

An Article on Education Data Mining and Their Application

Ankita*

677, Huda Sector - 01, Shahabad Markanda, District-Kurukshetra-136135, Haryana, India

Abstract – Data mining methods are used to derive valuable knowledge from raw data. The information gained is important and greatly impacts the decision-maker. Educational Data Mining (EDM) is a means of collecting useful information that could potentially affect an entity. Increasing the use of technology in education systems has contributed to the storing of vast amounts of student data, which makes it important to use EDM to enhance teaching and learning processes. Educational institutions are one of the most essential parts of our society and play a vital role in the growth and development of any country. Data Mining is an evolving methodology that can easily learn from historical data and use that information to predict future activity in areas of concern.

Keywords – Data Mining, Educational Data Mining (EDM), Applications in Education Data Mining

-----X-----

INTRODUCTION

Data mining is valuable at whatever point a framework is managing enormous data sets. In any instruction framework, understudy records for example enlistment subtleties, course qualification criteria, course intrigue and scholastic execution might be a significant thought to break down different patterns since every one of the frameworks are presently PC based data framework so data accessibility, alteration and updatation are a typical procedure now. Data warehousing might be taken as acceptable decision for keeping up the records of previous history. The data stockroom can be effectively created in any instruction organization with the adjustment of basic data alteration before stacking this for a data distribution center.

One of the essential objectives of any educational framework is to outfit students with the knowledge and abilities expected to change into fruitful vocations inside a predetermined period. How successfully worldwide educational frameworks meet this objective is a significant determinant of both monetary and social advancement. A few nations give free instruction to all residents from grade one through the college years. In this manner, an enormous number of students enter colleges consistently. For instance, King Khalid University (KKU) acknowledged around 23,000 students in 2013. It has gotten hard to give top notch instructing and direction to such countless students. Thus, numerous students neglect to finish their degrees inside the necessary time frames. EDM can give colleges an away from of explicit impediments to understudy learning. For instance, students can

bomb in cutting edge subjects since they didn't take in the fundamental data from the essential subjects. Utilizing data mining (DM) procedures to investigate understudy data can help distinguish potential explanations behind understudy disappointments.

Data mining gives numerous systems to data examination. The huge measure of data as of now in understudy databases surpasses the human capacity to dissect and separate the most valuable data without assistance from robotized examination strategies. Knowledge discovery (KD) is the procedure of nontrivial extraction of certain, obscure, and conceivably valuable data from a huge database. Data mining has been utilized in KD to find designs as for a clients needs. The example definition is an articulation in language that portrays a subset of data. A case of a KD design definition shows up in [1]. The expanding utilization of innovation in educational frameworks has made a lot of data accessible. EDM gives a lot of important data [2] and offers a more clear image of students and their learning forms. It utilizes DM systems to investigate educational data and fathom educational issues. Like other DM strategies extraction forms, EDM separates intriguing, interpretable, valuable, and novel data from educational data. In any case, EDM is explicitly planned for creating techniques that utilization extraordinary kinds of data in educational frameworks [3]. Such strategies are then used to upgrade knowledge about educational phenomena, students, and the settings where they learn [4].

DATA MINING

DM is an amazing artificial intelligence (AI) device, which can find valuable data by breaking down data from numerous edges or measurements, order that data, and outline the connections recognized in the database. Along these lines, this data helps settle on or improve choices. In DM arrangements, calculations can be utilized either autonomously or together to accomplish the ideal outcomes. A few calculations can investigate data; others separate a particular result dependent on that data. For instance, bunching calculations, which perceive designs, can assemble data into various n-gatherings. The data in each gathering are pretty much predictable, and the outcomes can help make a superior choice model. Various calculations, when applied to one arrangement, can perform separate assignments. For instance, by utilizing a regression tree strategy, they can get money related conjectures or affiliation rules to play out a market examination. A lot of data in databases today surpasses the human capacity to investigate and extricate the most valuable data without assistance from robotized examination systems. Knowledge discovery is the procedure of nontrivial extraction of verifiable, obscure, and possibly helpful data from an enormous database. Data mining utilized in KD has found examples concerning a client's needs. The example definition is an articulation in the language that portrays a subset of data; a model is appeared in [1]. The exact discovery of examples through DM is impacted by a few variables, for example, test size, data uprightness, and backing from space knowledge, all of which influence the level of assurance expected to distinguish designs. Commonly, DM reveals various examples in a database; in any case, just some of them are intriguing. Helpful knowledge establishes the examples important to the client. It is significant for clients to consider the level of trust in a given example while assessing its legitimacy.

EDUCATIONAL DATA MINING

Educational data mining is an evolving field associated with creating methods for analyzing the particular types of data that come from educational settings and using these tools to better understand students and the environments they study in [3]. Like data mining techniques, EDM, when used specifically, covers for (and allows use of) a multi-level structure and ignores objective educational data [3].

TECHNIQUES OF EDUCATION DATA MINING

The challenges faced by the production of Big Data technology are addressed by the use of different techniques. The most common strategies used in the field of educational data mining are described following.

- **Regression** – Regression was used to estimate the significance of a variable by measuring the association between variables utilizing analysis.
- **Nearest Neighbor** – In this method, the values are anticipated on the basis of the predicted ideals of the records which are closest to the record of the desires to be predicted.
- **Clustering** – Clustering means combining documents which are identical by defining the difference among them in an n-dimensional space because n is the number of variables.
- **Classification** – Classification is the definition of the category / class at which the attribute corresponds on the basis of previously classified values.

APPLICATIONS IN EDUCATION DATA MINING

Three applications for educational data mining are discussed here, with particular attention paid to them.

- **Predicting Student Performance**

Lin applied order and regression trees to anticipate what kinds of students would drop out from school, and afterward profit to class later for. The models had the option to give transient precision to foreseeing which kinds of students would profit by understudy maintenance programs. Chacon and Spicer et al. built up a framework dependent on data mining enables the foundation to distinguish and react to students in danger. Their work is profoundly illustrative of the control, since it follows with a severe data mining process, which is quantitative. The examination by Yeats, Reddy and Wheeler found that students who go to composing focuses will in general work superbly in their classes. Yu and DiGangi, et al. found that east coast students in USA will in general keep enlisted longer than their west coast partners do.

- **Course Management System**

EDM is frequently utilized in course the board frameworks, as Moodle, which contains utilization data that incorporates various exercises. García, Romero, Ventura, and de Castro built up a streamlined data mining toolbox that works inside the course the board framework and enables students and their learning clients to get data mining data for their courses. This exploration and application commitments will permit non-specialized workforce to take part in educational data mining exercises. Rather than customary static course designs, data mining can be applied

to redo learning exercises and adjust the pace for students to finish courses. It will make huge and ideal learning encounters for every understudy. Additionally, Blikstein found various kinds of programming practices in an online course.

- Planning and Scheduling

Inquires about on portable learning conditions as of late recommend that data mining can be applied to help give customized substance to various versatile clients, notwithstanding the contrasts between cell phones and traditional PCs. EDM applications will permit non-specialized clients take part in data mining devices and exercises making handling progressively available for all EDM clients. There are a few models, including factual and representation devices, investigating informal organizations and related impact on learning results.

CHALLENGES

As the related innovation created, expenses and difficulties related with executing EDM applications, such as putting away logged data and overseeing data frameworks [5]. Besides, picking which data to mine and break down may likewise be a test. Furthermore, singular protection is a proceeded with worry for the use of educational data mining devices. With free, available apparatuses in the market, students and students might be in danger giving data to the learning framework. Securing singular protection ought to be considered for the long haul improvement of EDM. In addition, it's hazy what data presentations, visualizations, and visual investigation are generally educational and support effective decision making for various partners.

LITERATURE REVIEW

In this report Author researches the impression of Knowledge Management inside Higher Education, and presents the idea of scholastics and colleges. It centers around two parts of the contextual investigation – the attributes of colleges and scholastics that support the usage of KM, and the impression of Knowledge Management and its difficulties for execution inside the higher education division [6].

In numerous areas Data Mining strategies are significant and from that Data Mining methods are utilized to improve the capacity of higher education establishment. On the off chance that Data Mining Techniques like clustering, decision tree and affiliation are applied to higher education process, it would improve in students life cycle the board, their exhibition, just as in determination of courses and so on. In this report creator had given brief presentation on data mining procedures like group investigation, decision tree, factor examination, regression examination [7].

In this report Author had utilized Classification method to assess understudy execution. There are numerous methodologies that are utilized for arrangement and from that creator had utilized decision tree strategy. Data resembles understudy's participation in homeroom, course, class test, and task characteristic of understudy were gathered from the understudy's administration framework, to foresee the understudy execution toward the finish of the semester [8].

Creator had utilized J48 Algorithm for foresee understudy's presentation and for this assignment characterization strategy is utilized in this report. J48 calculation is utilized to groups the data as decision tree and utilizing this decision tree we can without much of a stretch recognize the feeble understudy [9].

In this report creator review positive relationship understudy's presentation and the idea of the college. Creator had select understudy utilizing group inspecting strategy. This procedure is utilized for gathering the students [10].

In this report Author talked about which data mining procedures can be applied in the field of education and to distinguish which data mining system is appropriate for what sort of use as a theoretical model. For Example Classification system is smarter to anticipate understudy's presentation [11].

In this paper creator had conquered issue in ID3 Algorithm and created weighted ID3 Algorithm. ID3 calculation is one of the acclaimed calculation to produce decision trees. ID3 calculation has a deficiency that it arranged properties with numerous qualities. Thus, Author beats this issue by utilizing gain proportion. In this procedure creator offers weight to each trait at each decision making point [12].

Creator led study on the understudy by utilizing order strategy on classification, language and foundation capability and it was discovered that whether new comer students will entertainer or not. Creator directed a relative report to discover private coaching rate in various nations. Creator utilized decision tree model to foresee the last grade of students [13].

Higher education teachers are regularly intrigued by forecast of understudy's outcome before their tests. During classes they attempt to anticipate execution of students. For this issue creator portrayed that a few data mining systems like neural systems will be unable to achieve the learning task as little datasets and it can't give enough data to fill the holes between too little examples [14].

In this report creator examined on quality education. It doesn't mean generation of elevated level knowledge yet significance of value education is that education is created to students in efficient way so they can learn with no sort of issue [15].

Dough puncher spoke to the history and current patterns in the field of Educational Data Mining (EDM). We consider the methodological profile of research in the early long periods of EDM, contrasted with in 2008 and 2009, and talk about patterns and moves in the exploration led by this network. Specifically, we talk about the expanded accentuation on forecast, the development of work utilizing existing models to make logical disclosures ("discovery with models"), and the decrease in the recurrence of relationship mining inside the EDM people group. We examine two different ways that analysts have endeavored to arrange the assorted variety of research in educational data mining examination, and audit the kinds of research issues that these techniques have been utilized to address. The most referred to papers in EDM somewhere in the range of 1995 and 2005 are recorded, and their effect on the EDM people group (and past the EDM people group) is talked about [16].

Ritu Gautam present Data mining is utilized to discover new and valuable data from huge measure of data. Strategies of data mining are helpful in different application regions like extortion discovery, organizations, banking and broadcast communications. The significant use of Data Mining Technique is educational data mining so as to extricate valuable data from educational data. In Education Quality Assurance has constrained scholastics for continually investigating various ways for making enhancements in educational procedures. It has prompted expanding enthusiasm for educational data mining. Developing enthusiasm for data and investigation in education, instructing, and learning raises the need for expanded, top notch look into Data Mining is a procedure used to discover perhaps new data from enormous measure of data. Educational data mining is a rising pattern, worried about creating techniques for investigating the tremendous data that originate from the educational framework. . The target of this examination is to present Educational Data mining, by depicting a bit by bit process utilizing an assortment of procedures. In this paper a survey is directed on bit by bit procedures and application regions [17].

Pratiyush Guleria said that Data mining (DM) is where data is broke down and abridged into valuable data. To put it plainly, data mining is procedure of getting designs from huge databases. DM examinations huge dataset to separate shrouded examples, for example, comparable gatherings of data records utilizing clustering procedure. Knowledge Discovery in Databases is the way toward discovering knowledge in monstrous measure of data where data mining is the center of this procedure. Data mining can be utilized to mine

reasonable significant examples from enormous databases and these examples may then be changed over into knowledge. Data mining is the way toward removing the data and examples determined by the KDD procedure which helps in vital decision-making. Data mining works with data stockroom and the entire procedure is separated energetically plan to be performed on data: Selection, change, mining and results understanding. In this paper, we have explored Knowledge Discovery viewpoint in Data Mining and solidified various zones of data mining, its systems and techniques in it [18].

ASHISH DUTT by and by, educational establishments accumulate and store colossal volumes of data, for example, understudy enrolment and participation records, just as their assessment results. Mining such data yields invigorating data that serves its handlers well. Quick development in educational data focuses to the way that refining huge measures of data requires an increasingly advanced arrangement of calculations. This issue prompted the rise of the field of educational data mining (EDM). Conventional data mining calculations can't be straightforwardly applied to educational issues, as they may have a particular target and capacity. This suggests a preprocessing calculation must be implemented sole then some particular data mining strategies can be applied to the issues. One such preprocessing calculation in EDM is clustering. Numerous examinations on EDM have concentrated on the use of different data mining calculations to educational qualities. Along these lines, this paper gives more than three decades in length (1983–2016) methodical writing survey on clustering calculation and its materialness and ease of use with regards to EDM. Future bits of knowledge are illustrated dependent on the writing checked on, and roads for additional exploration are recognized [19].

Cristóbal Romero spoke to Educational data mining (EDM) is a rising interdisciplinary research zone that manages the advancement of techniques to investigate data beginning in an educational setting. EDM utilizes computational ways to deal with break down educational data so as to examine educational inquiries. This paper overviews the most applicable examinations did in this field to date. To begin with, it presents EDM and portrays the various gatherings of client, kinds of educational conditions, and the data they give. It at that point proceeds to list the most commonplace/basic undertakings in the educational condition that have been settled through data-mining procedures, lastly, probably the most encouraging future lines of research are discussed[20].

Suhrman present Management of higher education must keep on assessing on a progressing premise so as to improve the nature of foundations. This will have the option to do the essential assessment of

different data, data, and knowledge of both interior and outer foundations. They intend to utilize all the more effectively the gathered data, create devices so that to gather and direct administration data, so as to support administrative decision making. The gathered data could be used to assess quality, perform examinations and findings, assess trustworthiness to the benchmarks and practices of educational programs and schedules, and propose options in decision forms. Data mining to support decision making are appropriate techniques to give decision support in the education conditions, by creating and showing significant data and knowledge towards quality improvement of education forms. In educational space, this data is exceptionally valuable since it very well may be utilized as a base for exploring and improving the current educational principles and administrations. In this paper, a survey on data mining for scholarly decision support in education field is exhibited. The subtleties of this paper will survey on ongoing data mining in educational field and layouts future explores in educational data mining [21].

Siti Khadijah Mohamad clarifies the Data Mining is exceptionally valuable in the field of education particularly while examining conduct in internet learning condition. This is because of the capability of data mining in dissecting and revealing the concealed data of the data itself which is hard and very tedious if to be done physically. The reason for this audit is to investigate how the data mining was handled by past researchers and the most recent patterns on data mining in educational research. A few confinements of existing exploration are talked about and a few headings for future research are proposed [22].

CONCLUSION

Educational associations are one of the significant pieces of our general public and assuming an imperative job for development and advancement of any country. Data Mining is a developing method with the assistance of this one can effectively learn with chronicled data and utilize that knowledge for predicting future conduct of concern zones. Development of current education framework is without a doubt upgraded if data mining has been received as a futuristic strategic management instrument. The data mining of educational data (EDM) is making advancement techniques for the extraction of fascinating, interpretable, valuable, and novel data, which can prompt better comprehension of students and the settings in which they learn. EDM can be utilized in a wide range of regions including recognizing in danger students, distinguishing needs for the adapting needs of various gatherings of students, expanding graduation rates, effectively surveying institutional execution, boosting grounds assets, and optimizing subject curriculum reestablishment.

REFERENCES

- [1]. S.T. Wu (2007). "Knowledge discovery using pattern taxonomy model in text mining,".
- [2]. J. Mostow and J. Beck (2006). "Some useful tactics to modify, map and mine data from intelligent tutors," *Natural Language Engineering*, Vol. 12, No. 02, pp. 195–208.
- [3]. S. K. Mohamad and Z. Tasir (2013). "Educational data mining: A review," *Procedia-Social and Behavioral Sciences*, vol. 97, pp. 320–324.
- [4]. R. Baker et. al. (2010). "Data mining for education," *International encyclopedia of education*, vol. 7, pp. 112–118.
- [5]. Baker, R.S.J.D., Yacef, K. (2009) *The State of Educational Data Mining in 2009: A Review and Future Visions*. *Journal of Educational Data Mining*, 1 (1), pp. 3-17.
- [6]. Dr. R. Shukla: "KNOWLEDGE MANAGEMENT IN HIGHER EDUCATION", *International Journal of Reviews, Surveys and Research (IJRSR)*.
- [7]. M. Goyal & R. Vohra (2012). "Applications of Data Mining in Higher Education", *International Journal of Computer Science Issues*, Vol.9, Issue 2, No1, March 2012.
- [8]. B. Baradwaj, S. Pal (2011). "Mining Educational Data to Analyze Students' Performance" (*IJACSA*) *International Journal Of Advanced Computer Science and Applications*, Vol. 2, No. 6.
- [9]. M. Bhullar, A. Kaur (2012). "Use of Data Mining in Education Sector", *Proceedings of the World Congress on Engineering and Computer Science*, Vol. I, Oct 2012
- [10]. A. Laoufi, S. Mouhim, E. Megder, C. Cherkaoui and D. Mammass (2011). "Using Knowledge Management in Higher Education: Research Challenges and Opportunities" *Journal of Theoretical and Applied Information Technology*, Vol.31 No.2, pp. 100-108, Sep-2011
- [11]. S. Bhanap & R. Kulkarni (2013). "Student - Teacher Model for Higher Education System", ISSN: 2279-0535 Volume: Volume –II Issue: Issue-III.
- [12]. S. Dhanda, S.D, R.L, (2013). *Predicting Students' Performance using Modified ID3 Algorithm*, *International Journal of*

Engineering and Technology, ISSN:0975-4024, Vol. 5 No. 3 Jun-Jul 2013

Science Procedia - Social and Behavioral Sciences.

- [13]. U. Pandey & S. Pal (2011). "A Data Mining View on Class Room Teaching Language", ISSN:1694-0814, International Journal of Computer Science Issues, Vol. 8, Issue.2, Mar-2011
- [14]. Dr. Y. Cader (2007). "Knowledge Management and Knowledge-based Marketing "Journal of Business Chemistry, vol.4 , Issue 2 ,pp. 46-58.
- [15]. U.Pandey, B. Bhardwaj and S.Pal: "Data Mining as a Torch Bearer in Education Sector", Technical Journal of LBSIMDS.
- [16]. Baker, R. S., & Yacef, K. (2009). The State of Educational Data Mining in 2009: A Review and Future Visions. JEDM | Journal of Educational Data Mining, 1(1), 3-17. <https://doi.org/10.5281/zenodo.3554657>
- [17]. Ritu Gautam¹ , Deepika Pahuja (2012). "A Review on Educational Data Mining", International Journal of Science and Research (IJSR) ISSN (Online): 2319-7064 Impact Factor: 3.358 Volume 3 Issue 11, November 2014
- [18]. Pratiyush Guleria and Manu Sood (2014). "DATA MINING IN EDUCATION: A REVIEW ON THE KNOWLEDGE DISCOVERY PERSPECTIVE", International Journal of Data Mining & Knowledge Management Process (IJDMP) Vol.4, No.5, September 2014
- [19]. ASHISH DUTT¹ , MAIZATUL AKMAR ISMAIL¹, AND TUTUT HERAWAN (2017). "A Systematic Review on Educational Data Mining", IEEE accepted December 24, 2016, date of publication January 17, 2017, date of current version August 29, 2017.
- [20]. Cristóbal Romero; Sebastián Ventura (2010). "Educational Data Mining: A Review of the State of the Art", IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews) (Volume: 40 , Issue: 6 , Nov. 2010)
- [21]. Suhirman Suhirman, Tutut Herawan, Haruna Chiroma, Jasni Mohamad Zain (2017). "Data Mining for Education Decision Support: A Review", International Journal of Emerging Technologies in Learning (IJET).
- [22]. Siti Khadijah Mohamada , Zaidatun Tasir: "Educational data mining: A review", The 9th International Conference on Cognitive

Corresponding Author

Ankita*

677, Huda Sector - 01, Shahabad Markanda, District-Kurukshetra-136135, Haryana, India