A COMPARATIVE ANALYSIS OF TAXONOMY OF EDUCATIONAL DATA MINING APPLICATIONS

Supported by tetfund TERTIARY EDUCATION TRUST FUND

T. J. Odule*, A. O. Adesina and K-K. A. Abdullah
Department of Mathematical Sciences Olabisi Onabanjo University, Ago-Iwoye. Nigeria
*Corresponding author: tola.odule@oouagoiwoye.edu.ng

Received: November 29, 2019 **Accepted:** February 21, 2020

Abstract: This paper investigates the different potential uses of EDM and related techniques that can be utilized to address the issues of these uses. The emphasis is on the end goals of these uses while seeking to assemble the uses in classifications with comparative drives and ensure that the proposed taxonomy covered all the contemporary research results. The proposed rundown of utilizations was contrasted with the current productions, it was discovered that the rundown is progressively comprehensive as far as EDM themes contrasted with past studies and writings, and represents an original and more qualified order in a developing area of speciality.

Keywords: Educational data mining, EDM uses, taxonomy, EDM responsibilities

Introduction

Data mining has been utilized for making disclosures in instructive situations throughout the previous couple of years back. In the most recent period, the accessibility of accessible datasets and more employments of web-based scholarship frameworks have gathered additional consideration regarding this area of specialisation. Educational Data Mining (EDM) in the field of utilizing information-digging systems in instructive conditions. Romero *et al*. (2013) characterize an LMS as "a brought together online data frameworks where the learning substance is overseen and learning exercises are composed. LMS speaks to a progressively broad term for an innovation structure that supports all parts of formal and casual learning forms, including learning executives, content administration, course management, and so on" (Wang, 2014). There exist different techniques and uses in EDM. These uses can pursue both practical inquiry destinations, for example, refining and upgrading scholarship worth, just as unadulterated inquiry goals, which will, in general, advance comprehension of the scholarship procedure. Besides the characterization of users depending on their destinations, EDM uses can likewise be arranged dependent on the focused on end client. The utilizations of EDM can focus on any of the partners associated with instructive frameworks, for example, students, teachers, heads and scientists themselves. Giving input, personalization and suggestions can improve the learning procedure of understudies. Settling on disclosures and giving choice emotionally supportive networks can help the instructors by improving showing execution and deciding. Heads are given assets and instruments to settling on choices and sorting out the organizations.

This paper investigates the different potential uses of EDM and related techniques that can be utilized to address the issues of these uses. The emphasis is on the end goals of these uses while looking to assemble the uses in classifications with comparative drives. Although endeavours were made to draw a line between various classifications of uses, it ought to be noticed that sometimes, there is no reasonable limit between the applications. Some examination may have a place with something beyond one of these classifications. In some different circumstances, an application can be utilized as an instrument for arriving at a different use.

In past examinations, potential utilizations of EDM have been presented in some cases in no particular request, some of the time-dependent on the number of research papers done in every one of the classifications (Romero and Ventura, 2010). Potential utilization of EDM has been gathered into classes dependent on their end objective. Various applications were gathered however much as could reasonably be expected to more readily feature the likenesses and contrasts.

End clients in EDM responsibilities

For a better explanation of the distinguished uses, we can investigate the objective clients of every use. This has the additional estimation of likewise demonstrating the potential uses for the target clients which have not been focused on yet. Target clients in instructive conditions are scholars (understudies), teachers (instructors), managers and scientists. Students have been the objective of EDM in different uses, for example, gathering understudies, creating suggestions and versatile frameworks. One significant objective of EDM, all in all, is refining the nature of scholarship; and during the time spent scholarship, dual gatherings of clients strike a chord first, for example, students and teachers. The greater part of the applications in classes of understudy demonstrating and choice emotionally supportive networks target teachers as their end clients. Understudy displaying gives a superior comprehension of understudies' condition of learning and choice emotionally supportive networks can straightforwardly enable instructors to settle on better choices for refining the scholarship procedure. This likewise relates to the director of instructive foundations settling on more elevated level choices.

Scientists additionally speak to a gathering of end clients, as the goal of the exploration is to comprehend the learning procedure, create speculations and test them (Chrysafiadi and Virvou, 2013). As a, for example, scientists can utilize interpersonal organization analysis (IOA) to locate profitable attributes in the forecast of understudy execution. Table 1 shows conceivable focused on clients of every application. Note that any examination in EDM may relate to more than one of these classes on the double.

Table 1: Expected beneficiaries of EDM uses

Item	Scholars	Instructors	Managers	Investigators
Foreseeing execution and attributes		✓	✓	
Recognizing unwanted understudy conduct		✓	✓	
Describing and Classifying understudies	✓	✓		
Interpersonal Organisation Analysis		✓	✓	✓
Giving feedbacks	✓	✓	✓	
Making alarms for partners		✓	✓	
Arranging and booking	✓	✓	✓	
Developing curriculum		✓		
Developing concept maps		✓		✓

Materials and Methods

The strategies for EDM are equivalent to those in the information mining field as a rule. There are numerous strategies to use in EDM for every one of the different applications. Among these strategies, the most utilized are order and relapse (1), bunching (2), affiliation standard mining (3), revelation with models (4), anomaly recognition (5), interpersonal organization investigation (6), content mining (7), consecutive example mining (8) and perception procedures (9) additionally alluded to as refining of information for human judgment. Given research by Romero and Ventura in 2009, the most normally applied information mining errands are relapse, bunching, order and affiliation principle mining.

Uses and undertakings in EDM can be sorted dependent on various attributes. Various reviews of EDM exist, which have recorded potential uses of EDM. Considering past reviews and perusing into the examination models they have given, just as the on-going investigations distributed in diaries of instructive information mining, another rundown of EDM classes is therefore proposed. In this rundown, endeavours were made to consider every one of the classes referenced in past overviews and the writing, just as new classifications which are viewed as important to be included. These new classifications of uses can be clarified by the development of enthusiasm for EDM

Thirteen (13) classifications of uses were distinguished, as appeared in Fig. 1, shaping another scientific classification custom-made explicitly to EDM, along these lines making EDM a particular subfield of information mining. Four uses are assembled under "Understudy displaying", six under "Choice Support Systems" while remaining three are introduced as "Other" because they contrast from different applications.

In the sections that follow, these applications are described and delineated with the assistance of research models identified with every class of utilizations for greater lucidity.

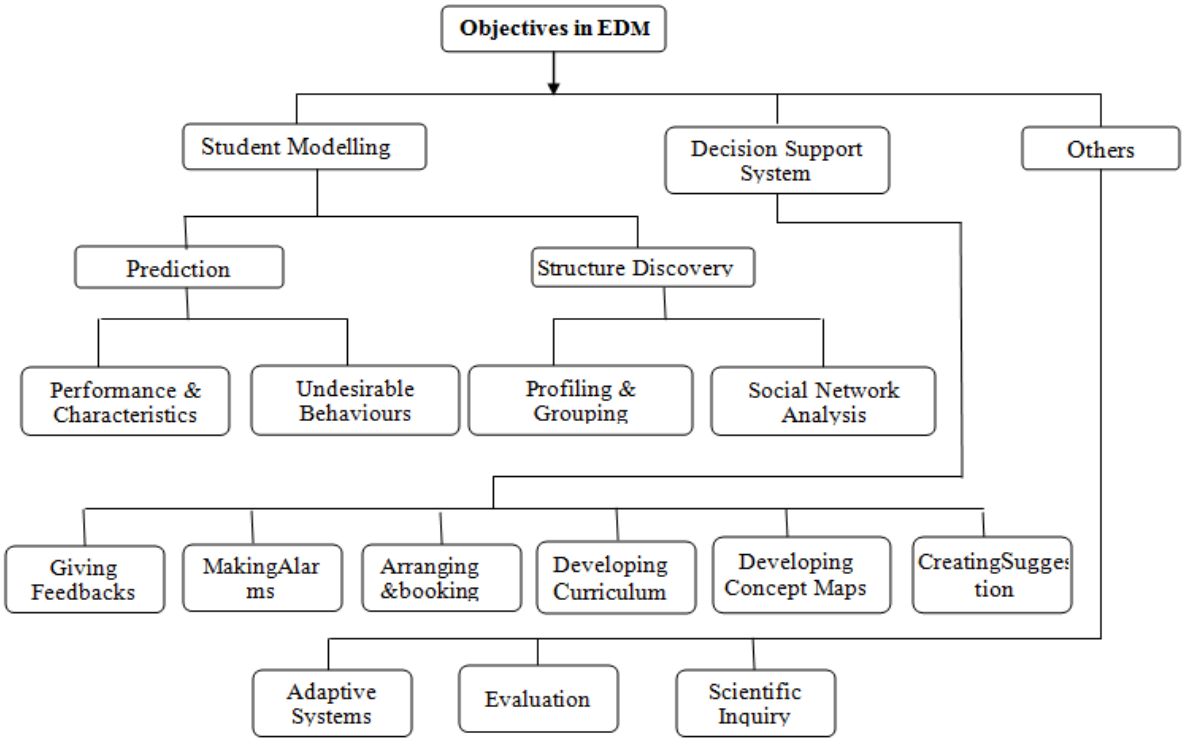

Fig 1: Taxonomy of applications in EDM

Understudy modelling

Understudy displaying is a procedure dedicated to speaking to subjective parts of understudy exercises, for example, dissecting the understudy's presentation or conduct, segregating fundamental confusions, speaking to understudies' objectives and plans, recognizing earlier and procured learning, keeping up a long-winded memory and depicting character qualities.

All uses in this classification provide a template depicting understudies endeavouring to arrive at their target. In light of the writing, there are various attributes in understudy demonstrating, to be specific, information and abilities (1), blunders and misguided judgments (2), scholarship styles and inclinations (3), emotional and subjective components (4) and meta-psychological elements (5).

Demonstrating understudy exercises and conduct can be utilized for foreseeing esteems speaking to understudies or finding structures that depict understudies. Thus, there are two sub-classes in understudy demonstrating: forecast and structure revelation. In the forecast, a particular property (speaking to attributes) that will be anticipated is typically known and in structure disclosure, the specific attribute may not be known or on the other hand, it might just be characterized as a structure, rather than a solitary property. It is additionally critical to specify that there probably won't be an unmistakable difference in these two sub-classifications in every scenario; yet as there are sufficient contrasts in the goal of these two gatherings, it appears to be desirable over recognize them.

Foreseeing scholar rating, attainment of scholarship results or features

In this arrangement of utilizations, the goal is to assess a worth or variable portraying understudies. This worth can demonstrate understudies' exhibition, accomplishment of learning results or trademark. A large portion of the current

productions are centred around the expectation of understudies' scholastic execution, yet thinks about likewise investigate trademark, for example, a joint effort with different understudies. The most generally utilized strategies for anticipating understudy execution and qualities are relapse and characterization however different procedures have likewise been utilized, for example, grouping and highlight determination. Highlight determination and forecast procedures have been utilized in this investigation

Recognizing unwanted understudy conduct

This arrangement of assignments are like the expectation of understudy execution and qualities, however in this classification, the emphasis is basically on recognizing unfortunate understudy conduct, for example, low inspiration, incorrect activities, tricking, dropping out, scholarly disappointment, and so on. The primary information mining strategies utilized in this classification of utilizations are grouping and bunching yet different procedures are likewise relevant, for example, include determination and anomaly location. A case of this gathering of utilizations is the exploration done by Bravo and Ortigosa (2009) where a methodology was suggested for recognizing potential side effects of low execution in e-getting the hang of utilizing creation rules.

Describing and classifying understudies

As the title of this classification suggests, the target in this arrangement of uses is to describe understudies depending on various factors, for example, attributes and learning, or utilizing the data to gather understudies for different purposes. Gathering understudies should be possible dependent on different properties of profile data. This undertaking is frequently not the same as bunching comparable understudies with one another, as the intention is to gather understudies to supplement one another. Likewise, when bunching understudies, one is searching for the best divergence between groups, however, this may not be the situation in gathering errands. When utilizing a gathering task for framing groups in a course venture, one likes to have bunches that are comparative, while involving different understudies that can supplement one another. In a manner like different classifications of utilizations, various information mining techniques can be exploited for these errands, for example, highlight choice and bunching.

Interpersonal organization analysis

In this classification of uses, the reason for existing is to acquire a model of understudies as a diagram, indicating various potential connections among them. In different uses of demonstrating, the emphasis is generally on people, however, in interpersonal organization analysis (IOA), the attention is on the connections between people. For instance, a joint effort is a property allocated to the connection among people, and to contemplate it, one must model the connections just as the people.

Decision support systems

Another real gathering of utilizations in EDM is the choice of emotionally supportive networks. Applications committed to this classification upgrade the way toward learning by helping partners decide. Instances of this class are: giving input, making alarms, arranging, producing suggestions and upgrading the courseware. The objective of this choice emotionally supportive networks is, for the most part, the teacher, however, it can likewise be the understudy, directors or analysts.

Giving feedbacks

Information investigation and perception can be utilized as one piece of numerous different applications, yet it can likewise be an application-independent from anyone else in instructive conditions by giving valuable data to instructors and heads to assist them with basic leadership. Thus, the motivation behind this classification of uses is to discover and feature the data identified with course exercises which might be useful to instructors and chairmen and give them criticism. The aftereffects of a large portion of the applications gathered in "Understudy displaying" can be utilized for making reports. Instances of this are: giving criticism on understudy execution or qualities, depicting the associations and coordinated efforts through interpersonal organization examination and making feedback from the summary data extricated with the assistance of profiling strategies.

Making alarms for partners

This class of utilizations is like uses in the understudy demonstrating classification. As a rule, the reason for existing is to anticipate understudy qualities and distinguish undesirable conduct, however, it's, for the most part, fills in as an online instrument for educating partners or making cautions continuously. Instances of circumstances in which cautions might be required are in instances of low inspiration, abuse, tricking, and so forth. A case of concentrate in this classification is the examination of Knowles, which presents a failure timely cautioning framework utilizing factual models and relapse (Knowles, 2015).

Arranging and Booking

The goal of this classification of uses is to assist partners with the undertaking of arranging and planning. It can support instructors and overseers with arranging upcoming options or asset distribution, aiding the confirmation and advising forms or some other errands associated with arranging and planning (Romero and Ventura, 2010). It can likewise assist understudies with option enlistment arranging, where there is a shared opinion with recommendation frameworks. In scientific investigation to arrange and book, different strategies were utilized, for example, a revelation with models, group investigation, and order.

Making courseware

Courseware is known as instructive programming giving substance, recordings, assessments and supplementary scholarship resources. In this classification of uses, the goal is to enable the instructor to make or grow study material consequently utilizing understudy use data.

Creating concept maps

Idea maps are "graphical instruments for sorting out and speaking to learning" (Novak and Alberto, 2008). In this classification of uses, the goal is to create idea maps of different angles to enable instructors to characterize the procedure of training. At the end of the day, idea maps give space prototypes to instructors. The prototype assists in relating various ideas to one another (for example finding out connections). Instances of idea maps are pecking order of points in the course material, connections of aptitudes and assessment things, relationship of test things and information parts, and so on.

Creating recommendation

In the majority of the studies and books on EDM applications, creating a proposal is introduced for the most part as making suggestions to understudies. In any case, suggestions can be focused on any partners. Instances of this classification of uses are option proposals to understudies or assessment thing suggestions to teachers. The most widely recognized techniques in recommendation frameworks are collective separating, content-based strategies, affiliation principle-based calculations and mixture approaches additionally utilized in EDM. Another strategy for producing suggestions is utilizing revelation with models.

Other applications

Adaptive Systems

This class of uses is identified with the utilization of smart frameworks in PC based learning, in which we need the framework to adjust to the client's conduct (otherwise called

"personalization"). This application is significant because, in numerous web-based learning frameworks, various students with various needs are engaged with the framework. What's more, as the quantity of members develops, it winds up more diligently to meet the particular needs everything being equal. Versatile frameworks can enable us to address the issues of each student. This adjustment can appear as adjusting the course material, guidance pace, giving clues, requesting and producing tests, and so on.

Evaluation

In this class of EDM uses, the goal is to give an assessor to support the teachers. This should be possible in experimental scholarship conditions and PC-based options. Assessment is one of the parts of the instructive condition which may not generally be natural, particularly in PC based learning situations. For instance, assessment is not well-characterized areas has been looked into in shrewd mentoring frameworks and is known as a space that comes up short on an authoritative arrangement (or the arrangement is reliant on the issue's origination) (Lynch *et al*., 2006). Therefore, the assessment in these spaces is testing.

Scientific inquiry

Like different area of specialisation, speculations and theories on the way toward scholarship and potential enhancements are utilized in instruction. One utilization of instructive information mining can be trying or notwithstanding creating hypotheses dependent on the different records in enormous datasets. This classification of utilizations, for the most part, focuses on the scientists as the end client, yet any of the created or tried speculations can be utilized in different applications focusing on different partners later.

Comparative Discussion of Findings

In the principal study distributed in Journal of Educational Data Mining (JEDM) [6], four regions of use have been referenced, for example improving understudy prototypes (1), refining area prototypes (2), concentrating the academic help given by studying programming (3) and logical examination into scholarship and students (4). Understudy demonstrating all in all or, as expressed by Baker and Yacef (2009), refining understudy prototypes is among the most referred to spaces of scientific examination in EDM. This gathering has been unfurled into increasingly nitty-gritty classes in the past area. Improving space models dependent on the goal of use can be known as a component of choice emotionally supportive networks, for example, creating idea maps or demonstrating reports. Likewise, concentrating academic help and logical inquiry can be outlined as a logical request.

In another study about EDM uses composed by Romero and Ventura (2010), 11 classifications of use were recommended dependent on 300 research concentrates finished before 2010. This overview has been incredibly valuable as a source of perspective for this paper, as it gives numerous guides to every one of the presented classifications just as strategies and systems utilized in them. The classes of usages exhibited in this investigation are: Analysis and Visualization of Data, Providing Feedback for Supporting Instructors, Recommendations for Students, Predicting Student's Performance, Student showing, Detecting Undesirable Student works on, Grouping Students, Social Network Analysis, Developing Concept Maps, Constructing Courseware and, Planning and Scheduling

These uses are altogether referenced in the proposed rundown of utilizations with a couple of modifications with increments. A piece of the end target examination and information perception is to give feedbacks (or input, as referenced previously). Also, giving reports (or criticism) isn't just constrained to supporting educators; it can likewise target understudies and heads.

In another review distributed in 2013 by similar creators, a similar rundown, with a couple of changes, was presented. The principle point displayed in this study is ordering applications and assignments dependent on the end client. This is possibly useful in building up a superior comprehension of the issues identified with EDM. The end clients proposed by Romero and Ventura (2013) are students, instructors, directors and specialists. A similar order has been utilized in this investigation. The conceivable focused on end clients of every application have additionally been indicated dependent on a similar grouping.

The book Educational Data Mining: Applications and Trends (Pena-Ayala, 2013) incorporates a prologue to the speciality area that talks about various uses, techniques and datasets. In the second, third and fourth piece of the book, there are some contextual investigations sorted in three gatherings (understudy displaying, appraisal and patterns). The general objectives of EDM are presented in the initial segment of the book to demonstrate the wide scope of uses. The accompanying themes are examined. Understudy displaying, Forecasting understudy execution and learning results, Generating proposal, Analysing student's conduct, Collaborating with partners, Field organization examination, Retaining and refining options, Learning the impacts of academic help that can be given by studying programming and, Improving logical information about scholarship and students via structure, finding or refining prototypes of the understudy.

This rundown imparts numerous points to past works, however, the contextual analyses in the third piece of the book entitled "Evaluation" include something new. These investigations spread assessment in cases in which assessment isn't a simple assignment. Improving assessment as a conceivable use of EDM had not been tended to previously. That was one significant explanation behind this investigation to indicate a classification of uses under "Assessment", as examined prior.

At last in one of the latest overviews on EDM in advanced education, four classifications of utilizations have been checked; specifically course the executives frameworks, understudy practices, choice emotionally supportive network in advanced education and understudy maintenance and wearing down. This rundown of utilizations presents various EDM errands, however, its treatment isn't evident, and does not spread the extensive scope of uses. For instance, understudy maintenance can be adjudged as a piece of understudy conduct and there are various subjects in the course the board frameworks which may bring about various uses.

Figure 2 is an outline of the relative investigation of discoveries from the writing and the proposed scientific classification of EDM applications. This Figure indicates what applications every one of the referred to references spread. A portion of the marks for classifications of uses may fluctuate; however, at last, they address similar applications utilizing an alternate wording. For instance, the proposed scientific categorization likened space structure investigation, referred to in a couple of sources (Romero *et al*., 2010) with developing the idea maps. The descriptions accommodated these ideas are not the equivalent; nonetheless, in light of the models given by the references, it is thusly gathered that they speak to a similar gathering of uses. This Figure demonstrates the themes that were referenced more regularly than others or the subjects for which there has been fewer spotlights as on others.

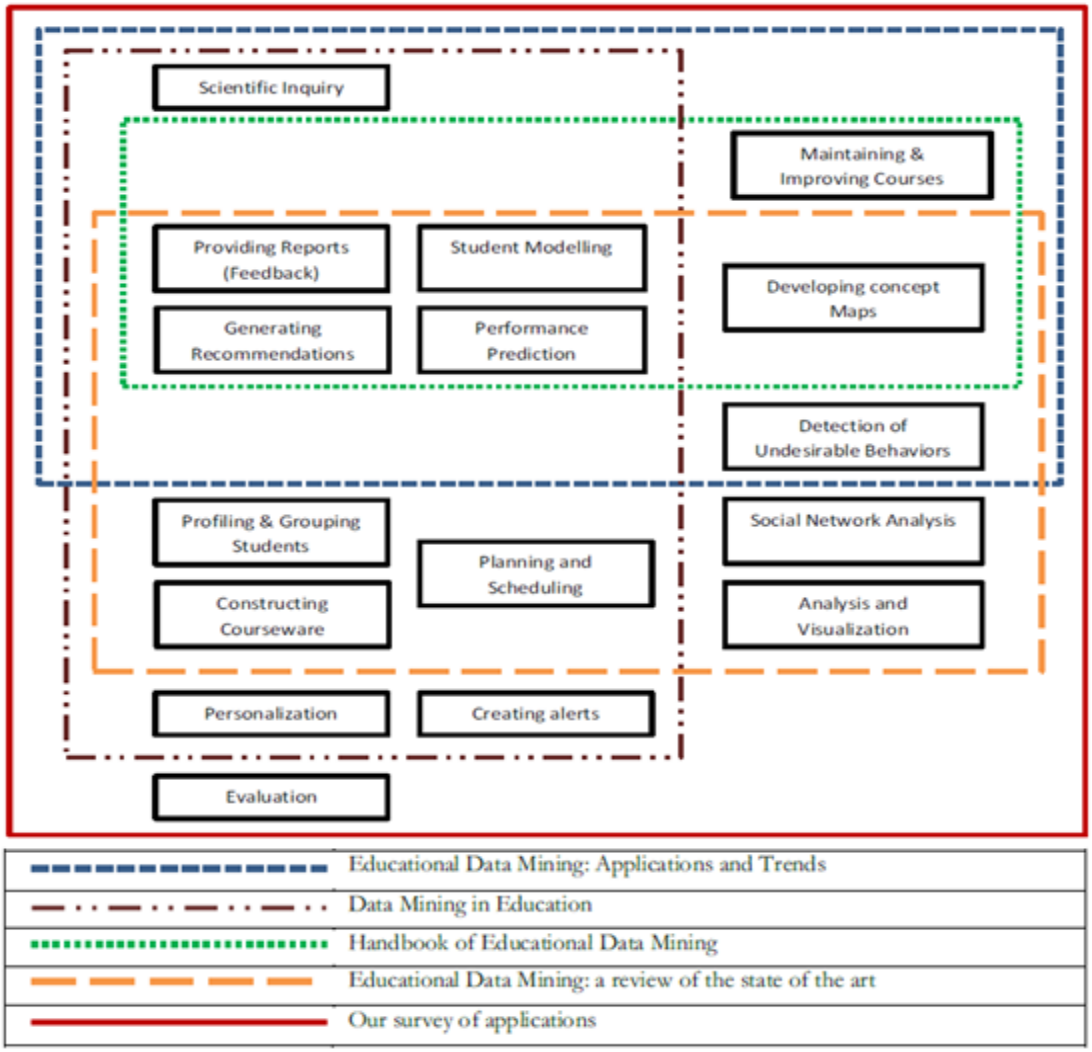

Fig. .2: Evaluation of described uses in EDM

This section condenses the overviews and writing identified with utilization of EDM and the issues which have been tended to in each.

Mill operator *et al*. utilized Lasso highlight determination for recognizing understudy attributes which affect the learning. They contrast the information mining strategy and different models and inspect if the chose highlights can be utilized for foreseeing understudy execution (Dee Miller *et al*., 2015). Zimmermann *et al*. (2015) acquainted a prototype-based methodology with anticipate graduate-level execution utilizing markers of undergrad level execution. Galyardt and Goldin (2015) have utilized the on-going understudy utilization information to improve the forecast precision of the framework in ITS. Research by Waters *et al*. (2014) distinguishes the joint effort of understudies in accessible options utilizing Bayesian grouping.

In another investigation, Dekker *et al*. (2009) have utilized a choice tree classifier to foresee understudy drop in an electrical building program. Lykourentzou *et al*. (2009) utilized various AI strategies, for example, Support Vector Models and neural systems for understudies dropout forecast. Azarnoush *et al*. (2013) proposed a technique for student division utilizing a difference measure dependent on an arbitrary woods; Kinnebrew *et al*. (2013) utilized succession mining strategies to recognize learning personal conduct standards separating unmistakable gatherings of understudies. Harley *et al*. (2013) considered the assignment of grouping and reporting of students depending on their associations with a canny mentoring framework.

Rallo *et al*. (1999) utilized information mining and interpersonal organization examination to display the elements and structure of educative online networks. Likewise, Reffay and Chanier (2003) utilized informal community examination to gauge attachment in communitarian separation learning situations. In a similar vein, Reyes and Tchounikine (2005) contemplated basic properties of learning gatherings dependent on a social point of view utilizing informal community investigation procedures.

In the region of utilizing choice help to give reports, Romero *et al*. (2013) utilized affiliation principle mining to give input to educators from the numerous decision test information. In another examination, Leah *et al*. (2010) have built up an early cautioning framework to make alarms for teachers utilizing execution expectation. Hsia *et al*. (2008) improved course arranging by building up the likelihood of understudies finishing their programmes depending on the understudy's inclinations and calling. The investigation of Delavari *et al*. (2008) finds unequivocal information that can be valuable for basic leadership forms just as proposing investigative rules for advanced education establishments. Huang *et al*. (2009) utilized group examination, choice tree calculation and neural systems for arranging courses.

A case of Courseware creation was cited in the exploration of Garcia *et al*., (2009) in where a framework was proposed for

creating, refining and keeping up online programmes utilizing affiliation principle mining and community separating. The examination of Agrawal *et al*. (2014) presents an investigation pilot for considering electronic course readings by referring to ideas which understudies are perusing. As anextra model, Lee *et al*. (2009) utilized an Apriori calculation to build up a naturally developed idea guide of learning, gave to instructors. Vialardi *et al*. (2009) utilized an exhibition indicator model for creating suggestions. The indicator prototype forecasts the achievement of every understudy in each course and will prescribe options the understudy is well on the way to be effective in. In another investigation, O'Mahony and Smyth (2007) build up a programme recommendation framework utilizing cooperative separating.

In the region of versatile frameworks, Alaofi *et al*. (2015) investigate the individualization of an advanced library utilizing the understudy's summary data to improve list items. In another model, Tang *et al*. (2000) suggest a strategy for customizing curriculum development utilizing information-digging strategies for separation learning conditions. A case of Evaluation use is the exploration of Mallavarapu *et al*. (2015) which proposes a computational technique to quantify and follow understudies' altitudinal thinking in an open-finished re-enactment. As another model, Hao, Shu and von Davier (2015) proposed another strategy for scoring a game/situation based undertaking utilizing separation work.

Conclusion

This study undertook a comparative analysis of the taxonomy of applications in EDM to bring it up-to-date with what currently obtains in the field and ensure that the proposed taxonomy covered all the contemporary research results. Comparable applications are assembled into classes and sub-classifications to propose scientific classification of assignments in EDM. This arrangement of uses depends on the end targets. Some delegate models were picked for every class. These models can enable clients to comprehend classifications better. Be that as it may, they don't cover all the potential undertakings in every class. The proposed rundown of utilizations was contrasted with the current productions. It was discovered that the rundown is progressively comprehensive as far as EDM themes contrasted with past studies and writing, and represents an original and more qualified order in a developing area of speciality. With the development of utilizing PCs and accessibility of information, it is accepted that the employments of EDM will likewise develop, prompting yet new applications.

Conflict of Interest

Authors declare there is no conflict of interest related to this study.

References

Agrawal R, Sreenivas G, Anitha K & Krishnaram K 2014. Study navigator: An algorithmically generated aid for learning from electronic textbooks. *J. Edu. Data Mining* (JEDM), 6(1): 53-75.

Alaofi M & Rumantir G 2015. Personalisation of generic library search results using student enrolment information. *J. Edu. Data Mining* (JEDM), 7(3): 68-88.

Azarnoush B, Jennifer MB, George CR, Bianca L, Bernstein BL & Atkinson KR 2013. Toward a framework for learner segmentation. *J. Edu. Data Mining* (JEDM), 5(2): 102-126.

Baker Ryan SJD & Yacef Kalina 2009. The state of educational data mining in 2009: A review and future visions. *J. Edu. Data Mining* (JEDM), 1(1): 3-17.

Bravo J & Ortigosa A 2009. Detecting symptoms of low performance using production rules. *Int. Working Group on Edu. Data Mining*.

Chrysafiadi K & Virvou M 2013. Student modelling approaches A literature review for the last decade. *Expert Sys. with Applic.*, 40(11): 4715-4729.

Dee Miller L, SohLeen-Kiat SA, Kupzyk K & Nugent G 2015. A comparison of educational statistics and data mining approaches to identify characteristics that impact online learning. *J. Edu. Data Mining* (JEDM), 7(3): 117-150.

Dekker GW, Pechenizkiy M & Vleeshouwers JM 2009. Predicting students drop out: A case study. *Int. Working Group on Edu. Data Mining*.

Delavari N, Phon-Amnuaisuk S & Mohammad Reza Beikzadeh MR 2008. Data mining application in higher learning institutions. *Informatics in Edu. Int. J.*, 7: 31-54.

Galyardt A & Goldin I 2015. Move your lamp post: Recent data reflect learner knowledge better than older data. *J. Edu. Data Mining* (JEDM), 7(2): 83-108.

Garcia E, Romero C, Ventura S & De Castro C 2009. An architecture for making recommendations to courseware authors using association rule mining and collaborative filtering. *User Modelling and User-Adapted Interaction*, 19(1-2): 99-132.

Hao J, Shu Z & von Davier A 2015. Analysing process data from game/scenario-based tasks: An edit distance approach. *J. Edu. Data Mining* (JEDM), 7(1): 33-50.

Harley JM, Trevors GJ & Azevedo R 2013. Clustering and profiling students according to their interactions with an intelligent tutoring system fostering self-regulated learning. *J. Edu. Data Mining* (JEDM), 5(1): 104-146.

Hsia Tai-Chang, An-Jin Shie, & Li-Chen Chen 2008. Course planning of extension education to meet market demand by using data mining techniques an example of Chinkuotechnology University in Taiwan. *Expert Sys. with Applic.*, 34(1): 596-602.

Huang Chiao-Tzu, Lin Wen-Tsann, Wang Shen-Tsu & Wang Wen-Shan 2009. Planning of educational training courses by data mining: Using china motor corporation as an example. *Expert Sys. with Applic.*, 36(3): 7199-7209.

Kinnebrew JS, Loretz KM & Biswas G 2013. A contextualized, differential sequence mining method to derive students' learning behaviour patterns. *J. Edu. Data Mining* (JEDM), 5(1): 190-219.

Knowles JE 2015. Of needles and haystacks: Building an accurate state-wide dropout early warning system in Wisconsin. *J. Edu. Data Mining* (JEDM), 7(3): 18-67.

Lee Chun-Hsiung, Gwo-Guang Lee & Yungho Leu 2009. Application of automatically constructed concept map of learning to conceptual diagnosis of e-learning. *Expert Sys. with Applic.*, 36(2): 1675-1684.

Lykourentzou Ioanna, Ioannis Giannoukos, Vassilis Nikolopoulos, George Mpardis & Vassili Loumos 2009. Dropout prediction in e-learning courses through the combination of machine learning techniques. *Computers & Education*, 53(3): 950-965.

Lynch Collin, Kevin Ashley, Vincent Aleven & Niels Pinkwart 2006. Defining ill-defined domains; a literature survey. In *Proceedings of the Workshop on Intelligent Tutoring Systems for Ill-defined Domains 8th International Conference on Intelligent Tutoring Systems*, pp. 1-10.

Leah P, Macfadyen & Shane Dawson 2010. Mining LMS data to develop an early warning system for educators: A proof of concept. *Computers & Education*, 54(2): 588-599.

Mallavarapu A, Lyons L, Shelley T & Slattery B 2015. Developing computational methods to measure and track

learners' spatial reasoning in an open-ended simulation. *J. Edu. Data Mining* (JEDM), 7(2): 49-82.

Novak JD & Alberto J Canas 2008. The Theory Underlying Concept Maps and how to Construct and use them. http://cmap.ihmc.us/docs/theory-of-concept-maps.

O'Mahony MP & Barry Smyth 2007. A recommender system for online course enrolment: an initial study. In *Proceedings of the 2007 ACM Conference on Recommender Systems*, pp. 133-136.ACM.

Pena-Ayala Alejandro 2013. *Educational Data Mining: Applications and Trends*, volume 524. Springer.

Rallo R, Gisbert M & J. Salinas J 1999. Using data mining and social networks to analyse the structure and content of educative online communities. *Analysis*, 468(472): 473.

Reflay C & Chanier T 2003. *How social network analysis can help to measure cohesion in collaborative distance-learning*, pages 343-352. Designing for change in networked learning environments. Springer.

Reyes P & Tchounikine P 2005. Mining learning groups' activities in forum-type tools. In *Proceedings of the 2005 Conference on Computer Support for Collaborative Learning: Learning 2005: the Next 10 Years!*, pages 509-513. International Society of the Learning Sciences.

Romero C & Ventura S 2013. Data mining in education. *Wiley Interdisciplinary Reviews*: *Data Mining and Knowledge Discovery*, 3(1):12-27.

Romero C, Zafra A, Luna JM & Ventura S 2013. Association rule mining using genetic programming to provide feedback to instructors from multiple choice quiz data. *Expert Systems*, 30(2): 162-172.

Romero C & Ventura S 2010. Educational data mining: A review of the state of the art. Systems, man, and cybernetics, Part C: Applications and reviews. *IEEE Transactions on*, 40(6): 601-618, 2010.

Tang Changjie, Rynson WH Lau, Qing Li, Huabei Yin, Tong Li, & Danny Kilis 2000. Personalized courseware construction based on web data mining. In *Web Information Systems Engineering, 2000.Proceedings of the First International Conference on*, 2: 204-211. IEEE.

Vilardi Cesar, Javier Bravo Agapito, Leila ShilaShafti & Alvaro Ortigosa 2009. Recommendation in higher education using data mining techniques. Barnes T, Desmarais M, Romero C & Ventura S (eds.)

Wang CX Victor 2014. Handbook of research on education and technology in a changing society. IGI Global.

Waters Andrew, Christoph Studer & Richard Baraniuk 2014. Collaboration type identification in educational datasets. *J. Edu. Data Mining* (JEDM), 6(1): 28-52.

Zimmermann Judith, Kay H Brodersen, Hans R Heinimann & Joachim M Buhmann 2015. A model-based approach to predicting graduate-level performance using indicators of undergraduate-level performance. *J. Edu. Data Mining* (JEDM), 7(3): 151-17615.