

PAPER • OPEN ACCESS

E-Learning in Data Analytics on Basis of Rule Mining Prediction in DM Environment

To cite this article: T. Nagarathinam *et al* 2021 *J. Phys.: Conf. Ser.* **1963** 012166

View the [article online](#) for updates and enhancements.

You may also like

- [Systematics in the ALMA Proposal Review Rankings](#)
John Carpenter
- [Psychological and experiential factors affecting climate change perception: learnings from a transnational empirical study and implications for framing climate-related flood events](#)
Gala Munoz-Carrier, Dana Thomsen and Gary J Pickering
- [Peer review declaration](#)

PRIME
PACIFIC RIM MEETING
ON ELECTROCHEMICAL
AND SOLID STATE SCIENCE

HONOLULU, HI
Oct 6–11, 2024

Abstract submission deadline:
April 12, 2024

Learn more and submit!

Joint Meeting of
The Electrochemical Society
•
The Electrochemical Society of Japan
•
Korea Electrochemical Society

E-Learning in Data Analytics on Basis of Rule Mining Prediction in DM Environment

T. Nagarathinam¹, V.R. Elangovan², Ahmed J. Obaid³, D. Akila⁴, Dang Quang Tuyen⁵

¹Department of Computer Science, Mass college of Arts & Science, Kumbakonam, India

²Department of computer Applications Agurchand Manmull Jain college Meenambakkam Chennai, India

³Faculty of Computer Science and Mathematics, University of Kufa, Iraq

⁴Department of Information Technology, Vels Institute of Science, Technology and Advanced Studies, Chennai, India

⁵Vietnam National University, Hanoi (VNU), Vietnam

atnaga123@gmail.com, elangotesting@live.com, ahmedj.aljanaby@uokufa.edu.iq, akiindia@yahoo.com, quangtuyen88@gmail.com

Abstract Nowadays, most of the information science inquires about a spotlight on affiliation rule to decide explicit examples and rules from huge information. Affiliation guideline is worked by basic information duration device, for example, WEKA which incorporates arrangement, bunching, affiliation runs, etc. When all is said in done, the affiliation standard could be pertinent for enormous E-Learning datasets like they were pertinent towards the formless arrangement. Within this exam, the uses for affiliating rule after gathering under straight information for a considerable length of time from LMS have been explored. Intriguing quality measurements and other applicable understudy inclinations are caught in the wake of thinking about a few affiliation rule calculations. Some imagined introductions of rules and their applicable outcomes will be exhibited in wording reasonableness in eLearning conditions.

Keywords: Big data, Integrated, Governance, Backup, Restoration, Maintain, data management system, Bayesian Frequency Correction (BFC)

1 Introduction

The Volume of data store up in the data warehouse are increasing in day to day and data to generate the occurrence of Big Data. Big data likely [2] can define through three main concepts, such as Variety, Velocity and Volume (VVV). Big data knowledge offers novel approaches as well as prospects guiding principles to defeat challenges via analysis methods in the duration of Big Data. Data duration in the education field is rising like a novel way toward discovered along with developing new educational situations, particularly in E-Learning, wherever the information could compose digitally recognize the performance of the learners [1].

E-Learning system gathers a huge quantity of data more than a period as of student behavior concerning surroundings studies, online everyday jobs, a variety of course recitals that is precious to assessing the efficiency of courses accessible along with



students on the whole learning ability. The evaluated data can assist the development of E-Learning taken as a whole and augment its efficiency and give a capable online education atmosphere. Investigating the E-Learning data can be fairly hard owing to its complication and discovery of the association flanked by diverse information. So, apply the exact method such as ARM is capable of disclosing some important models. Every one of belongings measured, the significance of Big Data analysis is for understanding the secreted models or patterns, co-incident among attributes of data.

During spitefulness of reality to the irregular itemsets able to measured like uninteresting or outliers, on times they are precious, particularly those to bear a resemblance to harmful associations, consider the instance within a meadow of education, safety as well as healthcare[4]. For the work, ARM is applied to E-Learning data like Predictive Apriori & Apriori[5] algorithms being employed in the direction of assessing their recital, specifically accuracy, rule generation & time consumption. Compare the Interestingness of the number of experiments by different configurations. The consequences from every ARM produce will be scrutinized to simplify the enrolled student behavior in E-Learning courses intended for future improvement.

2 Association Rule

ARM or incessant example mining has been primarily projected by [3] has been a standout amongst the broadly utilized systems in information duration. A case of an affiliation guideline would be "If a client purchases bread, the person within the question is probably going to purchase jam additionally" and the level of this likelihood being genuine is rely upon the certainty level and supports. The technique utilized contingent proclamations (assuming/at that point) to uncover the association between data in the database [5]. The point of the affiliation principle is to discover common examples happening into a dataset more than some time stretch. The guidelines created through this system able to assessed and direct to increment in deals via particular promoting, deals battle, stock administration, better client relationship and item arrangement strategy.

Everything is by all accounts a simple errand when the information is little and can be prepared on the double. Yet, in enormous information affiliation guideline mining, some things should be tended to, for example, information legitimacy, various information measurements, mining calculation and the correct apparatus. The adequacy of affiliation principles is well-characterized by the base help, certainty and standards made. The means are the continuous item sets being found by calculations; then, affiliation principles are created from those item sets. It is not vital that all affiliation standards are helpful, just those that legitimized the base help edge and least certainty limit set by the user [8].

3 Literature Survey

B.F Skinner [12] hypothesizes that learning takes place during observation, reiteration and alteration to an agreed behavior. A learner's mentality is perceived as a clean line up on knowledgeable about wiping out during plunder and punishments. In a behaviorist, restricted learning environment, the education list presumes the position of a single provider and mediator of come again is learned. As a consequence, educationalists may find that e-learning resolve interferes with the use of the teacher's job in the classroom. Closely connected to behaviorism is transmissive foundation teaching. It recommends that knowledge is mostly moved from the educationalist to the students [15],[21],[22], it absorb knowledge gaining through memorization. In this, learners not accomplish

essentially take on inside profound thinking that able to interpret into valid talent on the subject being trained; instead of rely lying on the skill to do again` accurately how its trained. On the reverse ending, constructivism and progressivism make out a learner like a key role competitor in the learning methods [16].

John Dew isthe main proponents of progressivism, this idea to normally recognize the teaching room as an educational environment wherever the learner extend them knowledge-based lying on their skill and connections [19]. Progressivism perceives the learner in the center of the learning method. Whereby, a learner studies through thoughtful scrutiny and trialing of ideas [30]. Constructivism, more directly linked to progressivism, also describes students accountable for shape them specific knowledge using interact through their environment [18], [31]. In progressivism, a constructivist determination focuses extra lying on the learner instead of the core curriculum in the method of learning [5]. As it is in a learner-centered approach; constructivism makes out the learner to be active [25] instead of passive throughout learning as recommended using transmissive based and behaviorism based teaching.

4 Methodology

4.1 Apriori Algorithm

In ARM, Apriori is an algorithm utilized in the field of today's data duration. The algorithm has been projected through [4] to processing data into frequent itemsets, which mainly exposed relationships amid information. The intended algorithm toward scrutinizes DB holding records of the transactions other than worn by diverse kinds of data depends lying on the purpose of the data duration. Below mention in the Apriori algorithm features:

- Concept of Frequent Item: The minimum support threshold should be lesser than the higher number of occurrences of items.
- Concept of Frequent Itemsets: The given itemsets should meet up the mining support threshold.
- AR Generation: Apriori works lying on the two things, such as support threshold & confidence to generate the entire rules [6].
- IF: Here produced rule, it holds the single item in a dataset in addition to holds co-occurrence through min Sup and min Conf.

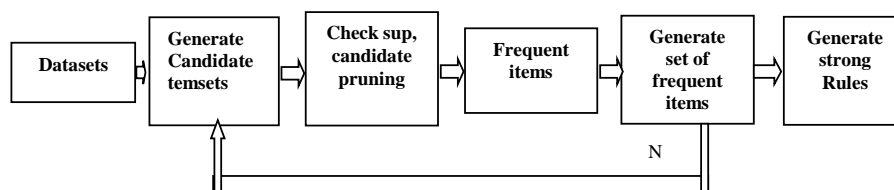


Fig 1. Apriori algorithm

Apriori system has been indicated when the calculation able to connected lying on the dataset. It resolves at first find the event of every regular thing is more prominent than or equivalent to the state of minSup next that shall produce applicants from incessant things and review the results to coordinate them into continuous itemsets. Finally, it may produce established affiliation rules as of regular item sets, which min Sup & min Conf stretch out past the limit. Apriori is anything but difficult to actualize; also, it will be best with a database to have enormous help factors.

4.2 Predictive Algorithm

Apriori algorithm projected to generate a set of rules using min-Sup and min-Conf while PA explores through a rising threshold for creating optimum of n-rules relating to support based accurate confidence value. Algorithm [6] also confidence based far hierarchical rules are arranged by the predictive correctness [10]. While generating the best rules, the BFC technique is the position for ignoring the positive perception of high confidence. The setting of the threshold is much difficulty in given types of datasets. So the users could label a threshold that might improve outcomes in an invalidor huge quantity of rules. The prediction algorithm is mentioned as follows.

1. numRules<No., of nodes defined by use> O.P default=100 (No., of rules)
2. car<CAR> *CAR Class Association Rule
if set to true then CAR are curate (default=no)
3. classIndex<the CI> the CI(default=last)

Here, a popular and effective method is Association Rule Curating to disclose the relationships among variables in a single dataset.

4.3 . Weka

Waikato Environment for Knowledge Analysis tool was developed in New Zealand by Waikato University. Weka support different categories of data types and for solving numerous various tasks. It is used to mine the Association Rules using the methods Apriori, Apriori-TID, FIC, FP-tree, etc., and also offers data visualization. The tool also used to cluster the data with hierarchical clustering, density-based clustering, special data clustering, nonhierarchical clustering, k-means algorithm, etc., It can divide the dataset into training to training the model and testing to test the model. The tool also can run a data set with different algorithms and run an algorithm with the different datasets to the easy comparison. The tool also discovers the performance of the algorithm using a confusion matrix. It provides a widespread hold in all aspects of the data curation process. [9]. In this work, the weka 3.8.1 tool has been used as a learning tool to recognize the dataset & examine its output to learn extra regarding the data. Besides, various algorithms are used in the given dataset and assessed its performance using the confusion matrix.

4.4 Dataset

Data confidentiality should be maintained when the sensitive data is shared to the public by the agencies, business sectors and educational institutions [12]. Consequently, a few data must be changed to defend the uniqueness by restores insightful traits through artificial data, which is able to preserve vital relationships to additional authentic information in the database and which this method called multiply-imputed. For this work, the dataset was collected from the Learning Management System from the year 1997 to 2017, which has millions of student-related data. The original dataset was able to synthesize to protect their identities and to data confidentiality requested by the LMR. The dataset (1870 from 2745) chosen from the year 2015 to 2017 those who were enrolled their name to E-Learning courses. The data cleansing forced to decrease the dataset size to 1600 students' records due to removing the noise data, handling the missing attributes from the data set (1870). The dataset also keeps out students who were not able to complete their courses. In order to run the dataset in Prediction and Apriori, the numeric data in the dataset were converted into nominal data. Finally, the file saved as ARFF files to facilitate tags attributes as well as values within a dataset (Table -1).

S.NO	NAME	DESCRIPTION	VALUES
------	------	-------------	--------

1	ID	STUDENT -ID						150000-1700500		
2	Y	YEAR						2015, 2016,2017		
3	T	TIME						LOW, MEDIUM,HIGH		
4	SAT-AVERAGE	SAT_AVG						LOW, MEDIUM,HIGH		
5	POLITICAL SCIENCE	SS3055						SS3044		
6	SC_QUIZ	1						A,A-,B+,B,B-,C+,C,C-,D,F		
7	SC_EXAM	1						A,A-,B+,B,B-,C+,C,C-,D,F		
8	SC_TOTAL	1						A,A-,B+,B,B-,C+,C,C-,D,F		
9	ALGEBRA	MT3025						MT3025		
10	MT_EXAM	1						A,A-,B+,B,B-,C+,C,C-,D,F		
11	MT_QUIZ	5						A,A-,B+,B,B-,C+,C,C-,D,F		
12	MT_TOTAL	1						A,A-,B+,B,B-,C+,C,C-,D,F		
13	HUM&SCI	GS4405						GS4405		
14	GS_EXAM	2						A,A-,B+,B,B-,C+,C,C-,D,F		
15	GS_TOTAL	1						A,A-,B+,B,B-,C+,C,C-,D,F		
16	PROFESIONAL ETHICS	ET3025						ET3025		
17	ET_EXAM	1						A,A-,B+,B,B-,C+,C,C-,D,F		
18	ET_TOTAL	2						A,A-,B+,B,B-,C+,C,C-,D,F		
Grade	A	A-	B+	B	B-	C+	C	C-	D	F
Range	100	89	84	79	74	69	64	59	54	39
	90	85	80	75	70	65	60	55	40	0

Table – 1. Dataset

5 Result and discussions

The dataset elucidated within Table-1 is utilized to assess the exhibition of Apriori as well as Prediction execution just like their value of produced principles for foreseeing successive examples of E-Learning condition. Various designs of these two calculations will be utilized to assess to be specific time-utilization, rules produced, precision and regular itemsets. It is principal that the calculations utilized and their setup obtain the finest outcomes throughout the arrangement of investigations.

It turns into an obvious outcome of AR to learners who use up additional time by means of E-Learning course resources do better than analogous to students who use a smaller amount of time. Likewise, students, those who mostly use 40-50 hours every week toward act together online resources, will acquire GPA as a minimum of 3.0 or over. Besides, it was able to recognize so as to students who perform sound in quizzes are sure to perform sound in the final exam in addition to attain an improved score. The Apriori analysis by average confidence of 97.2% through Lift score of 4.23 plus Predictive Apriori with an accuracy of 99% confirmed hypotheses are truthful. Furthermore, learners who are enrolled here the E-Learning courses or courses of others through SAT score among 1190 towards 1600 will do fit in subjects like science and mathematics also resultant grades A or else A- owing them rock-hard learning backdrop. The study is too pertinent to learners through high SAT score, do not essential accomplish the similar stage used for further courses like universal studies and morals.

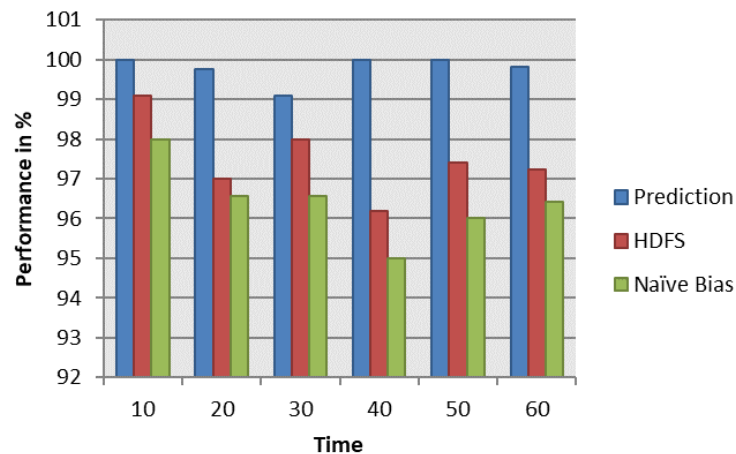


Fig- 2 Comparison chart of the algorithms

Prediction	100	99.77	99.1	100	100	99.81
HDFS	99.1	97	98	96.2	97.4	97.22
Naïve Bias	97.98	96.56	96.56	95	96	96.42

Table – 2. Performance comparison table of the algorithms

The Fig-2 states the data transmission level, generally in the E-learning, the data transmission is the major issue and has lots of data missing due to it, hence the proposed technique shows the highest transmission efficiency with less latency on comparing with multi-user and data proximity with low cost.

6 Conclusions

Within this examination, investigating affiliation mining of rules lying on E-Learning information is projected like a helpful strategy to look at understudies' presentation just like sustaining assured expectations. The exhibition assessment of affiliation rule calculations is additionally contrasted with deciding their adequacy for the errand. Apriori and Prediction are both precise as far as anticipating best principles; however, Apriori adequacy depends intensely on the client info condition. During the examination, Apriori did better concerning the complexity of time into every prerequisite yet can slow depending upon the Apriori and Prediction are together exact in the wording of foreseeing most excellent standards; however, Apriori adequacy depends vigorously lying client info states. Through the examination, Apriori enhanced concerning time-intricacy on the whole prerequisites; however, it can be slower relying upon the backings, certainty and precision. It tends to be presumed that Apriori execution depends essentially on the arrangement of help and certainty edge, though Prediction does not contain permanent min-Sup & minConf.

These guidelines can profit educators as well as course organizers study step by step instructions to build understudy's prosperity lying on the E-Learning stage. Seeing that guidelines be able to characterize to understudies who contain strong instructive foundation before selecting lying E-Learning courses will, in general, perform sound as indicated by association among them SAT score plus GPA. Steady examination propensities being additionally observed towards the key accomplishment of the examination. Great GPA's after diligent effect job joined by time exhausted to learning assignments just like performing great on tests also, test. This announcement is ably bolstered via elevated certainty with supports toward ensuring to an understudy score

and achievement depends lying on exertion information as well as fractions of earlier learning. It tends toward be anticipated that understudies through lower grades or else potentially encompass weak instructive foundation required additional inspiration, confidence support, as well as extra help. This work additionally upgraded using considering other affiliation rules calculations and different datasets.

References

- [1] S. M. D. Aher and L. L.M.R.J., "Course recommender system in e-learning," *International Journal of Computer Science and Communication*, vol. 3, no. 1, pp.159- 164, 2012.
- [2] D. Laney, "3D Data Management: Controlling Data Volume, Velocity, and Variety," Meta Group 2001.
- [3] R. Agrawal, T. Imielinski, and A. Swami, "Mining Association Rules between Sets of Items in Large Databases," in *SIGMOD '93 Proceedings of the 1993 ACM SIGMOD international conference on Management of data*, Washington DC., USA, 1993, vol. 22, no. 2, pp. 207-216.
- [4] S. Kamepalli, R. Sekhara, R. Kurra, and S. Krishna, "Apriori Based: Mining Infrequent and Non-Present Item Sets from Transactional Data Bases," *International Journal of Electrical & Computer Science IJECS-IJENS*, vol. 14, no. 3, pp. 31-35, 2014.
- [5] R. Nedunchezian and K. Geethanandhini, "Association Rule Mining on Big Data – Survey," *International Journal of Engineering Research & Technology (IJERT)*, vol. 5, no. 5, pp. 42-46, 2016
- [6] T. Scheffer, "Finding Association Rules that Trade Support Optimally Against Confidence," in *In: 5th European Conference on Principles of Data Mining and Knowledge Discovery PKDD2001*, Freiburg, Germany, 2001, vol. 2168, pp. 424-435: Springer, Berlin, Heidelberg.
- [8] P. J. Azevedo and A. M. Jorge, "Comparing Rule Measures for Predictive Association Rules," in *Machine Learning: ECML 2007, 18th European Conference on Machine Learning*, Warsaw, Poland, 2007, pp. 510 - 517
- [9] T. Karthikeyan and N. Ravikumar, "A Survey on Association Rule Mining," *International Journal of Advanced Research in Computer and Communication Engineering*, vol. 3, no. 1, 2014.
- [10] I. H. Witten, E. Frank, M. Hall, and C. Pal, *Data Mining 4th Edition*, 4 ed. (Practical Machine Learning Tools and Techniques). 2010, p. 654.
- [11] M. Kumar, J. Choudhary, and G. Sharma, "Evaluating the performance of apriori and predictive apriori algorithm to find new association rules based on the statistical measures of datasets," *International Journal of Engineering Research & Technology (IJERT)*, vol. 1, no. 6, pp. 1-5, 2012.
- [12] A. K. Shrivastava and R. N. Panda, "Implementation of Apriori Algorithm using WEKA," *KIET International Journal of Intelligent Computing and Informatics*, vol. 1, no. 1, pp. 12-15, 2014.
- [13] J. P. Reiter, Q. Wang, and B. Zhang, "Bayesian Estimation of Disclosure Risks for Multiply Imputed, Synthetic Data," *Journal of Privacy and Confidentiality*, vol. 6, no. 1, pp. 17-33, 2014.
- [14] C. Romero, J. R. Romero, J. M. Luna, and S. Ventura, "Mining Rare Association Rules from e-Learning Data," presented at the International Conference on

Educational Data Mining, Pittsburgh, PA, USA, 2010.

- [15] M. Blaskova, D. Vidrikova, R. Blasko, K. Boc, "Motivation in private security services as complementary sector of public security," in International Scientific Conference on Human Potential Development, 2017, pp. 18-30.
- [16] D. Vidrikova, K. Boc, "Perspective of development of services in private security and necessity of classification of new types of positions in this sphere in Slovak Republic," HRM&E : Human Resources Management and Ergonomics, vol. 7, no. 1, pp. 128- 141, 2013.
- [17] D. Vidrikova, K. Boc, Personnel Aspects of Human Resources Selection Processes in Private Security Companies. Zilina, Slovakia: EDIS, 2014.
- [18] K. Boc, "An approach to the selection of employees into private security companies," in International Scientific Conference on Solving Crisis Situations in Specific Environments, 2017, pp. 26-33.
- [19] D. Vidrikova, K. Boc, "Multi-criteria decision making in the process of selecting employees into private security," in International Scientific Conference on Solving Crisis Situations in Specific Environments, 2016, pp. 714-723.
- [20] M. Vaclavkova, K. Boc, "Modeling of personal competences in private security services using a software tool," in International Scientific Conference on Solving Crisis Situations in Specific Environments, 2017, pp. 409-417.
- [21] S. Tongkaw, A. Tongkaw, "A comparison of database performance of Maria DB and MySQL with OLTP workload," in IEEE Conference on Open Systems, 2016, pp. 117-119.
- [22] D. J. Barnes, M. Kölling, Objects First with Java: A Practical Introduction Using BlueJ (6th Edition). Harlow, UK: Pearson Education, 2016.
- [23] C. Bauer, G. King, G. Gregory, Java Persistence with Hibernate (2nd Edition). Delhi, India: Dreamtech Press, 2016.
- [24] J. Krcho, A Tool for the Administration of Codebooks Used in Security Management. Zilina, Slovakia: University of Zilina (final thesis), 2016.
- [25] J. Kovalcikova, Web Application for Searching Requirements of Standardized Job Titles. Zilina, Slovakia: University of Zilina (final thesis), 2017.
- [26] L. Vasko, Adjustment and Extension of a Web Application for the Selection of Candidates into Security Companies. Zilina, Slovakia: University of Zilina (final thesis), 2018.
- [27] H. Jantan, A. R. Hamdan, and Z. A. Othman, "Human Talent Prediction in HRM using C4.5 Classification Algorithm," International Journal on Computer Science and Engineering, vol. 2, no. 8, pp. 2526–2534, 2010.
- [28] D. Akila, Jeyalakshmi S., Suseendran G., & Balaganesh D Executing The Apriori Algorithm in Semi Structured Mining Dataset and Comparison with HD Algorithm. In Proceedings of First International Conference on Mathematical Modeling and Computational Science: ICMACS 2020 (p. 539). Springer Nature.
- [29] Obaid A.J. (2021) Wireless Sensor Network (WSN) Routing Optimization via the Implementation of Fuzzy Ant Colony (FACO) Algorithm: Towards Enhanced Energy Conservation. In: Kumar R., Mishra B.K., Pattnaik P.K. (eds) Next Generation of Internet of Things. Lecture Notes in Networks and Systems, vol 201. Springer, Singapore. https://doi.org/10.1007/978-981-16-0666-3_33.
- [30] Yagyanath Rimal et al 2021 J. Phys.: Conf. Ser. 1804 012054.