Journal of Algorithms and Computations

journal homepage: http://jac.ut.ac.ir

A Recommender System Based on Markov Process Using Web Usage Mining Method and Neural Network

H. Shafaei[1], R. Razavi[2]

[1] Department of Computer Engineering, Tehran Science and Research Branch, Islamic Azad University, Iran

[2] Department of Management and Information Systems, Kent State University, New York, United States

ABSTRACT

Web recommender systems provide the most appropriate recommendations by analyzing user's navigation behavior. This recommender system can be considered in different cases, such as e-commerce, search engines, etc. The aim of the proposed approach is to create users' profiles and find their common navigation patterns implicitly. The web log file is utilized to analyze browsing history and discover users' navigation models. This research used K-means for the clustering and a neural network is designed for classification, along with the recommended Markov model. The novelty of this research is considering a threshold for the proposed Markov model. The efficiency (accuracy and precision) of the proposed method is confirmed in a test case.

Keywords: Keywords: Web usage mining, k-means algorithm, neural network, Markov model, Recommender systems Controller Oscillator

ARTICLE INFO

Article history:
Research paper
Received 22, February 2024
Accepted 18 April 2024
Available online 05, August 2024

AMS subject classification: 68T09.

* Corresponding author: H. Shafaei, Email: Shafaei_h@ymail.com

† rrazvi@kent.edu

1 Introduction

Web recommender systems provide the most appropriate recommendations by analyzing user’s navigation behavior, and helping web users to access their favorite pages from a wide range of options. The analysis of user’s navigation behavior is done using web mining methods, with web usage mining being the most common method and widely used method for extracting user’s behavior patterns from log files. Web personalization based on web mining consists of two main steps. In the first step, which is done offline, the system is trained by using data taken from users’ behavior on the web, this training helps access’s patterns and extract user models. In the second step, which is done online, the extracted model from the first step is used to interpret and compare users’ navigation patterns and recommendations are provided based on this comparison. The goal of web personalization based on web usage mining is to recommend a set of objects to the current user, including links, ads, text, products, etc., with an orientation towards the user’s preferences and interests [1]. There are several types of recommender systems used to provide recommendations to users. These are some of the main types of recommender systems, and often, a combination of approaches is used to enhance recommendation accuracy and diversity. Information filtering or recommender system (RS) has become essential and important in social network applications and electronic commerce. Websites like Amazon and YouTube have used recommender systems to suggest personalized products and services. Useful advice can cause user adherence [2]. Web mining is divided into three parts: Web usage mining captures user’s activities and searches user’s behavior and navigation patterns. Web structure mining discovers knowledge from hyperlinks, including outlinks and inlinks of web pages. Web content mining extracts web page contents including text, images, etc. [3]. Web usage mining is the most common and widely used method for extracting users' behavior models from these files. Web usage mining focuses on techniques that can predict users' behavior while interacting on the web. The main task of web usage mining is to retrieve useful information from the cache of usage profiles of web servers based on user searches. The neural network is composed of interconnected processing units called neurons. The connection of the neurons is established using link. The value of these links is called the connection weights [4]. After building the mean vector of each cluster, the neural network is used to find most similar cluster to the user's session in the online phase. When entering the current user session, it can be detected through a trained neural network that which cluster this session belongs to. In fact, the mean vector of each cluster is considered as the input of the neural network and the number of the most similar cluster to the current user session is considered as the output of the neural network. Then appropriate pages are recommended to users by utilizing the neural network for the classification of users, and then the Markov model inside the clusters. Obviously, it is expected that the Markov process considers sequence of visited pages by users. In other words, a new user visits a relevant page with respect to the created tree by the Markov model. The usage of Markov models is for analyzing and discovering user’s navigation patterns and these states show pages while the links between states show the transition probability [5]. Today, a large number of pages are observed on the web. Users get confused between these different pages. Therefore, it is necessary to design a recommender system that suggests appropriate pages based on users' behavior. A recommender system is a technology that provides users with more accurate recommendations, but since users have different needs, the accuracy of the recommendation alone is not enough for the recommender system. Our proposed method provides greater accuracy and precision. In the present research, we propose a hybrid recommender system based on the Markov process, utilizing web usage mining and neural network. Initially, user behavior on the web is used to implicitly generate their features in the form of a profile. These features encompass the date of page visits, the frequency of page visits within a session and the duration of the user's presence on the page. In the next step, users are clustered. The optimized k-means cluster is used to cluster user sessions. Then, the neural network is used to train clusters, and the Markov model is used to correct pages suggested to users. The Markov process is expected to consider the sequence of pages that users visit. In other words, the new user who enters will see the related page according to the pattern created by the Markov model. Compared to previous researches, in this study expresses:

- We represent an optimized clustering, to reduce the overhead of offline modeling, runtime, and computational costs.
- We use the neural network to find the most similar cluster to the user session in the online phase, which can perform a task with high complexity with better confidence.
- We present the proposed Markov model, and accurately predict the next page visited by considering the threshold.
- Our solution improves the accuracy and precision of the recommender system by providing a proposed model.

Overall, the web usage mining, Markov models, and neural networks in a recommender system aim to provide personalized, accurate, and engaging recommendations. The advantages of incorporating a usage threshold for Markov models:

1. Improved Accuracy: By setting a usage threshold, the Markov model can ensure that it has sufficient data or observations to make reliable predictions. This leads to increased accuracy in the model's outputs, as it avoids making premature or erroneous predictions based on limited data.

2. Enhanced Predictive Capabilities: A usage threshold allows the model to capture and analyze a significant amount of relevant data. As a result, the Markov model becomes more effective in capturing patterns, dependencies, and transitions, leading to improved predictive capabilities.

3. Resource Optimization: By establishing a usage threshold, resources can be allocated efficiently. It prevents unnecessary processing and training of the model when data is insufficient or when the model is not yet adequately trained. This optimization of resources is particularly important in scenarios where computational power and time are limited.

4. Decision Confidence: With a usage threshold, decision-makers can have greater confidence in the predictions derived from the Markov model. They can rely on the model once it has surpassed the usage threshold, knowing that been trained on enough data to generate reliable insights.

5. Scalability: The usage threshold facilitates the scalability of the Markov model. As more data becomes available and the model accumulates a higher usage count, it becomes more robust and adaptable to new patterns and scenarios. This scalability is beneficial when dealing with evolving data and changing prediction requirements.

This present research is organized as follows, In Section 2, the required background is represented. Related works is introduced in section 3. The proposed method is described in section 4. In section

5, two criteria for estimating the proposed method are explained, also, in section 6, the evaluation of suggested method along with a discussion is represented. Finally, Section 7, concludes the paper and future works.

2 Background

Web mining is an interdisciplinary field consisting of data mining, statistics, machine learning, information retrieval, and network analysis.

It aims to address the challenges of extracting knowledge from web-related data, including web pages, links, user behavior, and other web-related entities.
Web mining is used for massive web data warehouses [6]. Web mining includes analysis of web server logs where data mining is used for finding relationships in large amounts of data. There are approximately three model discovery domains, that are related to web mining: Web content mining; web structure mining; and web usage mining [7].

Web usage mining is also known as log mining, which extracts patterns of interest in web access logs. The web personalization process based on web usage mining involves three steps: pre-preparation, pattern discovery, and recommendation [8].
The recommender systems are divided into two main types: personalized recommender systems and non-personalized recommender systems. Personalized recommender system models Extract user's profile information and history to generate recommendations. In contrast, the non-personalized recommender system generates a list of offers based on their popularity. Hence, all users receive a similar type of recommendation [9]. Personalized Recommender systems are classified into two groups: collaborative filters and content-based filters. Collaborative filters use a filtering method based on user items analysis or log of previous purchases [10].

3 Previous works

In this section, we categorize the recommender systems based on different techniques.

3.1 Previous works

In [11], the users on the site are clustered, and then users are placed on the same navigation path in the same cluster. The clustering approach is model-based, and web users are categorized according to the order of web page request [11]. The research [12] focuses on providing real-time dynamic recommendations for all website visitors regardless of whether they are registered or not registered [12]. In research [13], a bitwise view is given for web pages with large but scattered web data, then the display can be used with bitwise serial or parallel association algorithms [13]. A new clustering method using k-means ++ is offered for web source codes [14]. In this research, architecture is provided to integrate information about the products with web log data and generate a list of recommended products by using the LCS algorithm [15]. In this research, it assumed that users with similar behaviours have similar interests in an item, then hierarchical clustering technology is used to cluster users according to their profiles [16]. In research [17], Clustering of users is used to reconstruct a user-item bipartite network, which is significantly improved as network density [17].

3.2 Recommender system based on association rules

In research [18], a personalization framework is offered using a combination of methods traditional personalization system. To discover object association's rules, similar users of web object access data [20]. This research recommends a new clustering method called the HBM algorithm for user clustering based on time-based framed browsing sessions [19]. The established rules are based on the evaluation support, confidence, and lift criteria. According to final predictions, being interested in pages is essentially dependent on lift and support criteria, while confidence results in an equal amount on all pages [20]. The research [21] presents a recommendation system for online course learners. The recommender system proposes learning resources, such as relevant lessons, using association rules, content filtering, and collaborative filtering in formal online courses [21]. In research [22], the original data is cleared and pre-processed. Then, the association rules model, law model, and user value analysis model are used. Then, an Apriority algorithm is used to analyze the relationships between the user's previous access records and the user group of the K-means algorithm. Finally, the results of the experiment show that the output of association rules and the clustering analysis are meaningful [22].

3.3 Recommender system based on association rules

In this research, a new version of the web page recommendation model in web mining based on the KNN page ranking algorithm is introduced [23]. The purpose of this research is to identify the user's interest behavior and not the user's interest behavior classification [24]. In research [25], the recommender system uses a similarity measurement in KNN memory based on collaborative filtering techniques [25]. The mantle algorithm, which is a type of traditional genetic algorithm, is employed to determine the structure. It utilizes local search to expedite the attainment of an optimal response. Genetic algorithms are designed to explore the search space, while the local search explores the neighborhood of each solution generated by the genetic algorithm to find improved solutions [26]. In research [27], a new recommendation algorithm called the joined session-based context-aware recommendation model is proposed. This model involves mapping contextual information into low-dimensional real vector features and integrating them into a recommendation model based on recurrent neural networks [27]. This paper proposes a method that combines Recurrent Neural Network (RNN), Reinforcement Learning (RL), and Generative Adversarial Networks (GANs). By employing RL, the research leverages users' immediate feedback and simultaneously utilizes GAN to generate the necessary training data for RL [28]. In research [29], eight machine learning models are applied to different datasets from various domains to evaluate their performance and compare the results. The findings indicate that session-based KNN (SKNN) and its variants show promising results compared to other techniques [29]. In research [30], the focus is on exploring user-item relationships with multiple types of behaviors. The authors propose a new approach called the multi-behavior graph neural network (MBRec), which takes into account diverse interaction patterns and interdependencies among different behavior types. The framework employs a graph-structured learning approach to model high-order connectivity in a behavior-aware user-item interaction graph [30].

3.4 Recommender system based on Markov model

In study [31], profiling a web user in big data is examined. To improve the efficiency of profiling, the concept of contextual credibility is proposed. Human knowledge is represented by Markov logic statements. The MagicFG method is applied in an online aminer.org system to profile millions of researchers [31]. The research of [32] suggests a new hierarchical Markov model to identify changes in user referrals over time using latent text modeling [32]. In research [33], a collaborative recommender system for BI interactions is presented, and its advantage is the identification of user's interest. Users' interests are gained through the intent to interaction with the BI system [33]. In research [34] a recommendation framework is presented that takes into account the dynamics of user preferences. An approach based on Hidden Markov Model (HMM) is proposed to identify change points in a sequence of user interactions that reflect significant changes in preference according to the sequential behavior of all users in the data [34]. The research of [35] obtains accuracy of the Markov OP using updates from activity diagrams in maintenance and development of application [35]. In research [36], different types of datasets are used for analysis, the effectiveness of the hybrid model is compared using parameters such as accuracy, prediction, incorrect prediction [36]. A three-stage approach to GRS is devised, involving: 1) the utilization of three binary matrix factorization methods, 2) the development of an influence graph incorporating assertiveness and cooperativeness as personality traits, and 3) the application of an opinion dynamics model to achieve consensus [37]. To address these issues, this study proposes a collaborative filtering model that incorporates user and item feature information, along with ratings, to estimate preferences [38]. The research of [39] introduces the concept of intentions, which consist of categories and actions, to address these data issues. The intention-aware Markov chain-based sequential recommendation model (IMRec) is offered [39].

3.5 Hybrid Recommender system

In research [40], Markov models and the Bayesian theorem are combined, so that a two-level prediction model is designed [40]. In research [41], a hybrid model of the neural network and hidden Markov model is presented for time-sensitive systems [41]. In study [42], a Deep Learning model named RHMM was introduced for recommender systems. It incorporates the Hidden Markov Model and Artificial Neural Networks [42]. A personalized tourism recommendation system offers convenient and affordable travel information to individuals and also groups [43]. This study focuses on an e-commerce intelligent recommendation system (IRS) based on deep learning. Firstly, the overall design of the e-commerce recommendation system is presented, including the functional modules and system architecture of the e-commerce IRS [44]. This research aims to assess the effectiveness of deep learning networks and deep transformer models in review-based recommender systems [45]. The paper discusses research trends and future directions, concluding that POI recommender systems based on deep learning hold promise as a focus for future work in this field [46].

The brief summary for section related works is shown in Table 1:

Table 1. Comparison of parameters used in related work

Ref.	Accuracy	Precision	Recall	F-measure	Coverage
Aghdam, et al., 2018	-	√	√	√	-
Cheriyan, et al., 2017	√	-	-	√	√
Drushku, et al., 2018	√	-	-	-	√
Eskandanian, et al.,2019	√	-	-	-	-
Gu, et al,2018	-	√	√	√	-
Jindal, et al.,2019	√	-	-	-	-
Karami, et al.,2017	√	-	-	-	-
Khalil, et al., 2006	-	√	-	-	-
Lopes, et al., 2015	√	√	√	-	-
Narvekar, et al., 2015	√	√	-	-	-
Sneha, et al., 2011	-	√	√	√	-
Wang, et al., 2020	-	-	-	-	√
Xiao, et al., 2017	-	√	√	√	-
Yongqin, et al., 2018	√	-	-	√	-
Nigam, et al., 2020	√	-	-	-	-
Djellali, et al.,2020	-	√	-	√	-
Nan, et al.,2022	√	-	-	-	-
Zhao, et al.,2022	√	-	-	-	-
Abolghasemi, et al.,2022	-	√	-	-	-
Pujahari, et al.,2023	√	-	-	-	-
Safavi, et al.,2022	-	√	√	-	-

4 The Proposed method

The presented method is a web usage-based web mining. Therefore, the recommender of this research is created as follows: First, web logging is preprocessed, and then user sessions are extracted from the dataset. A session vector is created in this step, and then an algorithm is used to cluster users' sessions. A classification algorithm is used for data training, and then the recommended Markov model is applied considering the threshold for prediction of web pages. The preprocessing operation is performed on the web log file and user sessions are extracted. To create user sessions, three features of frequency, date, and duration of attendance at the session are applied to weighting the pages, and user sessions are divided into vectors with weighted pages. Because the ratio of the importance of frequency rate and duration of visit's page are equal, Therefore, the harmonic mean of these two metrics is used for weighting the pages.

The following two rules are considered to identify the session:

A. Rule of thirty minutes: If the total of user input is less than one minute, it is considered as a session.

B. Rule of Ten minutes: If rule 1 is not set, it should be considered a session until the time distance between any request is less than ten minutes.

Users usually do different activities at a period, and they may do multiple searches and activities at a period. Therefore, it is necessary to analyze the user's behavior in more detail. Because recently accessed web pages show users new interests better than previously visited pages. For this reason, to create a profile according to the sequence of access to web pages, increment linear weight is given to pages, and pages related to the latest user's sessions have a higher coefficient than previous sessions. Proposed method is shown in Figure 1:

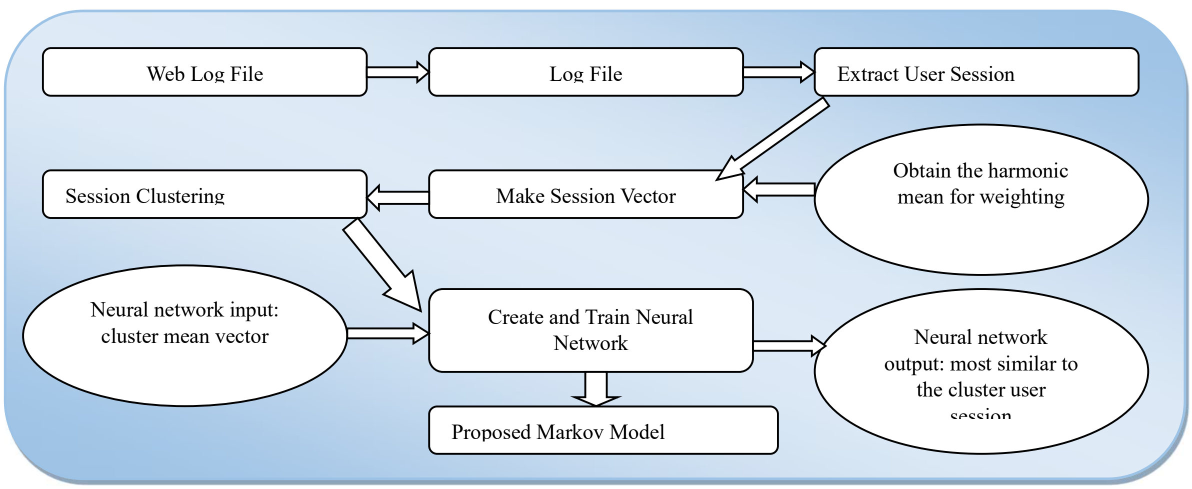

Figure 1. Proposed method

The use of common K-means methods increases the modeling overhead in the offline phase. Therefore, the optimized K-means algorithm is used to cluster the sessions. The initial clusters are randomly generated. Different results are obtained for each execution time of the K-means algorithm with different values of k. By selecting different values for k and analyzing the results, the best number of clusters is obtained per k=30. For this number of clusters, the accuracy of the result increases. The output of this step is the mean vectors of each cluster, which is considered as one of the main inputs of the neural network.

4.1 Pre-processing web logs

Pre-processing operation is done on web logs to identify user sessions. Dataset mining, data cleansing, user identification, and session identification are performed during this operation. This research used two NASA datasets and the University of Saskatchewan (UOFS) dataset. Details of this datasets are presented in Table 2.

Table 2. Details of data set used our research

	NASA	UOFS
Total log records	1891714	2408625
Total sessions	118718	172984
Average of session length	6.4	5.5
Number of pages	1005	5423
Dataset date (year/month)	1995/7	1995/6-12

4.2 Session clustering

K-means is used as a clustering strategy to reduce the training set; thus, it reduces algorithm execution time and computational cost. The use of K-means clustering methods increases the overload of the offline modeling. Therefore, in this research, the optimized K-means method is used to cluster sessions. The improved method causes an increase in speed. To cluster user sessions, at first, each session is converted to a vector of numbers ,and to increase the work speed, the fraction denominator in cosine similarity is stored in this array for each session. The square of the sum array is also placed at the end of each line of the array. In this research, after obtaining the interest of each user in pages, optimized K-means clustering is used to group users and cosine distance is used as a similarity metric. The steps are described below:

A. The matrix is obtained by the number of times that the user visits the page.
B. k is selected as the number of clusters.
C. The data set is equally divided into k parts and from each part a value is randomly selected as the center.
D. The cosine distance of these centers is obtained.
E. Similarity was more to each center, placed inside the it cluster.
F. When the first iteration is finished, the center will be updated. In order to, (the distance to the power 2) of each center from all the elements is obtained within the cluster, and at the end, the square root of that cluster is obtained. The new point is the center of the cluster.
G. For the second iteration, we need to calculate the cosine distance of all users from each center, and again, each user was closer to each cluster, add to that cluster.
H. We continue the steps until we reach a specific error, which is calculated.

I. After 5 to 10 times, if the error is not less than a determined value, we finish the clustering work.

This research uses an optimized K-means cluster for clustering the user sessions and considers the mean vector for each cluster. Primary clusters are randomly created. Different results are obtained at each time of the implemented K-means algorithm with different k values. The best number of clusters is obtained for k=30 by choosing different values for k and investigating results; hence, the precision and accuracy are increased for this number of clusters. The mean vectors of each cluster are the output of this step as one of the main inputs of the neural network.

4.3 Neural network for determination for corresponding cluster

The motivation to use the neural network comes from the fact that a neural network, unlike a linear program, can perform a task of high complexity, and the neural network provides better reliability due to its loop structure and parallel nature [47]. Neural networks are used as a classifier algorithm to train data. After creating the mean vector, a neural network is used to find the most similar cluster to the user's session in the online phase. Initially, the neural network is trained with each mean cluster. At the start of the current user session, we can determine which cluster this session belongs to. This is done by neural network training. In fact, the mean vector of each cluster is considered as the input of neural network, and the most similar cluster number to the current user session is considered the output of the neural network. A three-layer perceptron network with input, hidden and output layers is the applied neural network. In this research, 75% of the dataset is used for neural network training, and the remaining 25% for testing. The backpropagation algorithm is used to learn the network. A hidden layer is used in the network, and then it is trained in Matlab software with several different neurons. The sigmoid activation function is also used at the hidden and output layers. In the training phase Figure 2, the input of the neural network is a matrix that includes the mean vectors of the clusters. Each mean vector shows the motion models of the cluster users in a specific set of visited web pages. After building the neural network, it is determined which cluster each test session has the highest similarity to. Therefore, the output of the neural network is the number of clusters, that have sessions with the highest similarity to the test session being considered.

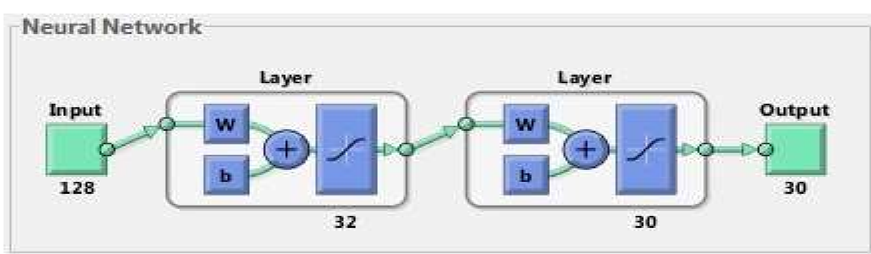

Figure 2. Neural network training

4.4 Recommended Markov Model

The innovation in this article refers to the usage threshold specifically for Markov models. In the context of Markov models, a usage threshold could be a predetermined level of data or usage required for the model to effectively capture and predict patterns. The usage threshold in a Markov model may be defined as the minimum amount of data or observations necessary for the model to produce reliable results. It helps ensure that the model is trained on a sufficient amount of data, thus improving its accuracy and predictive capabilities. Implementing a usage threshold in Markov models can have several benefits. It can help prevent premature or inaccurate predictions when the model doesn't have enough data to make reliable inferences. It can also aid in resource allocation and decision-making by providing a clear guideline on when the model is considered robust and can be relied upon. The recommended Markov Model is used to predict the next webpage. In many cases, the low-order (first) Markov models cannot accurately predict the next visited page by users because these models do not take a deep look at the user's past and only predict on the basis of a visited last page. Therefore, our recommended model does not consider the first-order Markov. The higher-order Markov models should be used for better precision. On the other hand, higher-order Markov model has limitations such as high complexity and many states. A method is recommended to solve this problem:

A) Calculation of All-Kth-Order Markov model for each cluster (in online mode), is shown in Table 3.

Table 3: Algorithm of All-Kth-Order Markov model [48]

Algorithm: All Kth Prediction
Input: user session, x, of length K
Output: Next page to be visited, p

1. $p \leftarrow predict(x, mk)$.
2. $If\ p\ is\ not\ 0\ then\ return\ p$
3. $x \leftarrow strip\ first\ page\ ID\ from\ x$
4. $K \leftarrow K-1$
5. $if (K=0)\ return\ 'failure'$
6. $Goto\ step\ 1$
7. $Stop$

B) Getting the best order within each cluster (in offline mode)
C) Considering a threshold (to determine that it continues to which Markov order in offline mode).
D) Calculation of threshold is as follows:

- Accuracy is obtained for each Markov order (for instance, 2, 3, 4, and 5).
- Threshold is equal to obtained mean Accuracy for each Markov order, is shown in Table 4.

Table 4: Algorithm Calculate Threshold

Algorithm: Calculate Threshold (Nth Markov,train)
1. double accuracy = 0
2. for i=2 to i ≤ Nth Markov
3. Hash(i,train)
4. accuracy = accuracy + Testi (train,test,i,i)
5. end for
6. accfinal + = (accuracy / (Nth Markov - 2))
7. threshold = (accfinal.Sum / accfinal.Count)
8. Stop

E) Accuracy number of each order is then compared with threshold number. If it is greater than threshold, then that order will be the best order and our target answer.

F) The proposed method is compared with the defined methods.

In online mode, the same order applies when cluster is specified.

5 Implementation

The recommender system aims to calculate a recommended set, rs, for the current user session, which has the highest match with the user's interest. This section is the only online component of the system and should be efficient and accurate. Therefore, any input vector in the test section is divided into two parts and it represents the user's interest. The first part of each vector is considered as the current user session, and the second part is considered as pages that should be recommended to users (rpt). Then, the output of the recommended system (rs) is compared with the second part, and the system efficiency is evaluated by obtaining the correct number of recommended pages (rp). The number of recommended pages and two metrics, precision and recall, are the parameters, that affect system efficiency.

5.1 Precision calculation method

Precision is the ability of the recommender system to generate precise recommendations. The precision of a recommendation is simply equal to the ratio of correct ones to the total number, as shown in equation 1.

According to the correct recommendation, if a session is generated based on the visited section, then the user session occurs.

$$Precision(rs, rp) = \frac{|rs \cap rp|}{rs} \tag{1}$$

Values of precision of comparative models for two datasets are shown in Table 5 and Table 6. As it is shown, in both tables our approach has the best results.

Table 5: Calculated Precision using of NASA Dataset

Rank	Our Proposed Method	Mamoun, et al., 2012	All-Kth-Markov	All-Kth-Dempster's Rule	Dempster's Rule	All-Kth-ARM	ARM	ANN
1	0.19	0.18	0.16	0.21	0.21	0.10	0.05	0.15
2	0.35	0.32	0.31	0.30	0.30	0.11	0.06	0.21
3	0.42	0.39	0.37	0.33	0.33	0.12	0.06	0.25
4	0.47	0.43	0.41	0.38	0.38	0.12	0.06	0.30
5	0.51	0.46	0.44	0.40	0.40	0.13	0.06	0.33
6	0.53	0.49	0.46	0.45	0.45	0.14	0.06	0.37
7	0.55	0.51	0.48	0.48	0.48	0.14	0.06	0.40
8	0.57	0.52	0.50	0.50	0.50	0.15	0.06	0.45
9	0.58	0.53	0.51	0.51	0.51	0.15	0.06	0.48
10	0.60	0.54	0.52	0.53	0.53	0.16	0.06	0.50

Table 6: Calculated Precision using of UOFS Dataset

Rank	Our Proposed Method	Mamoun, et al., 2012	All-Kth-Markov	All-Kth-Dempster's Rule	Dempster's Rule	All-Kth-ARM	ARM	ANN
1	0.18	0.16	0.15	0.19	0.18	0.09	0.04	0.11
2	0.32	0.30	0.32	0.25	0.27	0.10	0.04	0.18
3	0.40	0.37	0.35	0.30	0.30	0.11	0.05	0.22
4	0.45	0.42	0.38	0.35	0.35	0.11	0.05	0.27
5	0.48	0.44	0.4	0.38	0.38	0.12	0.05	0.30
6	0.50	0.46	0.45	0.41	0.42	0.13	0.05	0.34
7	0.52	0.48	0.47	0.44	0.46	0.13	0.05	0.38
8	0.54	0.50	0.49	0.48	0.48	0.14	0.05	0.44
9	0.56	0.51	0.52	0.50	0.49	0.14	0.05	0.46
10	0.58	0.52	0.54	0.52	0.51	0.15	0.05	0.48

5.2 Accuracy calculation method

The method introduced in [48] is used to calculate accuracy of our Proposed method. $Rank_i$ in this Method of Mamoun A. Awad et al [49] is used to calculate accuracy of recommended method. Method which means that user's i next page is predicted instead of a page. If one of these i pages is the user's next real page, then the prediction is right.

- ✓ Hit: If prediction is correct (or if the prediction is in the prediction list in the case of applied rank), hit will increase by a unit.
- ✓ Miss: If prediction is incorrect (or not available in prediction list).

- ✓ Match: Match will have one unit of increase if it recommends at least a page to user regardless of correct or incorrect prediction of the next page.

The last page is kept and prediction is done for all remaining pages in order to identify accuracy of prediction in each session.

Accuracy of prediction is equal to ratio of correct predictions to total number of matches, is shown in equation 2:

$$Accuracy = \frac{Hit}{Matc} \tag{2}$$

Table 7 shows the accuracy values obtained from the training set (NASA data) for training neural network. The best accuracy is obtained with 75% training set and in rank = 10.

Table 7: Calculated Accuracy with training set for training neural network (NASA Dataset)

	Training (%)		
	50	60	75
Rank			
1	0.06	0.07	0.16
2	0.07	0.11	0.31
3	0.09	0.17	0.38
4	0.12	0.22	0.43
5	0.14	0.25	0.46
6	0.15	0.28	0.48
7	0.17	0.29	0.51
8	0.18	0.31	0.53
9	0.19	0.32	0.54
10	0.21	0.34	0.55

Table 8 shows the accuracy values obtained from the UOFS trained dataset for training neural network. The best accuracy is obtained with 75% training set and in rank = 10. In 60% of training set, rank = 1and rank = 2 shows an equal value of accuracy.

Table 8: Calculated Accuracy with training set for training neural network (UOFS Dataset)

	Training (%)		
	50	60	75
Rank			
1	0.04	0.06	0.12
2	0.05	0.06	0.25
3	0.06	0.16	0.28
4	0.08	0.19	0.37
5	0.09	0.23	0.4

6	0.1	0.26	0.42
7	0.11	0.28	0.46
8	0.13	0.3	0.48
9	0.14	0.33	0.5
10	0.15	0.34	0.52

Values of accuracy of comparative models for two datasets is shown in Table 9 and Table 10. As it is show in Table 9, the accuracy of our approach is better than others except for rank = 1. The best performance is achieved for rank = 10.

Table 9: Calculated Accuracy using of NASA Dataset

Rank	Our Proposed Method	Mamoun, et al., 2012	All-Kth-Markov	All-Kth-Dempster's Rule	Dempster's Rule	All-Kth-ARM	ARM	ANN
1	0.16	0.24	0.14	0.14	0.12	0.05	0.02	0.10
2	0.31	0.26	0.28	0.20	0.17	0.06	0.03	0.14
3	0.38	0.28	0.34	0.24	0.20	0.07	0.03	0.17
4	0.43	0.29	0.38	0.28	0.25	0.07	0.03	0.20
5	0.46	0.30	0.41	0.31	0.26	0.07	0.03	0.23
6	0.48	0.31	0.43	0.34	0.30	0.07	0.03	0.25
7	0.51	0.32	0.45	0.38	0.33	0.07	0.03	0.28
8	0.53	0.33	0.46	0.40	0.35	0.09	0.04	0.30
9	0.54	0.34	0.47	0.42	0.37	0.09	0.04	0.32
10	0.55	0.35	0.49	0.44	0.39	0.09	0.04	0.34

Therefore, in Table 10, our approach has the best performance in all cases except for rank 1 and 2.

Table 10: Calculated Accuracy using of UOFS Dataset

Rank	Our Proposed Method	Mamoun, et al., 2012	All-Kth-Markov	All-Kth-Dempster's Rule	Dempster's Rule	All-Kth-ARM	ARM	ANN
1	0.12	0.20	0.1	0.12	0.11	0.04	0.01	0.07
2	0.25	0.22	0.22	0.18	0.15	0.04	0.02	0.11
3	0.28	0.24	0.26	0.20	0.18	0.05	0.02	0.15
4	0.37	0.26	0.32	0.24	0.22	0.06	0.02	0.18
5	0.40	0.28	0.34	0.28	0.24	0.06	0.02	0.20
6	0.42	0.29	0.38	0.31	0.27	0.06	0.03	0.23
7	0.46	0.30	0.40	0.35	0.30	0.06	0.03	0.26
8	0.48	0.32	0.42	0.37	0.33	0.08	0.03	0.28
9	0.50	0.33	0.43	0.39	0.35	0.08	0.04	0.30
10	0.52	0.34	0.45	0.42	0.37	0.08	0.04	0.32

6 Evaluation of results

In this section, comparative results of the accuracy of the proposed model are presented by considering the different values of 50%, 60% and 75% of training set (NASA and UOFS dataset) for training the neural network. The x-axis of the figure shows the rank and the y-axis shows the accuracy of the proposed method. Note that in 75% of the training set used for training the neural network provide highest accuracy. In addition, in rank = 10, the proposed method provides the highest accuracy Figure 3, Figure 4.

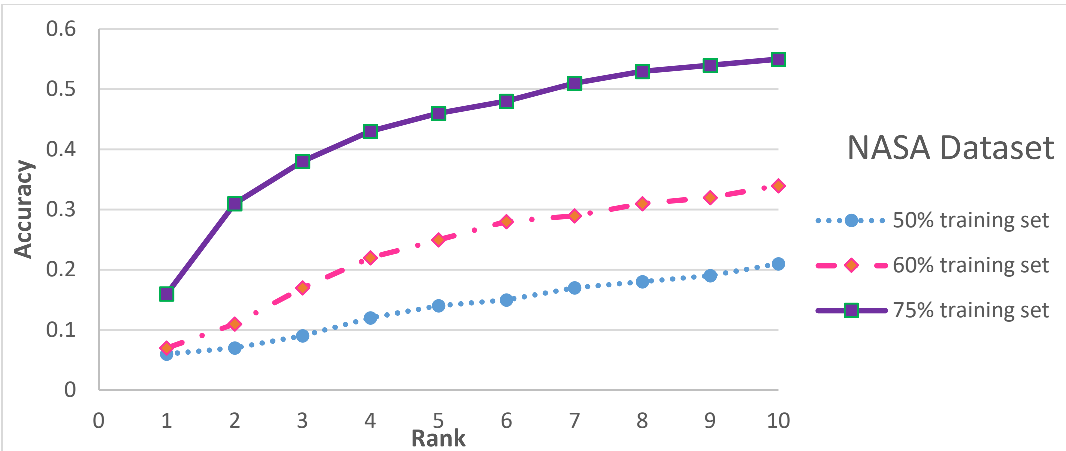

Figure 3. Improvement of Accuracy with training neural network (NASA Dataset)

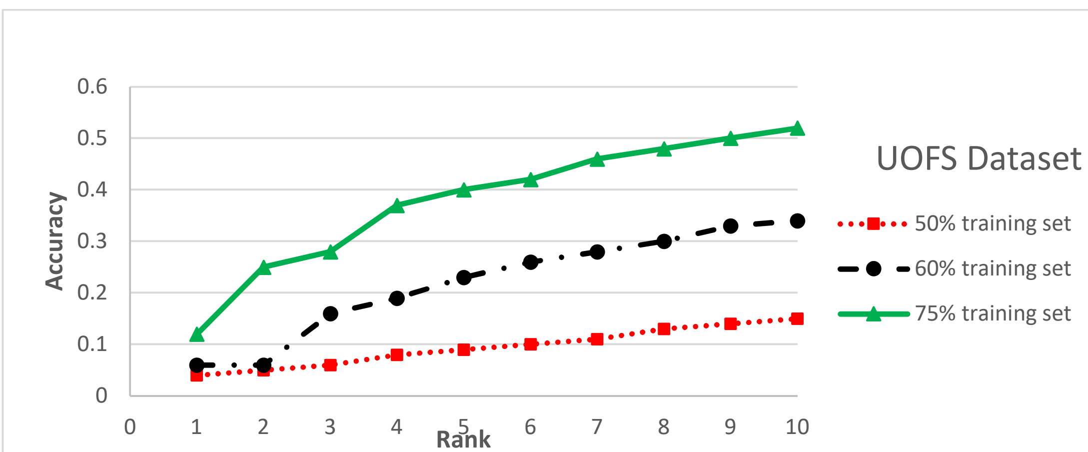

Figure 4. Improvement of Accuracy with training neural network (UOFS Dataset)

As it can be seen in Figure 3 and Figure 4, when the neural network is trained with a larger training set, the accuracy of the proposed method also increases.

In addition, we compare and present prediction results with seven different models and two datasets NASA and UOFS is used. These seven models include the proposed method in [48], ARM, ANN, model, Dempster's Rule and another three hybrid models. The hybrid models consist of All-Kth-Dempster's Rule, All-Kth-Markov and All-Kth-ARM. Furthermore, the combination of the Markov model with ANN is considered as Dempster's rule and ANN combination with All-Kth-Markov is considered as All-Kth-Dempster's rule. The concept of ranking is used in the results of this research; rank n means that performed prediction is correct if the predicted page is among the top n pages that have the highest reliability. The accuracy of our proposed method compared to the presented method in [48], and the accuracy of the proposed method in compared with the All-Kth-Order Markov model and ANN, ARM, All-Kth-ARM, Dempster's Rule, All-Kth-Dempster's Rule, the dataset of NASA is applied (Figure 5).

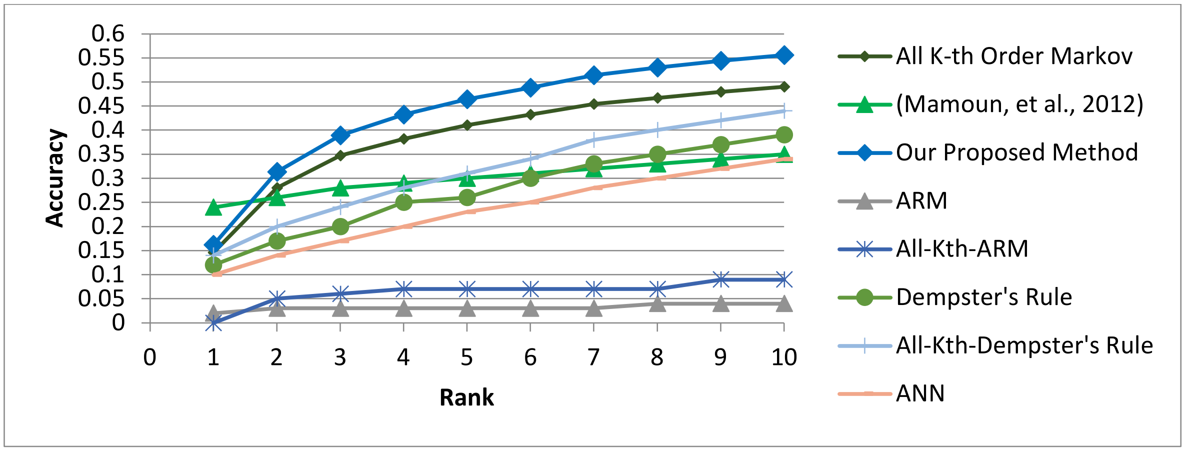

Figure 5. Comparison of accuracy of proposed method with [48] and All-Kth-Order Markov Model, ANN, ARM, All-Kth-ARM, Dempster's Rule, All-Kth-Dempster's Rule (NASA data set).

According to the obtained results of this research, it seems that a combination of classification and the Markov model can increase the accuracy of the recommended model. An increase in the accuracy of the recommended method is due to the existence of further similar sessions in a cluster that increases Hit. Furthermore, data clustering of this research is done based on three metrics: frequency, date and time of user presence on the page, so clustering users with the same interests will increase accuracy. Figure 6 shows the precision of our proposed method compared with the presented method in [48] and All-Kth-Order Markov, ANN, ARM, All-Kth-ARM, Dempster's Rule, and All-Kth-Dempster's Rule. The result shows the precision of the recommended method compared with [48], All-Kth-Order Markov model, ANN, ARM, All-Kth-ARM, Dempster's Rule, and All-Kth-Dempster's Rule. The dataset NASA is applied.

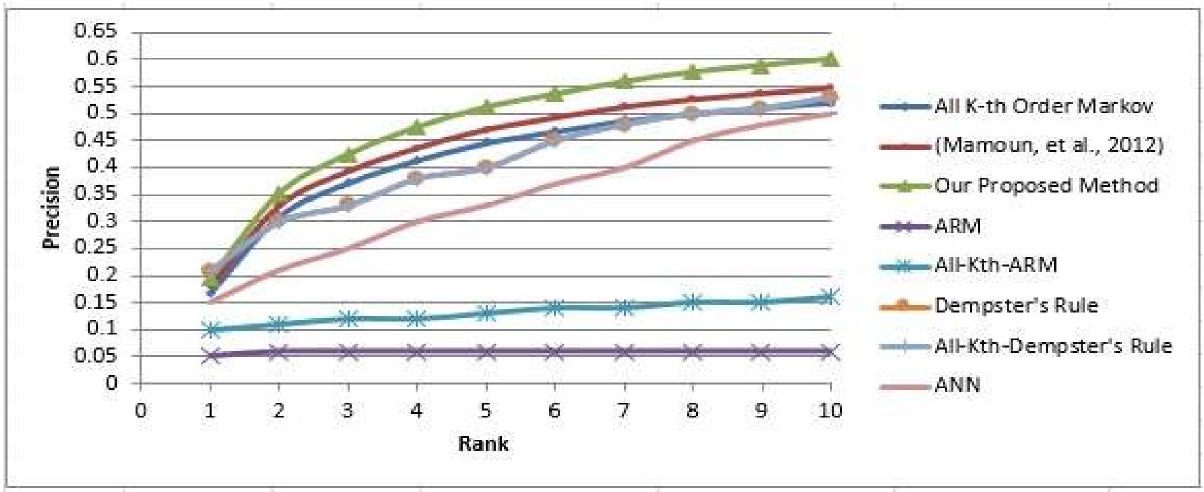

Figure 6. Comparison of precision of proposed method with [48] and All-Kth-Order Markov model, ANN, ARM, All-Kth-ARM, Dempster's Rule, All-Kth-Dempster's Rule (NASA data set).

Accuracy of our proposed method compared to the presented method in [48], and the accuracy of the proposed method in comparison with the All-Kth-Order Markov model and ANN, ARM, All-Kth-ARM, Dempster's Rule, All-Kth-Dempster's Rule, the dataset of UOFS is applied (Figure 7). Simulation results indicate that the accuracy of our proposed method is higher than the each in [48] and All-Kth-Order Markov model and ANN, ARM, All-Kth-ARM, Dempster's Rule, All-Kth-Dempster's Rule. The result is shown in Figure 5 and Figure 7. Figure 6 shows the precision of our proposed method compared with the presented method in Mamoun A. Awad et al [49] and All-Kth-Order Markov, ANN, ARM, All-Kth-ARM, Dempster's Rule, All-Kth-Dempster's Rule. The result shows the precision of the recommended method compared with [48], All-Kth-Order Markov model, ANN, ARM, All-Kth-ARM, Dempster's Rule, and All-Kth-Dempster's Rule. The dataset UOFS is applied. Simulation results indicate that precision of our proposed method is much higher than recommended method in [48] and All-Kth-Order Markov model, ANN, ARM, All-Kth-ARM, Dempster's Rule, All-Kth-Dempster's Rule. As shown in the figures, the ARM5 and All-Kth-ARM methods have the lowest accuracy and precision. Result is shown in Figure 6 and Figure 8.

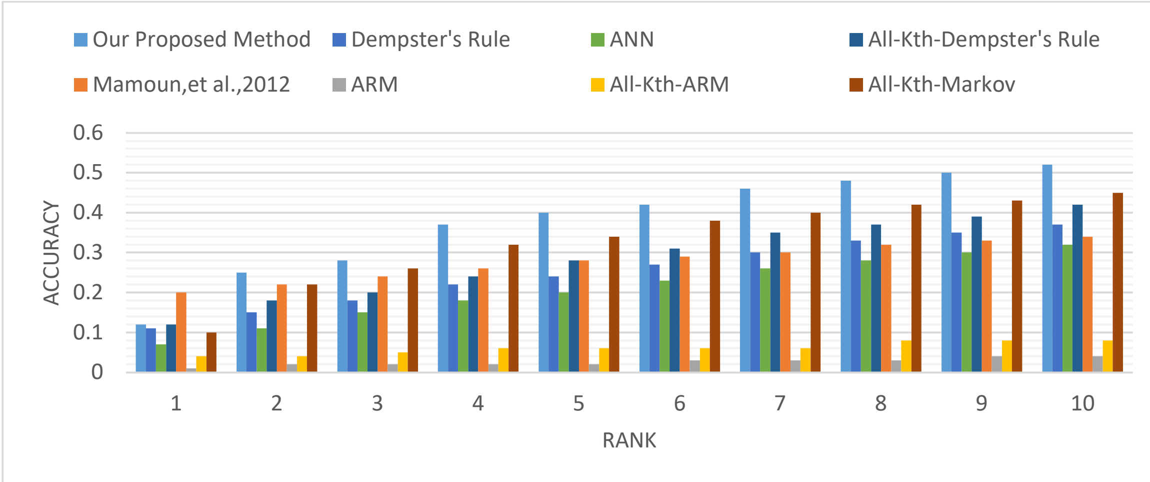

Figure 7. Comparison of accuracy of proposed method with [48] and All-Kth-Order Markov Model, ANN, ARM, All-Kth-ARM, Dempster's Rule, All-Kth-Dempster's Rule (UOFS data set).

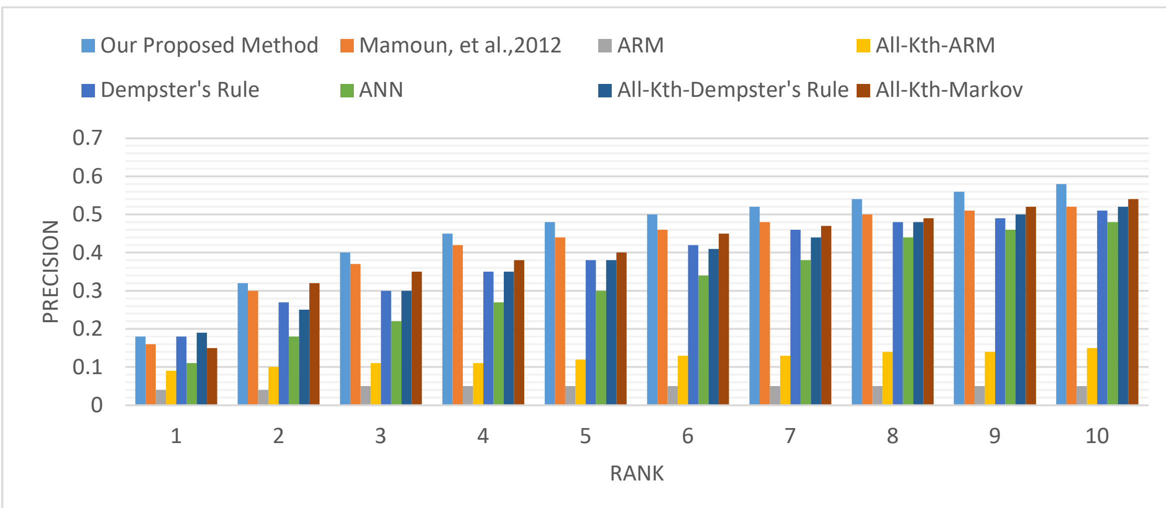

Figure 8. Comparison of precision of proposed method with [48] and All-Kth-Order Markov Model, ANN, ARM, All-Kth-ARM, Dempster's Rule, All-Kth-Dempster's Rule (UOFS data set).

❖ Recall parameter is not evaluated in this research because recommended pages are compared only with the last page.

6.1 Discussion

Accuracy and precision are two important issues in designing recommender systems. UOFS datasets are much larger than NASA databases, clustering very large databases, as UOFS is very time-consuming, which is done faster using the optimized K-mean clustering method. The results obtained in this research show that the size of the sessions is effective in the clustering process so the smaller session size reduces the accuracy and precision which is one of the disadvantages of clustering by the K-means method, but this reduction is not significant. Previous methods have evaluated a limited number of Ranks, but in this research, evaluation improved the accuracy and precision of the recommender system by considering more Ranks. Therefore, in our proposed method, by considering the greater number of pages and in higher ranks. In the experiment performed on the trained neural network using by training set (NASA and UOFS dataset), the effect of the percentage of the training set was determined. If the neural network is trained with more training sets, the accuracy and efficiency of the proposed method also increases, and the effect of rank is still significant because with increasing rank, the accuracy of the proposed method increase. In both the NASA and UOFS training sets, 75% of the training set for training neural network provides the best results. Therefore, the effect of trained a neural network for obtaining the highest accuracy is completely clear. Because the performed experiments on low percentages of training sets provide low accuracy and time-consuming, are not studied in this research. The accuracy and precision increase and efficiency of the proposed method increase compared to the previous methods. The proposed method can be expanded, because there is the ability to add several new prediction methods to it, which increases the improvement of accuracy and precision for prediction.

7 The Proposed method

In the present research, user sessions were first identified by pre-processing web server log files, and then their sessions were clustered based on three features: frequency, date and duration of the user's visit to the page. As a result, clustering users with the same interests will increase accuracy and precision. The neural network was used in the online phase to identify clusters associated with the recommended Markov model and then used by considering of threshold for predicting users' requested web pages. The evaluation was done based on comparing the recommended system with the All-Kth-Order Markov model and presented methods in [48] and ANN, ARM, All-Kth-ARM, Dempster's Rule, and All-Kth-Dempster's Rule. According to the results, the precision of recommended system is higher than all three methods, in other words, efficiency of the recommended system is higher than the All-Kth-Order Markov model and presented methods in [48], and ANN, ARM, All-Kth-ARM, Dempster's Rule, and the All-Kth-Dempster's Rule in prediction by consideration of a higher number of pages. The precision of the recommended system is also higher than the All-Kth-Order Markov model and presented methods in [48] and ANN, ARM, All-Kth-ARM, Dempster's Rule, All-Kth-Dempster's Rule. Results of tests indicated that our recommended method has higher precision in the recommendation of web pages to users. In addition, there are some suggestions for future work: (1) In this research, the neural network was used as a classifier, and other classification algorithms can be used for future work. (2) In this research, the optimized K-means clustering algorithm was used to cluster the user's session, this method can be combined with other clustering methods. (3) In this research, features of frequency, time and date of page visit are used to determine the weight of pages, for future research it can be used from other features extracted from the web log file.

References

[1] M.Riyahi, M.K.Sohrabi, Providing effective recommendations in discussion groups using a new hybrid recommender system based on implicit ratings and semantic similarity, Electronic Commerce Research and Applications, 40, 2020.

[2] S.Natarajan, S.Vairavasundaram, S.Natarajan, A.H.Gandomi, Resolving Data Sparsity and Cold Start Problem in Collaborative Filtering Recommender System Using Linked Open Data, Expert Systems with Applications, 149, 2020.

[3] B.Pooja M, Prof.Tushar J.Rava, Improving Web Page Access Prediction using Web Usage Mining and Web Content Mining, Proceedings of the Third International Conference on Electronics Communication and Aerospace Technology (ICECA). India, 2019.

[4] R.M.Mohammad, F.Thabtah, L.McCluskey, Predicting phishing websites based on self-structuring neural network, Neural Comput & Applic, Springer, London, 25, 443-458, 2014.

[5] H.Jindal, N.Sardana, R.Mehta, Efficient web navigation prediction using hybrid models based on multiple evidence combinations, International Journal of Computers and Applications, 2019.

[6] Priyanka S. Panchal, Prof. Urmi D. Agravat, Hybrid Technique for User's Web Page Access Prediction based on Markov Model, 4th ICCCNT, IEEE, Tiruchengode, India, 1-8, 2013.

[7] International Journal of Emerging Technology and Advanced Engineering, Website: www.ijetae.com, 2, 269-271, 2012.

[8] C.Romero, S.Ventura, A.Zafra, P.de Bra, Applying web usage mining for personalizing hyperlinks in Web-based adaptive educational systems, 2009.

[9] Z.Ali, P.Kefalas, kh.Muhammad, B.Ali, M.Imran, Deep learning in citation recommendation models survey, Expert Systems with Applications, 162, 113790, 2020.

[10] H.D. Park, K.H. Kim, C.Y.Il. A literature review and classification of recommender systems research". Expert Systems with Applications, 39, 10059–10072, 2012.

[11] I.Cadez, D.Hecherman P.Smyth, S.White, Model-based clustering and visualization of navigation patterns on a web site". Data Mining and Knowledge Discovery, 7, 399-424, 2003.

[12] P.Lopes, B.Roy, Dynamic Recommendation System Using Web Usage Mining for E-commerce Users, International Conference on Advanced Computing Technologies and Applications (ICACTA), 60-69, 2015.

[13] C.K.Leung, F.Jiang, Web Page Recommendation from Sparse Big Web Data, IEEE/WIC/ACM International Conference on Web Intelligence (WI), Santiago, Chile, 592-597, 2018.

[14] M.M.Ozturk, U.Cavusoglu, A novel defect prediction method for web pages k-mean++, Expert Systems with Applications, 42, 6496-6506, 2015.

[15] Sneha.Y.S, Dr.G.Mahadevan, M.Parkash.M, An Online Recommendation System Based on Web Usage Mining and Semantic Web Using LCS Algorith, IEEE, India, 223-226, 2011.

[16] L.Yongqin, X.Budong, Personalized Recommendation Method based on Web Log Mining, International Conference on Smart Grid and Electrical Automation, ICSGEA, Changsha, China, 416-419, 2018.

[17] F.Zhang, Sh.Qi, Q.Liu, M.Mao, A. Zeng, Alleviating the data sparsity problem of recommender systems by clustering nodes in bipartite networks, Expert Systems with Applications, 149, 113346, 2020.

[18] C.H.Lee, Y.H.Kim, P.K.Rhee, Web personalization expert with combining collaborative filtering and association rule mining technique, Expert Systems with Application, 21, 131-137, 2001.

[19] F.H.Wang, H.M.Shao, Effective personalized recommendation based on time framed navigation clustering and association mining, Expert Systems with Applications 27, www.elsevier.com/locate/eswa, 365-377, 2004.

[20] R.Geetharaman, P.Revathy, S. G.jacob, Prediction of users webpage access behaviour using association rule mining, Indian Academy of Sciences, Sadhana, 40, 2353-2365, 2015.

[21] J.Xiao, M.Wang, B.Jiang, J.Li, A personalized recommendation system with combination algorithm for online learning, J Ambient Intell Human Comput, 9, 667-677, 2017.

[22] Sh.Zhang, Ch.Liu, Q.Li, Website Intelligent Recommendation Based on K-means and Apriori Algorithms, International Conference on Advanced Control, Automation and Artificial Intelligence (ACAAI), Advances in Intelligent Systems Research, 155, 2018.

[23] Ch.Nigam, A.K.Sharma, Experimental performance analysis of web recommendation model in web usage mining using KNN page ranking classification approach, Materials Today: Proceedings, Elsevier, 2214-7853, 2020.

[24] P.G.Om.Prakash, A.Jaya, Analyzing and Predicting User Navigation Pattern from Weblogs using Modified Classification Algorithm, Indonesian Journal of Electrical Engineering and Computer Science, 11, 333-340, 2018.

[25] D.Wang, Y.Yih, J.Tang, M.Ventresca, et al, Improving Neighbor-Based Collaborative Filtering by Using a Hybrid Similarity Measurement, Expert Systems with Applications, 30475-30479, 2020.

[26] R.Asgarnezhad, S.S.AbdullMajeed, Z.A.Abbas, S.S.Salman, An Effective Algorithm to Improve Recommender Systems using Evolutionary Computation Algorithms and Neural Network, Wasit Journal of Computer and Mathematic Science, 1, 27-35, 2022.

[27] T.Wu, F.Sun, j.Dong, Z.Wang, Y.Li, "Context-aware session recommendation based on recurrent neural networks". Computers and Electrical Engineering, 100, 2022.

[28] J.Zhao, H.Li, L.Qu, Q.Zhang, Q.Sun, H.Huo, M.Gong, DCFGAN: An adversarial deep reinforcement learning framework with improved negative sampling for session-based recommender systems, Information Sciences, 596, 222-235, 2022.

[29] Ch.Kumar, M.Kumar, User session interaction-based recommendation system using various machine learning techniques, Multimedia Tools and Applications, 82, 21279-21309, 2023.

[30] L.Xia, et al, Multi-Behavior Graph Neural Networks for Recommender System, IEEE Transactions on Neural Networks and Learning Systems, 1-15, 2022.

[31] X.Gu, H.Yang, J.Tang, J.Zhang, et al, Profiling Web users using big data, Social Network Analysis and Mining, 8-24, 2018.

[32] M.H.Aghdam, Context-aware recommender system using hierarchical hidden markov model, Physica A: Statistical Mechanics and its Application, 518, 89-98, 2018.

[33] K.Drushku, J.Aligon, N.Labroche, P.Marcel, Interest-based recommendations for business intelligence users". Information Systems, 89, 79-93, 2018.

[34] F.Eskandanian, B.Mobasher, Modeling the Dynamics of User Preference for Sequence-Aware Recommendation Using Hidden Markov Models, The 32nd International FLAIRS Conference in Association with AAAI. Florida, 2019.

[35] G.Karami, J.Tian, Maintaining accurate web usage models using updates from activity diagrams". Information and Software Technology, 96, 68-77, 2017.

[36] M.Narvekar, Sh.Banu, Predicting User's Web Navigation Behavior Using Hybrid Approach". International Conference on Advanced Computing Technologies and Applications (ICACTA), Elsevier, 3-12, 2015.

[37] R.Abolghasemi, A personality-aware group recommendation system based on pairwise preferences, Information Sciences, 595, 1-17, 2022.

[38] A.Pujahari, D.S.Sisodia, Modeling user's preference changes in recommender systems via time-dependent Markov random fields, 243, 2023.

[39] S.Ni, Sh.Hu, L.Li, An Intention-Aware Markov Chain Based Method for Top-K Recommendation". IEEE Transactions on Automation Science and Engineering, 1-12, 2022.

[40] S.Cheriyan, Dr.K.Chitra, Web page prediction using Markov Model and Bayesian Statistics, Second International Conference on Electrical, Computer and Communication Technologies (ICECCT), IEEE. India.

[41] H.Chen, Z.Lin, A Hybrid Neural Network and Hidden Markov Model for Time-aware Recommender Systems. 11th International Conference on Agents and Artificial Intelligence, ICAART, Prague, 2, 204-213, 2019.

[42] Ch.Djellali, M.adda, A New Hybrid Deep Learning Model based-Recommender System using Artificial Neural Network and Hidden Markov Model, The 17th International Conference on Mobile Systems and Pervasive Computing (MobiSPC) , 175, 214-220, 2020.

[43] X.Nan, K.kanato, X.Wang, Design and Implementation of a Personalized Tourism Recommendation System Based on the Data Mining and Collaborative Filtering Algorithm, Hindawi Computational Intelligence and Neuroscience, 1-14, 2022.

[44] G.Huang, E-Commerce Intelligent Recommendation System Based on Deep Learning, IEEE Asia-Pacific Conference on Image Processing, Electronics and Computers (IPEC), 2022.

[45] Sh.Gheewala, Sh.Xu, S.Yeom, S.Maqsood, Exploiting deep transformer models in textual review based recommender systems, Expert Systems With Applications, 235, 2023.

[46] S.Safavi, M.Jalali, M.Houshmand, Toward Point-of-Interest Recommendation Systems: A Critical Review on Deep-Learning Approaches, MDPI, electronics, 11, 2022.

[47] A web server log file sample. available on 25 April 2014[online]. *www.jafsoft.com/*searchengines*/log_sample.html,* 2014.

[48] Mamoun A. Awad, L.R.Khan, Web Navigation Prediction Using Multiple Evidence Combination and Domain Knowledge, IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS–PART A: SYSTEMS AND HUMANS, 37, 1054-1062, 2007.

[49] Mamoun A. Awad, Issa Khalil. "Prediction of User's Web Browsing Behavior: Application of Markov Model". IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS–PART B: CYBERNETICS, 42, 1131-1142, 2012.