hybrid system. Although in some techniques of hybridization, the order is not a matter. The techniques that order is not important are [26]:

**Mixed:** the result from both techniques will be presented in one grouped or separate list. It has been utilized in [27] to provide the TV shows to the users. The mixed hybrid system provides recommendations based on the characteristics of each show and preferences of other users.

**Weighted:** The score for each technique is computed, and the weighting of final score will be the basis for the recommendation. In personalized Tango (P-tango) for online newspapers, equal weights are assigned to both filtering techniques. Gradually each weight is increasing regarding the user rating. Based on the rating, the absolute error is computed and is decreasing through the better recommendation.

**Switching:** This technique uses some criterion to switch between filtering techniques and based on the specific chosen filter, recommends the item. In the DailyLearner switching hybrid system, content-based filtering with  $k$  nearest neighbor is first applied. If it does not produce sufficient recommendations, collaborative filtering takes advantage of similar users' interests to recommend desired items. In another system, item-based collaborative filtering is triggered if the accuracy of the content-based filtering part is low [28].

**Feature combination:** The technique takes advantage of one filtering type such as collaborative filtering as feature allied with data. Then content-based filtering is applied. Through this kind of hybrid system, the absolute dependency on users is dropped by applying collaborative filtering as a feature

combination. In the movie recommender domain [29], the RIPPER algorithm is implemented with item features and users rating.

There are three other models of hybrid systems that are ordered by their intrinsic structure:

**Feature augmentation:** One of the filtering techniques is applied to compute rating scores or to classify items. The output of this filtering is the input for the other filtering technique. In Libra system, content-based filtering through Naïve Bayes is done on data that comes from Amazon. The data from Amazon that show related authors and titles were implemented using collaborative filtering. Collaborative filtering is done first.

**Meta-level:** It provides a model through one of the filtering methods as an input for the other one. The model is the complete one, not a learned model like feature augmented techniques. In Fab [30] at first by means of relevance feedback and the Rocchio algorithm, collections of items (the need of users in mass of dataset in web) are composed (content-based).  $K$ -nearest neighbor is then used with collaborative filtering to complete the recommendations. Meta-level is the only ordered technique that applies content-based filtering first.

**Cascade:** Approximately similar to the other ordered techniques, it refines the result of candidates that have been filtered by the previous technique. But if the items in the first filtering have very low priorities, they will not be in the second filtering stage. In fact, the second filtering step is only applied to provide more accurate recommendations and if an item has not enough rating score, it will not be in the second phase. Fab [31] is the example of this technique. With collaborative filtering on the selection stage, the items are chosen with an exact score and presented to the user.

According to the implemented hybrid systems in news recommender system (such as Daily Learner), switching schema is the most common strategy. It can start with content-based filtering and utilize Naive Bayes to categorize the news articles based on the content of the articles and apply item-based collaborative filtering to calculate the similarity between the news articles and the user profile. On the other hand, it is also possible to apply collaborative filtering to find the closest users to the active user (through KNN) and then with content-based filtering identify much more similar items based on the similarity computation of user profile and news articles.

Table 2 shows the applied machine learning techniques to deal with the issues of news recommender systems.

Table 2. Machine learning techniques and challenges addressed

ML Techniques	Challenges addresses of news recommender system
<i>Decision Tree (C4.5)</i>	Capturing short term interest [1].
<i>Rule-based (RIPPER)</i>	Serendipity can be supported with new category reasoning [32].
<i>KNN</i>	Short-term interests and provide the latest news to the user based on their interests [1].
<i>Rocchio and Relevance Feedback</i>	Handling long-term interest of the user [1].
<i>Support Vector</i>	Sparse Problem and huge data after a long time usage of the application[33].

<i>Machine</i>	
<i>Probabilistic methods and Naive Bayes</i>	<p>Handling long-term interest of the user</p> <p>Sparse problem</p> <p>Noisy data</p> <p>Cold Start</p> <p>Precious interest of the user [28].</p>
<i>Neural Network</i>	<p>Short term and long term [34].</p> <p>Tied Boltzmann with residual parameter could outperform on non cold-start problem in comparison with simple method of collaborative filtering, Pearson correlation for the items. It also is competitive with the cold-start problem in content-based filtering. (Netflix)</p> <p>Changing interest of the user [24].</p>
<i>Clustering</i>	<p>Cold start</p> <p>Through fuzzy membership new and interesting news articles are possible to be represented to the user [25].</p>

## 5. Conclusion

The news recommender system is somewhat different from other recommender systems. It is used to provide a variety of personalized news articles that have very short life spans. In addition the range of the user's interests is wide and changing over time and contexts. These characteristics necessitate very dynamic analyses of user profiles.

In this paper the distinguishable characteristics that affect recommendation strategies are assessed. The user feedback on recommended items is one of them. Different algorithms of machine learning (that fall into the categories of supervised and unsupervised) are discussed to build up user profiles. On the other hand, as the user profile is dependent on the whole framework of filtering methods, the techniques are also studied. They utilize user profiles in diverse ways which affect the accuracy of the corresponding recommendations.

## References

- [1] Ricci, F., et al., Recommender Systems Handbook. 2010: Springer-Verlag New York, Inc. 842.
- [2] Özgöbek, Ö., J. A. Gulla., R. C. Erdur, A Survey on Challenges and Methods in News Recommendation, in In Proceedings of the 10th International Conference on Web Information System and Technologies April 2014: Barcelona.
- [3] Bouneffouf, D., Towards User Profile Modelling in Recommender System. 2013.
- [4] Gabrilovich, E. and S. Markovitch, Computing semantic relatedness using Wikipedia-based explicit semantic analysis, in Proceedings of the 20th international joint conference on Artificial intelligence. 2007, Morgan Kaufmann Publishers Inc.: Hyderabad, India. p. 1606-1611.
- [5] Jon Atle Gulla, J.E.L., Arne Dag Fidjestøl, John Eirik Nilsen, Kent Robin Haugen, and Xioameng Su, Learning User Profiles in Mobile News Recommendation. Journal of Print and Media Technology Research, September 2013. Vol II, No. 3: p. pp. 183-194.
- [6] Jon Atle Gulla, A.D.F., Xiaomeng Su and Humberto Castejon, Implicit User Profiling in News Recommender Systems, in In Proceedings of the 10th International Conference on Web Information System and Technologies April 2014: Barcelona.
- [7] Abel, F., et al., Analyzing user modeling on twitter for personalized news recommendations, in Proceedings of the 19th international conference on User modeling, adaption, and personalization. 2011, Springer-Verlag: Girona, Spain. p. 1-12.
- [8] Adomavicius, G. and A. Tuzhilin, Context-aware recommender systems, in Proceedings of the 2008 ACM conference on Recommender systems. 2008, ACM: Lausanne, Switzerland. p. 335-336.
- [9] Liu, J., et al., Personalized news recommendation based on click behavior, in Proceedings of the 15th international conference on Intelligent user interfaces. 2010, ACM: Hong Kong, China. p. 31-40.
- [10] Bellogín, A., et al. Discovering Relevant Preferences in a Personalised Recommender System using Machine Learning Techniques. in Preference Learning Workshop (PL 2008), at the 8th European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML PKDD 2008). 2008.
- [11] Witten, I.H., E. Frank, and M.A. Hall, Data Mining: Practical Machine Learning Tools and Techniques. 2011: Morgan Kaufmann Publishers Inc. 664.
- [12] VENKATADRIM, L.C.R. A Comparative Study On Decision Tree Classification Algorithms In Data Mining. 2010; Available from: [https://www.academia.edu/1374211/A\\_Comparative\\_Study\\_On\\_Decision\\_Tree\\_Classification\\_Algorithms\\_In\\_Data\\_Mining](https://www.academia.edu/1374211/A_Comparative_Study_On_Decision_Tree_Classification_Algorithms_In_Data_Mining).
- [13] Pazzani, M.J. and D. Billsus, Content-based recommendation systems, in The adaptive web, B. Peter, K. Alfred, and N. Wolfgang, Editors. 2007, Springer-Verlag. p. 325-341.
- [14] Deokar, S., WEIGHTED K NEAREST NEIGHBOR. 2009.
- [15] Webb, G., M. Pazzani, and D. Billsus, Machine Learning for User Modeling. User Modeling and User-Adapted Interaction, 2001. 11(1-2): p. 19-29.
- [16] Manning, C.D., et al., Introduction to Information Retrieval. 2008: Cambridge University Press. 496.

- [17] Prügel-Bennett, M.A.G.a.A., Building Switching Hybrid Recommender System Using Machine Learning Classifiers and Collaborative Filtering. IAENG International Journal of Computer Science.
- [18] Margaritis, D., Learning Bayesian Network Model Structure. 2003, University of Pittsburgh.
- [19] Barber, D., Bayesian Reasoning and Machine Learning. 2010.
- [20] Peretto, P., An Introduction to the Modeling of Neural Networks. 1992: Cambridge University Press.
- [21] Kotsiantis, S.B., Supervised Machine Learning: A Review of Classification Techniques, in Proceedings of the 2007 conference on Emerging Artificial Intelligence Applications in Computer Engineering: Real World AI Systems with Applications in eHealth, HCI, Information Retrieval and Pervasive Technologies. 2007, IOS Press. p. 3-24.
- [22] Bian, J., et al., Exploiting User Preference for Online Learning in Web Content Optimization Systems. ACM Trans. Intell. Syst. Technol., 2014. 5(2): p. 1-23.
- [23] Rajaraman, A. and J.D. Ullman, Mining of Massive Datasets. 2011: Cambridge University Press. 326.
- [24] Gunawardana, A. and C. Meek, A unified approach to building hybrid recommender systems, in Proceedings of the third ACM conference on Recommender systems. 2009, ACM: New York, New York, USA. p. 117-124.
- [25] Mouton, C., Unsupervised Word Sense Induction from Multiple Semantic Spaces with Locality Sensitive Hashing, in International Conference RANLP. 2009.
- [26] Burke, R., Hybrid Recommender Systems: Survey and Experiments. User Modeling and User-Adapted Interaction, 2002. 12(4): p. 331-370.
- [27] Cotter, P. and B. Smyth, PTV: Intelligent Personalised TV Guides, in Proceedings of the Seventeenth National Conference on Artificial Intelligence and Twelfth Conference on Innovative Applications of Artificial Intelligence. 2000, AAAI Press. p. 957-964.
- [28] Ghazanfar, M.A. and A. Prugel-Bennett, An Improved Switching Hybrid Recommender System Using Naive Bayes Classifier and Collaborative Filtering. International Multiconference of Engineers and Computer Scientists (Imecs 2010), Vols I-Iii, 2010: p. 493-502.
- [29] Basu, C., H. Hirsh, and W. Cohen, Recommendation as classification: using social and content-based information in recommendation, in Proceedings of the fifteenth national/tenth conference on Artificial intelligence/Innovative applications of artificial intelligence. 1998, American Association for Artificial Intelligence: Madison, Wisconsin, USA. p. 714-720.
- [30] Balabanovi, M., #263, and Y. Shoham, Fab: content-based, collaborative recommendation. Commun. ACM, 1997. 40(3): p. 66-72.
- [31] Balabanovi, M. and #263, An adaptive Web page recommendation service, in Proceedings of the first international conference on Autonomous agents. 1997, ACM: Marina del Rey, California, USA. p. 378-385.
- [32] Markward Britsch, N.G., Michael Schmelling, Application of the rule-growing algorithm RIPPER to particle physics analysis. 2008.
- [33] Anatole Gershman, T.W., Eugene Fink, Jaime Carbonell, News Personalization using Support Vector Machines.
- [34] Kyo-Joong Oh, W.-J.L., Chae-Gyun Lim, Ho-Jin Choi, Personalized News Recommendation using Classified Keywords to Capture User Preference, in Advanced Communication Technology (ICACT). 2014.