2019 International Conference on Deep Learning and Machine Learning in Emerging Applications (Deep-ML)

A Collaborative Recommendation System for Online Courses Recommendations

Raghad Obeidat
Department of Computer Science
Jordan University of Science and Technology
Irbid, Jordan
raobeidat16@cit.just.edu.jo

Rehab Duwairi
Department of Computer Information Systems
Jordan University of Science and Technology
Irbid, Jordan
rehab@just.edu.jo

Ahmad Al-Aiad
Department of Computer Information Systems
Jordan University of Science and Technology
Irbid, Jordan
aiaiad@just.edu.jo

***Abstract*—In this paper, we present a collaborative recommender system that recommends online courses for students based on similarities of students' course history. The proposed system employs data mining techniques to discover patterns between courses. Consequently, we have noticed that clustering students into similar groups based on their respective course selections play a vital role in generating association rules of high quality when compared with the association rules generated using the whole set of courses and students. In particular, the Apriori algorithm was used to generate association rules; once using the whole dataset and once using the clusters which are formed based on students' choices of courses. The results reveal that the coverage of the rules generated on clusters are better. Also, to assess the effect of course dependency on recommendations, we applied the SPADE algorithm on course sequences. The results are in harmony with the results obtained when Apriori was applied.**

Keywords— Course recommendation system, Collaborative Filtering, Data mining, Association Rule mining

I. INTRODUCTION

Most of the decisions people make are based on suggestions or recommendations coming from people experiences or from the Internet [1]. In E-commerce domain customers interested in reading product reviews or product ratings before buying. Recommendation systems (RS) make the process easier to assist people regarding their preferences [1]. It can be defined as a collection of techniques that increase the benefit of complex information and facilitate interaction with it [2, 3]. Identifying people's preferences and making decisions becomes a big challenge to build recommender systems that cope with information overload. Extracting users' data can be done implicitly by identifying users' actions or explicitly by collecting users' ratings [4].

Collaborative filtering [1, 5, 6] is one of RS techniques which is based on gathering and analyzing large data sets including users' activities, opinions or preferences. It is "user to user" correlation method, thus predicting items to the target user by finding similarities with other users and their preferences. Content-based filtering rely mainly on content knowledge source and description of the items [6]. It recommends items based on a comparison between the item's content and target user preferences. Sometimes Content-based filtering is used to support collaborative filtering, which called hybrid-based filtering [6] to get better results and overcome limitations.

The growth of educational resources represents a great revolution [7], given the increase in Massive Online Open Courses (MOOCs) [7, 8] such as Coursera, Udacity, EdX, etc., learners can find courses from almost every knowledge domain. Currently, educational systems represent an emerging research area, these systems use different recommendation techniques in order to suggest online learning activities [6], based on preferences, knowledge, and data from other students with similar interests.

The course recommendation in E-learning systems suggests to the students the best combination of subjects in which they are interested in [9]. Suggestions are developed using a set of rules and methods to propose a list of courses using the student's data regarding their career interest and goals.

This paper presents a recommendation methodology that recommends courses to students based on affinity between courses taken by the target student and other students. It aims to provide an effective course recommendation using multiple techniques. The students will be clustered into groups based on their courses' grades, and traditional data-mining (DM) techniques will be applied for generating association rules with a predicted course's grade.

This paper consists of five sections. Section 1 provided the introduction to this work. Section 2 describes related work. Section 3 explains the methodology and framework of this research. Section 4 describes and experiments and results. Finally, Section 5 draws the conclusions of this work.

II. RELATED WORK

This Section describes briefly some recommendation systems that employ association rule mining and apply different techniques to discover patterns between users. Tewari et al. in [10] used association rule mining to develop a book recommendation system based on the content and the quality of the book by combining features of content-based filtering and collaborative-based filtering. On the other hand, Parvatikar and Joshi [11] have used various techniques for recommending books to purchase. They used item-based collaborative filtering and apply association rule; the ensuing results show that their recommendation system solved data sparsity and scalability.

Additionally, a lot of research has been conducted on building a recommendation system that integrates both user clustering and association rules techniques. Al-Badarenah and Alsakran in [6] proposed a collaborative recommender system for course

978-1-7281-2914-3/19/$31.00 ©2019 IEEE
DOI 10.1109/Deep-ML.2019.00018

selection that recommends elective courses based on similarities between students, they used the *K-means* algorithm to cluster similar student and the *K-nearest neighborhood* technique to select the most similar cluster to the target student then apply association rule mining for each cluster. Wang and Shao in [12] developed a personalized recommendation system of web pages, their proposed model includes clustering time-framed navigation sessions using Hierarchical Bisecting Medoids algorithm and analyzing those navigation sessions with association-mining for recommendations in the future. As a result, their system improves recommendation service effectively.

From a user-fragmentation aspect, there is a new recommendation system created in [13] that uses online marketing dataset to cluster the customers, suppliers, and products into groups using Self-Organizing maps neural networks (SOM) and extract meaningful pattern using association rule mining. Wen-Shung Tai et al. in [14] proposed a hybrid recommendation system that combines SOM and data mining techniques. They used courses dataset which consist of 850 learners, SOM neural networks have been used to cluster learners based on similarity and adopted the Apriori algorithm to generate association rules for each cluster.

In [15], Aher and Lobo used a combination of machine learning algorithms (clustering technique using Simple K-means algorithm and association rule using Apriori algorithm) for course recommendation in E-Learning System based on historical data. Upendrana et al. in [16] proposed a course recommendation system based on student's cognitive ability and academic achievements, their system designed to help students to short-list courses that suit their grades. It recommends a list of courses that have a high probability of success.

III. THE PROPOSED RECOMMENDATION SYSTEM

Recently, the number of people learning online courses is rapidly increasing. The volume, availability, and diversity of information made it easier for students to find their own unique tendencies and goals [6-8]. One of the students' concerns is to make decisions about what course they intend to take and to achieve an acceptable grade in it. In fact, there is a lot of similarities between students' experiences and interests, students read feedback [7] or reviews for those who have taken the same course, which means that it can affect their decisions.

To give better suggestions and facilitate decision making, we have developed a course recommendation system that recommends courses based on similarities between students and meets the specified minimum acceptable grade for the recommended course. Our proposed method tries to recommend courses based on students' similarities. First, clustering is used to classify or group students with the same interest. When groups have been established, data mining techniques will be used to elicit the rules of the best learning path.

Throughout our experiments we explore and try to answer the following questions:

- What is the effect of the clustering method for generating better association rules covering all types of students?
- Which performs better in recommendation; clustering students into groups before generating association rules or generating association rules without clustering students?
- What is the effect of the sequential pattern mining on the recommendations?
- Which performs better in recommendations; association rule mining or sequential pattern mining which based on sequences and events?

Building a course recommendation system required many steps; Fig. 1 shows our recommendation methodology. In the first phase, we will use a dataset that contains the usage data for the educational system; it used to analyze students' information and identifies the substantial parameters that affect recommendation process. Next, the students will be grouped into a set of clusters based on their previous behavior and history. After that a rule generation technique will be applied for each cluster, the rules are used to recommend courses for the student based on the similarity within the groups. Finally, we will evaluate the system.

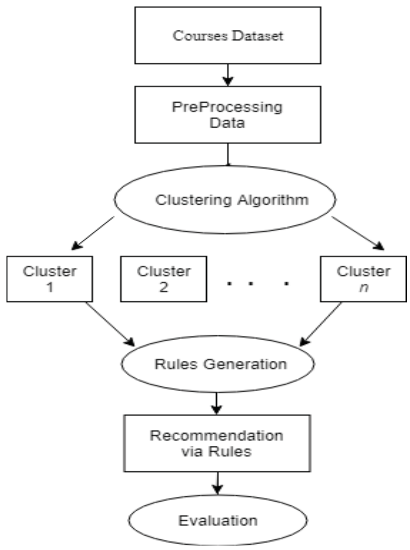

Fig. 1. Recommendation Methodology

A. Dataset Description

We performed our experiments using Open Online Courses dataset, where there were 841,687 registrations taken by 597,692 unique students of HarvardX and MITx in the year from the fall of 2012 to the summer of 2013 [17]. The total number of courses is 17 courses launched on the edX platform.

Our experiment conducted on 22,144 records taken by 10,000 unique users with 16 courses as maximum courses. Data preparation and preprocessing are the basic operations that include removing incomplete records and calculating the order of courses sequences events for each student. Grades is a number between 0 and 1, we have converted the grades into new grade scales. Table 1 shows courses grade scales.

Table I Course Grade Scale

Grade Range	Grade Scale
[0.7-1.0]	A
(0.7-0.5]	B
(0.5-0.1]	C
(0.1-0.0]	F
Not taken course	N

B. Clustering Phase

Clustering algorithm will be applied on the courses dataset, this step is necessary to group similar students to the same cluster based on their previous behavior and history. We have used Weka 3.8 data science tool to cluster similar students. The used clustering algorithm is k-means clustering algorithm, and Euclidean distance is used to measure the similarity between courses.

C. Course mining

This step is the core step in our recommendation experiments. Data mining techniques are used to obtain association rules, courses dataset is built by mapping each course to an item and each student to a transaction. Table 2 shows an example of courses dataset, where "N" grade scale denotes that the student does not take the course. Rules generation have been conducted on R programming language and free software environment, two algorithms have been used in this phase: Apriori Algorithm Association Rule Mining and SPADE Algorithm [19] which is known as vertical format sequential pattern mining method.

Table II An Example of Courses Dataset

Student ID	PH207X	CS50X	6.00X	7.00X
1	A	B	C	N
2	B	C	C	F
3	N	F	F	N
4	A	B	N	B

- *Association Rule mining*

Apriori algorithm is applied to each cluster created from the previous step, it takes courses transactional datasets which contains transaction Id that corresponds to each student's set of items which is form of course-grade pairs. Table 3 shows an example of courses transactional dataset. An example of courses association rules that can be generated: ***{CS50X-A** Λ **PH207X-C} =>{6.00X-B}***, this indicates if a target student got A in CS50X and C in PH207X courses thus, 6.00X course will be recommended with expected grade B.

Table III An Example of Courses Tansactional Dataset Used in Association Rule Mining

Student ID	Items
1	PH207X-A, CS50X-B, 6.00X-C
2	CS50X-C, 6.00X-C, PH207X, 7.00X-F
3	CS50X-F, 6.00X-F
4	7.00X-B, CS50X-B, PH207X-A

- *Sequential pattern mining*

Frequent sequence mining is used to discover patterns that have a specific order [19], hence the arrangement of taken courses is needed to be applied on SPADE algorithm. We have sorted taken courses for each student based on course start date and end date parameters in the original dataset. An example of courses association rules that can be generated: **< *{CS50X-A* Λ *6.00X-A}* > => < *{PH207X -B}* >**, this indicates if a target student took CS50X course with grade A before taking 6.00X course with grade A thus, PH207X course will be recommended with expected grade B.

D. Courses Recommendation

Association rules generated in the prior step will be used for courses recommendation. The association rules are in the form of $\{c_1\text{-}g \wedge c_2\text{-}g \ldots \wedge c_n\text{-}g\} \Rightarrow \{c_{n+1}\text{-}g \wedge \ldots \wedge c_k\text{-}g\}$, where c represents course name and g represents the course grade.

When evaluating the quality of association rules that will be selected later in our recommendation strategy, three common factors to be considered are the coverage, the support and the confidence factor of the rule. Generality / Coverage [18]: A pattern is general if it covers a relatively large subset of a dataset. If a pattern characterizes more information in the dataset, it tends to be more interesting (i.e. the number of tuples satisfied by the antecedent part of the rule). The recommendation strategy is described as follows:

1) If the target student has taken the course in the antecedent part of the rule course, then the system recommends the student to take courses in the consequent part of the rule.
2) If more than one rule recommends target student to different courses, then the student can select courses recommended by rues that have high quality (high coverage high support and high confidence).
3) In the case of not matching some courses in the antecedent part of the rule with courses that taken by target student, the rule is used to be recommended if the number of matching courses is greater than match threshold specified by the student.
4) The target student specifies the minimum acceptable grade, so the system recommends rules that have courses in consequent part of the rule if the expected grades are greater than the specified minimum grade.
5) If all the recommended courses in the consequent part of all the rules have less than minimum acceptable grade. The system will retrieve all students records with their course sequence, it will search for students who have taken the same course in the consequent part of the rule and return their course sequence. Course sequences are selected if the related course has a grade greater than the specified grade, thus, it will recommend the prerequisite courses from course sequences to achieve high grades based on best students' history.

In the following example, the system recommends courses for a student who has taken five courses, table 4 shows an example of courses taken by targt student

Table IV An Example of Courses Taken by Target Student

CS50X	ER22X	PH278X	6.002X	14.73X
A	C	F	F	B

1) {CS50X-A ∧ PH278X-F} =>{6.00X-B}
2) {CS50X-A ∧ ER22X-C ∧ 14.73X-B} => {CB22X-B ∧ PH207X-C}

If minimum accepted grade was **B**, then recommendation using association rules are:

1) Recommends 6.00X-B
2) Recommends CB22X-B and don't recommend PH207X-C because C < B.

If there was only one association rule that recommends course has less than accepted grade *{CS50X-A} => {6.00x-Spring-F}*. System will search for 6.00x-Spring-(A||B||C||D) Accepted mark in all students' course sequences, If Accepted mark was **A** then search for 6.00x-Spring-A.

The recommendation will recommend the prerequisite courses from found course sequences to achieve high grades based on best students' history. For example,

1) PH207X-A/14.73x-A/**6.00x-Spring-A.** Recommends PH207X-A and14.73x-A to achieve **6.00x-Spring-A**.
2) CS50X-A/**6.00x-Spring-A.** Recommends CS50X to achieve **6.00x-Spring-A**.
3) ER22X-F/8.02x-F/**6.00x-Spring-A**/7.00x-B/PH278X-C. Don't recommends pre-requisite courses because it not logical to recommends courses with fail expected grades.

IV. RESULTS

As early mentioned, the k-means clustering algorithm is used, with k = 5 as the best number of clusters. To evaluate our recommendation system, we did a comparison between the Apriori algorithm association rule mining and SPADE algorithm sequential pattern mining. Coverage measure has been used to study the performance of recommendation when clustering students into groups before generating association rules and generating association rules without clustering students.

The system takes two parameters to generate association rules: minimum support and minimum confidence. Our goal is to achieve high coverage for getting high-quality association rules that cover a large number of students rather than generating a large number of association rules not covering most students. Coverage denotes the number of tuples that satisfy the left part of the rule which means the number of students that took courses in the antecedent part of the rule. Table 5 shows the results of rules coverage using Apriori algorithm association rule mining, it shows the average coverage of all generated rules on both datasets; dataset without clustering and dataset with clustering by taking the average coverage for all clusters.

Table V Rules Coverage using Apriori Agorithm Association Rule Mining

Support\Confidence	**Rules Coverage using Apriori Algorithm Association Rule Mining**		
	Coverage	***Without clustering***	***With clustering***
Supt: 0.01, Conf: 0.01	Cov>=0.1	0.268	0.374
	Cov >=0.2	0.481	0.602
Supt: 0.01, Conf: 0.1	Cov>=0.1	0.199	0.29
	Cov >=0.2	0.407	0.494

As shown in Table 5, association rules using the clustering step have higher coverage values. We think finding similar students and predict rules to cover the whole dataset would be difficult, unlike the clustering step which classifies similar students into groups to generate rules for each group. Every rule has support, confidence, and coverage. When increasing the three parameters above 0.1 on both datasets, we noticed that the number of rules decreased, and in some cases, we did not get rules. Association rules may have high confidence and have low coverage on the dataset hence, one low coverage rule will affect the average coverage of the rules. specifying minimum coverage would perform better in the recommendation system, as shown in table 5 the increased coverage yield better results on both datasets, but clustering step has a significant impact on the performance.

We performed the second experiment to examine whether the arrangement of the courses had an impact on generated rules. Table 6 shows rules coverage using SPADE algorithm sequential pattern mining, SPADE algorithm acts better on both datasets. The values of the none-clustered dataset have improved, and values on clustered dataset were very close to the Apriori algorithm values and have higher values comparing with none-clustered datasets. From Fig. 2 we can see the significant impact of clustering step on the total coverage of association rule using Apriori and SPADE algorithms, we can notice the performance of non-clustered dataset has improved and the rules absence in SPADE algorithm when the coverage and confidence parameters were (20 %, 10%) respectively.

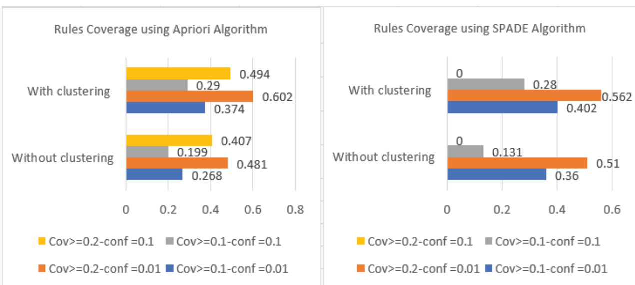

Fig. 2. Association Rules Coverage using Apriori Algorithm and SPADE Algorithm on 10,000 students' dataset

Table VI Rules Coverage using SPADE Algorithm Sequential Pattern Mining

Support\Confidence	**Rules Coverage using SPADE Algorithm Sequential Pattern Mining**		
	Coverage	***Without clustering***	***With clustering***
Supt: 0.01, Conf: 0.01	Cov>=0.1	0.360	0.402
	Cov >=0.2	0.51	0.562
Supt: 0.01, Conf: 0.1	Cov>=0.1	0.131	0.28
	Cov >=0.2	No rules	No rules

Most students take online courses to learn something new or to get information that helps them in certain areas. In fact, students do not take getting good grades in courses' assignments seriously. Students are not bound to continue learning courses or to be committed to doing assignments and most of them do not have a specific plan to get the certificates. Most of the students in our datasets have grades between 0 and 0.1, the generated rules contain courses with an expected grade "F". We filtered the data and excluded zeros grades, to take students records to those who have courses' grades between [0.01-1] and it has the same scale A-F as shown in table 1. The new experiment was performed on 11,915 records taken by 5,391 unique users with 10 course as maximum courses, and selected students with more than two courses.

Table 7 and Table 8 show Rules coverage on the second filtered data using Apriori and SPADE algorithms respectively. The generated association rules were better than the old dataset rules, and the coverage of the none-clustered dataset was fixed values for both algorithms, it achieved coverage of 0.122 as shown in Fig. 3. Coverage values on clustered dataset have significant values comparing with none-clustered datasets for both algorithms.

Table VII Rules Coverage using Apriori Algorithm Association Rule Mining

Support\Confidence	**Rules Coverage using Apriori Algorithm Association Rule Mining**		
	Coverage	***Without clustering***	***With clustering***
Supt: 0.01, Conf: 0.01	Cov>=0.1	0.122	0.534
	Cov >=0.2	0.122	0.582
Supt: 0.01, Conf: 0.1	Cov>=0.1	0.122	0.348
	Cov >=0.2	0.122	0.46

Table VIII Rules Coverage using SPADE Algorithm Sequential Pattern Mining

Support\Confidence	**Rules Coverage using SPADE Algorithm Sequential Pattern Mining**		
	Coverage	***Without clustering***	***With clustering***
Supt: 0.01, Conf: 0.01	Cov>=0.1	0.122	0.546
	Cov >=0.2	0.122	0.594
Supt: 0.01, Conf: 0.1	Cov>=0.1	0.122	0.28
	Cov >=0.2	0.122	0.376

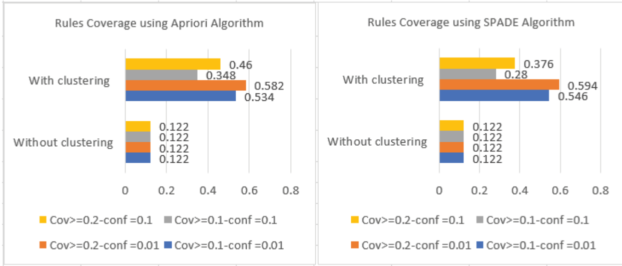

Fig. 3. Association Rules Coverage using Apriori Algorithm and SPADE Algorithm on 5,391 students' dataset

V. CONCLUSION

In this paper, we have proposed a course recommendation system that recommends online courses for students based on similarity and dissimilarity between target student and other students. Data mining techniques have been used by applying association rules algorithms to generate courses rules, coverage measure was used to study the performance of the recommendation. Through our experiments, we noticed that clustering dataset into similar clusters would have higher coverage values, unlike generating rules to cover the whole dataset. Clustering dataset has a significant impact on performance and choosing high coverage yields better results in the recommendation system. In future work, much work can be performed by applying our method in other domains of interest such as E-commerce domain by recommending products to customers and doing a comparison between our method and other typical methods.

REFERENCES

[1] M. Kunaver and T. Požrl, "Diversity in recommender systems – A survey", *Knowledge-Based Systems*, vol. 123, pp. 154-162, 2017. Available: 10.1016/j.knosys.2017.02.009.

[2] G. Adomavicius and A. Tuzhilin, "Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions", *IEEE Transactions on Knowledge and Data Engineering*, vol. 17, no. 6, pp. 734-749, 2005. Available: 10.1109/tkde.2005.99.

[3] J. Bobadilla, F. Ortega, A. Hernando and A. Gutiérrez, "Recommender systems survey", *Knowledge-Based Systems*, vol. 46, pp. 109-132, 2013. Available: 10.1016/j.knosys.2013.03.012.

[4] R. Farzan and P. Brusilovsky, "Social navigation support in a course recommendation system", in *International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems*, Berlin, 2006, pp. 91-100.

[5] H. Thanh-Nhan, H. Nguyen and N. Thai-Nghe, "Methods for building course recommendation systems", in *2016 Eighth International Conference on Knowledge and Systems Engineering (KSE)*, 2016, pp. 163-168.

[6] A. Al-Badarenah and J. Alsakran, "An Automated Recommender System for Course Selection", *International Journal of Advanced Computer Science and Applications*, vol. 7, no. 3, 2016. Available: 10.14569/ijacsa.2016.070323.

[7] C. Romero and S. Ventura, "Educational data mining: A survey from 1995 to 2005", *Expert Systems with Applications*, vol. 33, no. 1, pp. 135-146, 2007. Available: 10.1016/j.eswa.2006.04.005.

[8] D. Yang, M. Piergallini, I. Howley and C. Rose, "Forum thread recommendation for massive open online courses", in *Educational Data Mining 2014.*, 2014.

[9] D. Upendran, S. Chatterjee, S. Sindhumol and K. Bijlani, "Application of Predictive Analytics in Intelligent Course Recommendation", *Procedia Computer Science*, vol. 93, pp. 917-923, 2016. Available: 10.1016/j.procs.2016.07.267.

[10] A. Tewari, A. Kumar and A. Barman, "Book recommendation system based on combine features of content based filtering, collaborative filtering and association rule mining", in *2014 IEEE International Advance Computing Conference (IACC)*, 2014, pp. 500-503.

[11] S. Parvatikar and B. Joshi, "Online book recommendation system by using collaborative filtering and association mining", in *2015 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC)*, 2015, pp. 1-4.
[12] F. Wang and H. Shao, "Effective personalized recommendation based on time-framed navigation clustering and association mining", *Expert Systems with Applications*, vol. 27, no. 3, pp. 365-377, 2004. Available: 10.1016/j.eswa.2004.05.005.
[13] S. Changchien and T. Lu, "Mining association rules procedure to support on-line recommendation by customers and products fragmentation", *Expert Systems with Applications*, vol. 20, no. 4, pp. 325-335, 2001. Available: 10.1016/s0957-4174(01)00017-3.
[14] D. Wen-Shung Tai, H. Wu and P. Li, "Effective e-learning recommendation system based on self-organizing maps and association mining", *The Electronic Library*, vol. 26, no. 3, pp. 329-344, 2008. Available: 10.1108/02640470810879482.
[15] S. Aher and L. Lobo, "Combination of machine learning algorithms for recommendation of courses in E-Learning System based on historical data", *Knowledge-Based Systems*, vol. 51, pp. 1-14, 2013. Available: 10.1016/j.knosys.2013.04.015.
[16] D. Upendran, S. Chatterjee, S. Sindhumol and K. Bijlani, "Application of Predictive Analytics in Intelligent Course Recommendation", *Procedia Computer Science*, vol. 93, pp. 917-923, 2016. Available: 10.1016/j.procs.2016.07.267.
[17]"HarvardX-MITx Person-Course Academic Year 2013 De-Identified dataset, version 2.0 - MITx and HarvardX Dataverse", *Dx.doi.org*. [Online]. Available: http://dx.doi.org/10.7910/DVN/26147. [Accessed: 08- Mar- 2019].
[18] L. Geng and H. Hamilton, "Interestingness measures for data mining", *ACM Computing Surveys*, vol. 38, no. 3, p. 9-es, 2006. Available: 10.1145/1132960.1132963.
[19] M. J. Zaki. SPADE: An efficient algorithm for mining frequent sequences. Machine Learning, 42(1/2):31–60, 2001.