"News Recommendation System Using Online Learning"

Fifth-Semester Project Report

Submitted By
Sarabjyot Singh (IIT2016114)
Mohit Verma (IIT2016112)
Anand Singh (IRM2016007)
Sudhanshu Yadav (IWM2016003)

UNDER THE SUPERVISION OF
Dr. Sonali Agarwal
Assistant Professor

INDIAN INSTITUTE OF INFORMATION TECHNOLOGY, ALLAHABAD

(A UNIVERSITY ESTABLISHED UNDER SEC.3 OF UGC ACT, 1956 VIDE NOTIFICATION NO. F.9-4/99-U.3 DATED 04.08.2000 OF THE GOVT. OF INDIA)
A CENTRE OF EXCELLENCE IN INFORMATION TECHNOLOGY ESTABLISHED BY GOVT.OF INDIA

November 2018

Contents

Abstract

Nowadays, there has been a lot of research work in the field of recommendation systems. They are used almost in every field from recommending news , movies, articles ,magazines,books to name a few. Different approach to recommender systems include content-based, collaborative filtering , knowledge-based, hybrid and also using reinforcement learning.In this work we try to solve the problem of news recommendation using online learning with the help of clustering the user contexts.We try to increase the click through rate for Yahoo! R6b dataset.The improvement of recommendation systems would help the organizations with improving the user experience thereby increasing the revenue for the organization.

Chapter 1

Introduction

In the recent times , there has been a lot of research work on the recommender systems. This recommendations can be of news, ads,movies,or various products sold on online shopping platforms.We will focus on news recommendation system in our work.The system will be able to recommend a user , a news article from a given pool of articles based on the user context. In this system it will be assumed that the users having same context will have same interests so , similar articles will be shown to them.

1.1 *Problem Definition*

The problem can be framed as follows:
Whenever a user visits the web-page , the system should display the relevant news article , based on past experience . This past experience can be "whether some user with similar interests or context , when shown an article , clicked or not". So , the article should be such that it maximizes the chance of getting clicked by the user. The systems' job is not yet finished at this stage , it should further take in account , whether it gets a positive or negative feedback to improve it's recommendation in the future(whether the user clicked or not).The only input to the system is the user context , based on which it has to display an article and get the user feedback(click). The system will always maintain a fixed pool of recent news articles, which would be updated regularly.

1.2 *Recommendation Systems*

The recommendation systems are such systems which recommend the users of a website the correct content which they wish to see or read. The recommendation system will only have some information of the user based on which it needs to recommend the content to the user [1]. These recommendation systems can improve the user experience of the website, thereby increasing the revenue for the concerned website or organization, it's good to have some effective recommender systems.These recommendation systems are used presently in different domains like recommending news articles, ad-recommendation, movie recommendation ,etc. While designing such recommender systems , designers face several issues and challenges that need proper attention.Formal definition of recommendation systems can be found in [1, 2].

1.2.1 Content-based Filtering

According to [3],"This technique works on the basis of user's test. So when a user creates his account system asks him to give ratings to different areas of his interests. Then based on these ratings system can recommend articles." For example if a user rates category of politics and entertainment, system will suggest him only about these category and user may never get suggestions from other categories such as sports and bollywood. This problem can be solved by the another technique called collaborative filtering.

1.2.2 Collaborative Filtering

According to [3],"The logic behind collaborative filtering is that if two or more users have same opinion or interst to a particular issue then they may have same intersta in other areas as well. These users make a pool of users of same interests." So in this technique user gets recommendation based on his/her neighbourhood. In this technique users can get recommendations from those category which were not rated by the user but rated by the users for community.

1.2.3 Demographic

The assumption behind this technique is that if two people have similar attributes like age,gender,occupation and physical attributes then they may have interst in same area.So based on demographic system can recommend users about their interests.[3]. In this method suggestions can be customized based on country language and age of customers. This technique can be helpful in marketing litreture but proper research is need to be done in this technique.

1.2.4 Knowledge-Based

This technique is also called contrait based filtering and works well for the system where the users do not visit very often. This algorithm takes into consideration about the specifications of items and matches them with user's interests. When a user creates an account system asks him about his needs and interests.[1, 2]. These techniques work well in initial times and need to be equipped with proper learning components otherwise they fail in longer period.

1.3 *Applications of Recommendation Systems*

The recommendation systems are used in vast domain in recent times. Some of the important domains are listed below [3]:
Nowadays recommendation systems are used widely.Such as ,Social-social networks as facebook,twitter. E-commerce-for buying books, cameras, laptops. Content-newspaper,email,documents. Entertainment-movies,games,music. Services-travel services,houses,matchmaking. Now recommender system is also used for recommending questions for students and recommending many policies for bank users, And it is of great use when someone is planning to work on same topic or area.

1.4 *Issues and Challenges*

1.4.1 Cold Start

The two cases of cold start problems are- 1. When a new user creates his account 2. When a new item is added to the catalogue In both the cases system can not recommend items to users correctly because system do not know the interests of users and area of item.This problem can be solved in many ways as asking the user about his interests on th time of profile creation or showing him items based on demographic.[1].

1.4.2 Context-Awareness

Most of the recommender systems try to recommend the relevent topics without taking additional information as age,time,location in account.But this plays an important role in recommendation systems and should be taken in account by the system to increase efficiency.There are many techniques as contextual pre-filtering,postfiltering and modeling for solving this issue.System can use some basic methods to improve efficiency in this problem like facial expressions, physiological signals and recording speech interpretation.
footnoteDefinition taken as quoted in [2].

1.4.3 Scalability

So when the data is huge it is very difficult for system to process such data. Sites likes Amazon.com have millions items and users.There are some approaches to solve this problem like clustering,reducing dimensionality and Bayesian Network.

1.4.4 Latency

Generally new items are added frequently in systems and it is not possible for recommender system to recommend them as they are not yet rated.So the

[1]as quoted in [2]

problem of Latency arises in this scenario and creates difficulty for real life applications.

Chapter 2

Literature Review

In the year 2006, Netflix company announced a competition for data scientists which was a breakthrough event in the field of recommender systems.It was important as it showcased the importance of recommending items to people and sped up the research for new machine learning algorithms.The winner of the competition used RMSE(Root mean square error) technique .[3]. Collaborative filtering as described in [4], works by exploiting the similarity between the users depending on their past history and visited web-pages.Thus, it works very well for situation where the previously liked or disliked web-pages influence the future web-clicks. The content based filtering technique which works by matching the interests of a user , like the previously visited web-pages , but in contrary with collaborative , this technique does not take into account the interests of other users.The items that are recommended are similar to the items liked in the past by the user.There is another technique which combined the previous two techniques known as hybrid approach [5]. This technique aims at tackling the problem of recommendation by unifying the advantages of both techniques the.But in many applications based on web , the content changes frequently,since users' interests are not static , they keep on changing. Also , what if the users' are entirely new , this problem known as cold start problem [6], these techniques would fall flat in such situation since they depend majorly on the past history and interests of the users. Hence , these type of problems are major problems. There has been a lot of research in recommendation systems in recent times.In [7] work, the use of contextual bandit problem is described for recommending the news articles . This method uses the contextual information of the users to increase the clicks . As a user arrives , based on it's preferences the algorithm selects the correct news articles , not only this , but also it notes down whether the user clicked on the article or not as a feedback. Moreover , a new algorithm LinUCB is proposed . The dataset used is Yahoo! today module .LinUCB for contextual MAB(Multi-Armed Bandits) , assumes the user contexts as binary vectors.Using the Bayesian networks framework , the different parameters are modified by Gaussian distribution. It heavily assumes that there exists a linear relationship between the user contexts

and their feedback. The computation cost would be higher if the complicated fitting model is used . The recommender systems also face issues like scalability. Also , experiments were done which compare different algorithms[8] . In [8] , the recommendation system for news articles, is discussed. The recommendation system is formulated as a multi-armed bandit problem . It works in two phases , in first phase, the users are clustered using k-means algorithm offline, and in second phase , using online learning, the system learns to predict the news articles preferred by the clusters. It is assumed that the users with same contexts or features will have interests. Whenever a user arrives, the input to the system is the user-context, based on this information , the system finds the cluster having similar features as that of the user, and then Upper-confidence bound algorithm , it displays an article. If the user clicks the displayed article ,the system receives a numerical reward of 1 and if the user does not click , there is no reward . This reward system helps the recommender , to change it's policy , so that in the future it can improve it's recommendation.The data-set used is artificially generated.The contexts of users are generated randomly which have 50 features that are either 0 or 1 value. It uses an offline k-means algorithm to group the users based on their similarity in contexts into 10 groups or clusters with 10 articles.
In [8], in clustering phase, once the clusters are formed initially, afterwards , the users are only assigned a cluster based on their similarity , but in real world scenario , it might be the case , that there is separate group of users for which a separate cluster must be formed for correct operation of the recommender system.
In [9] , instead of clustering the users based on the contexts, the items are clustered using a cluster-tree. Whenever a user arrives , the system tries to find the nearest cluster of items, based on the context of the user , and then randomly suggests an article from that item cluster.
This algorithm is suitable for situation , when the items or articles to recommend do not change with time . But , in real world scenarios like recommending news articles, where the news articles keep on coming , it becomes infeasible to maintain and update item cluster tree . Moreover, it is also not necessary that the items will have a similarity based on which they can be clustered. So, in such scenarios this algorithm might not perform well.

In [10] ,Bayes method for learning is used in contextual bandit problems.It can easily exploit the huge contextual information efficiently. This technique won the second rank in the Challenge on Yahoo! dataset at ICML 2012 Workshop new Challenges for Exploration and Exploitation 3. This technique's complexity is linear in the number of binary features and also linear in the number of arms. These are a significant advantages over other algorithms.

Till now , all the discussed recommender systems only recommend one item to the user. This can be extended to a list of items that are recommended to the user. This is discussed in [11] . This paper describes recommender systems that using the user feed-backs can work in situations that change or in other words

are dynamic and provide the recommendation to the users. In these situations , the feedback feedback is given through the clicks.This response or feedback is used to improve the future click-through rate.Here they came up with an algorithm named multi-objective ranked bandits.
[12] paper uses these ensemble algorithms. They used two data sets which include Yahoo's news recommendation data and online advertising data of KDD 2012.

Since , in this problem we will be dealing with continuous data streams, it becomes necessary to review the different data-stream clustering algorithms. In [13] , a survey of algorithms,challenges and issues related to data stream clustering is done.The volume increase in current data and shortage of memory to preserve it conduct us to the dynamic processing data and extracting knowledge from it.This method converts data as a stream of data which assert from one side and goes to other side so it causes unavailability to visit data twice. This behaviour of stream data creates some troubles. The Two troubles associated with this behaviour of stream includes: 1) Scanning of processing data stream occurs only one time, 2) During the time the changes happen in progressive stream and concept which data includes can be restrained or immediate.So, according to [13] the data stream algorithms should be able to

- Identify the changes as soon as it take place.
- Identify both restrained and immediate changes equally.
- Differentiate between noise and real drift.
- Forget obsolete cluster and example.
- Recall few old clusters and examples. Second trouble related to efficiency.

Due to this clustering should be processed only one time over the incoming data and should be in the possibly varying inter-arrival times. Since,limitation of memory makes impossible to store whole data.We are also dealed with the challenge to maintain a current result which can be presented to the user at any given time for clustering . [14] provides a algorithm that ables to adapt to the speed of the data stream automatically and also it is parameter free. This causes the best utilization of the time available under the current constraints or parameters to generate a clustering of the objects up to that point. This [14] method includes the object's age to reflects the greater importance of more new data. For the purpose of effective and efficient handling, [14] introduces the ClusTree that based on a compact representation of data and adaptive structure to maintain stream summaries. Furthermore option to manage very fast streams through aggregation mechanisms and propose novel descent strategies that optimize the clustering result on slower streams .ClusTree [14] supports any-time clustering.
For the study of exploration-exploitation tradeoff in reinforcement learning Multi armed bandit problem is an efficient model. Even though many algorithms

are well-examined theoretically and practically for the problem. [15] gives a provisional study of the most acclaimed multi-armed bandit algorithms. The early four algorithms are e-greedy, Boltzmann exploration, pursuit, and reinforcement comparison. all these specific techniques propose distinct ideas on handling the exploration/exploitation tradeoff in reinforcement learing. While, the UCB1 and UCB1-Tuned algorithms, basically situated on sophisticated mathematical ideas that are exploited to give strong theoretical confirmation on the expected regret.

The combination of similar objects allows assertion of groups in place of single objects for even faster processing and settings, [14] uses alternative assertion techniques for slow stram that exploit possible idle times of the algorithm to provide more accurate and optimize clustering output. It can be shown that the method of ClusTree is able to maintain the exact number of of micro clusters at stream speeds that for granularity is exponential w.r.t. competing approaches for equal stream speed and that are faster by orders of magnitude. Moreover, Clustree is able to detect clusters of arbitrary shape.

Title/Broad Area including year of publications, etc.	**Details about Frameworks / Algorithms etc.**	**Details about Tools, Datasets etc.**	**Summary of the research outcome**
Online Learning-based Clustering Approach for News Recommendation Systems.	Online Learning K-Means Clustering, ϵ-greedy and UCB1 strategy		As in the paper, it only store the summary of cluster information and the cumulative number of selected arms, instead of making note of whole history of users feedback and contexts. On the basis of summarized cluster information, OL-KMC reconstruct the clusters from the users feedback and perform as an adaptive learning algorithm.
A Contextual-Bandit Approach to Personalized News Article Recommendation (2012)	ϵ-greedy, LinUCB with Disjoint Linear Models, LinUCB with Hybrid Linear Models	Yahoo! Today Module dataset, based on the webpage recommendations of Yahoo! Front Page	This paper provides a simple and efficient method for evaluating bandit algorithms directly from CTR. It concludes that upper confidence bound (UCB) methods generally outperform the unguided ϵ-greedy methods. THe LinUCB algorithm shows its effectiveness in personalized web services when data is sparse and large numbers of contents is available.
Linear Bayes policy for learning in contextual-bandits	Linear Bayes' Method, Contextual and Non-contextual bandits	Yahoo! Front Page Today Module obtained from the "User Click Log Dataset" (Yahoo! Academic Relations, 2012).	This paper presented a empirical Bayes-like method for learning in contextual-bandit problems. It shows that the Bayes' method has better result than a simple contextual bandit and simple contextual bandit has better than that of a simple non-contextual bandit as well. And user-click behaviour is independent of the contextual information.
Data Stream Clustering: Challenges and Issues (2010)	Data Stream Clustering Methods and K-Means Clustering Algorithms		This paper highlights the main problems in data stream clustering is evolving data, which is more difficult to detect therefore unsupervised methods would be required. Some of other major issues are 'once' visited data and space limitations.
Online Learning in Large-Scale Contextual Recommender Systems(2016)	Adaptive Clustering Recommendation (ACR) algorithm, multi-armed bandits algorithms and online learning	Yahoo! Today Module dataset, based on the webpage recommendations of Yahoo! Front Page.	The proposed ACR algorithm increases the learning speed of RS by making use of an adaptive item clustering method. It also provides the solution of cold start problem by providing different contents to different types of users .

Recommender Systems: Issues, Challenges, and Research Opportunities	Content-based filtering, Collaborative filtering, Hybrid filtering		This article focused on prominent issues and challenges which need proper attention that are sparsity, latency, grey sheep, cold-start problem, context-awareness and discussed what has to be done mitigate these issues.
Ensemble Contextual Bandits for Personalized Recommendation (2014)	HyperTS and HyperTSFB ensemble bandit algorithms are used for solving the contextual recommendation problem in the cold-start situation.	News recommendation data (Yahoo! Today News module, based on the webpage recommendations of Yahoo! Front Page) and Online advertising data from KDD Cup 2012	This paper uses these ensemble algorithms to distribute the trials to the base bandit policies. It explores these contextual bandit algorithms to obtain strong click through rate (CTR) of web contents. It employs a meta-bandit paradigm in which a hyper bandit is used with base bandits, to explicitly explore/exploit the base bandits based on user feedbacks.
Recommendation systems: Principles, methods and evaluation (2015)	Content-based filtering, Collaborative filtering and Hybrid filtering techniques		This paper discussed the two traditional recommendation strategies with diverse kind of hybridization techniques to improve their performances. Some other learning algorithms are used to measuring the quality and performance of discussed recommendation algorithms.
Multi-Objective Ranked Bandits for Recommender Systems (2016)	MO-MAB algorithms	Yahoo! Today News dataset and KDD Cup 2012 online ads dataset	This paper explores a multi-objective ranking bandits (MO-MAB) algorithm for online recommendation. The algorithm has four main parts: a set of recommendation quality metrics, a scalarization function, a dynamic prioritization scheme for weighting these metrics and a base Multi Armed Bandit algorithm. It also present gains in CTR (Click Through Rate) performance over baselines when multiple objectives are considered simultaneously.
Algorithms for the multi-armed bandit problem (2000)	ϵ-greedy, Boltzmann exploration, pursuit, UCB1, UCB1-Tuned and reinforcement comparison.		It presents a deep study of all the mentioned most popular multi-armed bandit algorithms.
Recommender Systems: Introduction and Challenges (2015)	Content-Based, Collaborative Filtering, Demographic, Knowledge-Based, Community-Based and Hybrid Recommender Systems.		It introduces all the core recommendation technique used to generate the recommendations. These all algorithms are customized to provide useful and effective suggestions for that specific type of item, what to read ,what to watch and what to buy etc.

Chapter 3

Proposed Methodology

Our approach to this problem is to formulate the problem as a "Multi-Armed Bandit" frame-work as described in [8]. Our approach will be a two phase approach. In first phase , we will cluster the users based on their contexts using a stream clustering algorithm since the user-context would arrive as a continuous stream of data . After the clusters are formed in the first phase , each cluster has its knowledge and can maintain an online learning process.In a way this is a collaborative approach combined with contextual-bandits approach. In the following subsection we first provide an introduction of multi-armed bandits and then provide the formulation of our problem in terms of multi-armed bandits.

3.1 *Multi-Armed Bandits and problem formulation*

In the early days, there were machines called bandits, which had a number/numbers of levers which could be pulled and based on some unknown distribution, the player received some reward(generally coins/cash).Such machines having k-levers are called k-armed bandits.Some machines have only one lever , but the player then has to choose among k such machines each having some unknown reward distribution.The problem of maximizing the reward obtained is deeply studied in [16] using reinforcement learning algorithms.Our problem can also be modelled as k-armed bandit problem , where a machine is equivalent to the group of targeted users and k levers or arms the corresponding news articles which we need to recommend. In our problem , our recommender system needs to learn the unknown distribution of user feedback(user clicks) ,i.e on showing a particular news article(on pulling a particular lever), whether the user clicks on that article or not(whether the reward is recieved or not). Here the reward would be +1 for user click and no reward otherwise.
Here we have assumed that there is only one machine having k-arms(only one type of user groups and the corresponding news articles preferred by them(k-arms).But, there would different types of people in real life having different

interests, who would have a completely different set of preferred articles. So , to capture this scenario , we would need to maintain K number of machines(K number of user groups or clusters) having k-arms(k news news articles) . For this we require clustering the users based on their interests or contexts , who would have their own set or pool of articles .These clusters would correspond to an independent machine having k-arms.
We would be using the Upper Confidence Bound Algorithm described in [8]. In our problem , we would try to maximize the click-through rate (number of clicks per recommendations of an article).So , in our multi-armed bandit problem the reward would be user clicks, more the reward more would be the click-through rate.

3.2 *UCB1 Algorithm*

The terminology is as follows:

- The system would receive user-context at discrete time t x_t.
- A_C is the set of news for the cluster C.
- a_t is a news article.
- r_{t,a_t} is the reward which is clicked or not. The reward r_{t,a_t} follows unknown distribution.

Algorithm 1 UCB1 Algorithm

```
1: Parameters: α > 0
2: for news a_t ∈ A_C do
3:     μ_{a_t} = r_{a_t}/n_{a_t} + α sqrt(2 log n_C / n_{a_t})
4: end for
5: a*_t ← arg max_{a_t ∈ A_C} μ_{a_t}
6: n_{a*_t} ← n_{a*_t} + 1
```

The expected total regret [8] is given by the below formula:

$$R_A(T) = E[\sum_{t=1}^{T} r^*_{t,a_t}] - E[\sum_{t=1}^{T} r_{t,a_t}]$$

The cluster C is the input for the above algorithm which is formed by using online K-means algorithm which is described below.

3.3 *Online K-means*

The online k-means algorithm is a clustering algorithm . It updates the means as soon as the data point arrives. Hence the name online k-means. In this algorithm the input is the number of clusters and their initial means or centres. Now these means are selected randomly and they should be distinct.Initial means are nothing but the data points itself. Thus, we need k representative data points which will represent their respective cluster and act as their mean. Now , after choosing these points , the algorithm iteratively updates it's clusters mean . Whenever a new point arrives , the algorithm compares this point with k means at that moment of time .More specifically it compares the similarity between them and chooses that point which is the most similar and assigns this point to that cluster as well as updates the mean of that cluster. Now there are different similarity indices like jaccard index, manhattan distance, hamming distance , etc. Here we will be using the euclid distance . But in our case the the euclid distance is same as manhattan distance or hamming distance , since we are using binary data .

3.4 *Combining UCB1 with Online K-means*

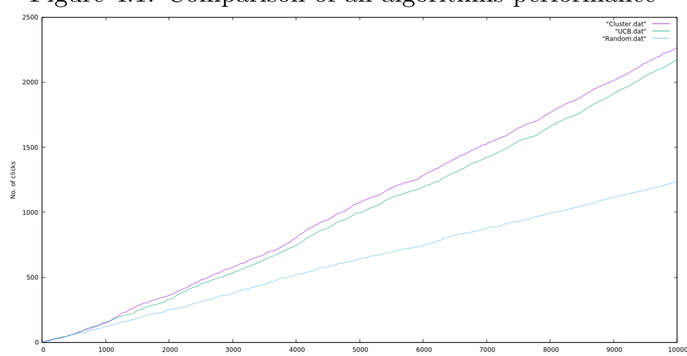
We can combine the UCB1 algorithm with Online K-means and model it as contextual armed bandit problem . It is contextual because now we will have different rewards for recommending different articles to different cluster , using this we are able to target specific groups of people having similar interests and also take into consideration their response . We can think each cluster as different bandit having it's own k-arms which are nothing but news articles to recommend.

Chapter 4

Experimental Results

4.1 *Comparison of algorithms*

Figure 4.1: Comparison of all algorithms performance



As can be seen from the figure above , it is clear that after sufficient number of recommendations , the UCB1 combined with online k-means outperforms other two algorithms , which are random and simple UCB1 without clustering or the context-free bandit algorithm. The worst performance is by the random algorithm , while UCB1 is better and the UCB1 with online k-means is the best. The dataset of 10000 entries is randomly generated . Each entry has first 10 columns which are user's attributes . These attributes have binary values. Similarly the next ten columns are also binary , which describes whether the user will click if a particular article is shown . 1 means that the user will click and 0 means the user won't click .

Chapter 5

Conclusion

From our experimental results , we conclude that indeed the method of combining the UCB1 algorithm with online k-means increses the click-through rate , though not much . Also , it is far better than the blind random recommendation , almost double the click-through rate.

Bibliography

[1] Folasade Isinkaye, Yetunde Folajimi, and Bolanle Ojokoh. Recommendation systems: Principles, methods and evaluation. 16, 08 2015.

[2] Shah Khusro, Zafar Ali, and Irfan Ullah. Recommender systems: Issues, challenges, and research opportunities. pages 1179–1189, 02 2016.

[3] Francesco Ricci, Lior Rokach, and Bracha Shapira. *Recommender Systems Handbook*. Springer Publishing Company, Incorporated, 2nd edition, 2015.

[4] J. Ben Schafer, Joseph Konstan, and John Riedl. Recommender systems in e-commerce. In *Proceedings of the 1st ACM Conference on Electronic Commerce*, EC '99, pages 158–166, New York, NY, USA, 1999. ACM.

[5] Sarabjot Singh Anand and Bamshad Mobasher. Intelligent techniques for web personalization. In *Proceedings of the 2003 International Conference on Intelligent Techniques for Web Personalization*, ITWP'03, pages 1–36, Berlin, Heidelberg, 2005. Springer-Verlag.

[6] Seung-Taek Park, David Pennock, Omid Madani, Nathan Good, and Dennis DeCoste. Naïve filterbots for robust cold-start recommendations. In *Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '06, pages 699–705, New York, NY, USA, 2006. ACM.

[7] Lihong Li, Wei Chu, John Langford, and Robert E. Schapire. A contextual-bandit approach to personalized news article recommendation. 02 2010.

[8] Minh N. H. Nguyen, Chuan Pham, Jae Hyeok Son, and Choong Seon Hong. Online learning-based clustering approach for news recommendation systems. *2016 18th Asia-Pacific Network Operations and Management Symposium (APNOMS)*, pages 1–4, 2016.

[9] L. Song, C. Tekin, and M. van der Schaar. Clustering based online learning in recommender systems: A bandit approach. In *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 4528–4532, May 2014.

[10] Jose Antonio Martin H and Ana Vargas. Linear bayes policy for learning in contextual-bandits. 40:74007406, 12 2013.

[11] Anisio Lacerda. Multi-objective ranked bandits for recommender systems. *Neurocomput.*, 246(C):12–24, July 2017.

[12] Liang Tang, Yexi Jiang, Lei Li, and Tao Li. Ensemble contextual bandits for personalized recommendation. In *Proceedings of the 8th ACM Conference on Recommender Systems*, RecSys '14, pages 73–80, New York, NY, USA, 2014. ACM.

[13] Madjid Khalilian and Norwati Mustapha. Data stream clustering: Challenges and issues. *CoRR*, abs/1006.5261, 2010.

[14] Philipp Kranen, Ira Assent, Corinna Baldauf, and Thomas Seidl. The clustree: Indexing micro-clusters for anytime stream mining. *Knowl. Inf. Syst.*, 29(2):249–272, November 2011.

[15] Volodymyr Kuleshov, Doina Precup, and Leslie Pack Kaelbling. Algorithms for the multi-armed bandit problem. 2010.

[16] Jeffrey D. Johnson, Jinghong Li, and Zengshi Chen. Reinforcement learning: An introduction: R.s. sutton, a.g. barto. 35:205–206, 11 2000.