

Role of Data Mining in Education Sector

Sunil MP ¹, Gernel Lumacad ²

¹ Assistant Professor Department of Electronics and Communication Engineering, Faculty of Engineering and Technology, JAIN (Deemed-to-be University), India

² St. Rita's College of Balingasag, Philippines

Corresponding Author E-Mail: ¹mp.sunil@jainuniversity.ac.in

Abstract

Data mining is a method used for categorizing and predicting the performance of a student and a teacher as well. It helps both student and teacher in developing the teaching method and the entire system. Every student who uses this method get a huge favor from this. This helps students to choose the right career option. Currently, the tendency of providing immense importance to data mining techniques in the educational sector is increasing as these procedures have a huge necessity in bringing efficiency in both learning and teaching procedures. Each and every person is gathering a large amount of data every day, in case these data are not further examined; only the large amount of data will remain. With the latest systems and technologies, people can utilise these data and examine those and benefit from it. The best technique for this issue is the data mining process. Data mining is a method of bringing out the useful and disclosed data and information from big data sets. Educational data mining is a process from which teachers and students get a lot of help. Teachers are able to observe every student's performance. On the other hand, students can choose a perfect and accurate career option by this process. This process utilises several techniques and methods such as statistics, machine learning, data analysis and data mining. Educational data mining is a method of raw information converting from a huge educational database to meaningful and effective information.

Keywords

Data mining, education, students.

INTRODUCTION

Educational data mining is utilized for predicting and categorizing performance and dropouts of the students and the performance of teachers as well. It may assist teachers to trace academic advancement to develop the method of teaching. On the other hand, it assists students in education management and course choosing to be more systematic and successful [1]. This is a method of taking out the pattern and some other functional information from a big data set. Data mining is an action of searching heavy capacity of data to discover patterns, find trends and increase perception into the usage of the data. Then the data miners may use those searching to guess the result and make decisions. This has the ability to convert unprocessed data into high level information that always assists businesses increase by taking the best decisions.

There are some different types of data mining such as social media mining, video and audio mining, text mining and pictorial data mining. The function of data mining in education is to develop learning results by mining and examining the data. Teachers can understand the interest of every student and can do a better job on developing the experience of learning for students by using data collected from the behavior investigation and several evaluations. Generally, the basic function of data mining is prediction function which is the method of searching patterns from data that uses many variables to discover parallel variables of undecided value [2]. Data analytics can help students in finding correct and genuine answers to difficult questions. It may save money and it is super easy and quick. This is extremely approachable, understandable, easily reachable and immensely convenient. This assists a student to adapt.

MATERIALS AND METHODS

This study has been conducted on the basis of particular research design, approach, data collection procedure, and research method. In this study, cross-sectional research design has been chosen as this specific design helps to portray different aspects of the study in an explicit way. Inductive research approach has been followed in present study; reason behind it is that certain interpretations related to the study can be developed based on this approach. Apart from this an appropriate conclusion is derived by the researcher with the help of an inductive approach. Secondary data has been collected in the study to develop knowledge about the importance of data mining in educational institutions.

In this particular case, the researcher has gathered secondary data by following a qualitative method. This specific method has a huge necessity in helping researchers to find the reason rather than the fact [3]. Secondary data in the present case has been collected from various peer-reviewed journals and articles. All journals and articles from which information and data have been gathered are published after the year 2019. Articles and peer-reviewed journals that have been published before 2019 have been excluded from this study. Despite this, any kind of irrelevant information that has no connection with this study is excluded.

RESULTS

Data mining has started global possibilities for several businesses around the world. This calculation statistical field contrasts billions of segregated pieces of data and utilized by organizations to predict and detect the behavior of customers.

The objective of data mining is to give rise to new market opportunities. This process transforms data into terrific good knowledge. Data mining in education can be utilized for a good understanding of students and the conditions of learning of the students [4]. It also helps in developing teaching support and the educational system as well. Educational data mining sector is utilizing a huge amount of information to authenticate the investigated findings. This also assists guesses on student dropout, knowledge and the inspirational state becomes more faults-less with the additional data. The benefits of data mining in educational places are countless [5]. There are several and numerous studies and research papers concerning the applications and usage of data mining process in education.

One of the most regular usage of data mining is developing the procedure of study, upgrading execution of course, helping students in find the correct course for own self, creating the best profile of a student, finding the main issue regarding the dropouts, aim and motto of every students, improving not only the study but also giving attention to extracurricular activities, forecasting the performance of a student and plays an extremely important role in helping a student as a support for taking a good decision of student enrolment. The benefits of data mining involve helping organizations to collect some information that is trustworthy. Compared to several data applications, data mining applications are quite well organized and low cost, time saving, flexible and extremely user friendly. This application helps organizations produce more beneficial production and functional adjustments. Data mining system in education utilizes both inheritance and new systems.

On the other hand, data mining techniques have some challenges and issues as well. Some frequently faced data mining challenges are - social and security challenges. Particular and private data about every person and sensitive information is collected for buyer's profiles, buyer standard of behavior appreciation - illegal access to data and the private plan of information transforming into a big problem.

Incomplete and obstreperous data

Data mining is a method to acquire data from extreme big sizes of information. This living truth of data is quite incomplete, noisy and miscellaneous. Information in big amounts frequently will be incorrect and untrustworthy [6]. These problems, errors and blunders happen because of humans only in the devices that estimate the data.

Compound data

Real data is diversified, and it can be data from media, adding common language text, specific data, material data, video and audio data, image data, time series data and many more. It is extremely difficult to work with these different types of data to focus on the essential information [7]. New systems and equipment will surely need to be made to detach information that is important and more private.

Disseminated data

Real and original data is generally reserved at numerous stages in spreaded processing state. Data mining can be on separate systems, database and internet [8]. It is extremely difficult to convey all the information to a merged data archive supremely because of organizational and technical reasons.

Performance

The demonstration of the data mining substructure normally depends on the capacity of algorithms and techniques. On the little chance that the algorithms and technique plan are not enough at that time, which will inspire the demonstration of the data mining estimate unfortunately.

Development of mining design

Matters of the struggling of data mining approaches, the huge size of the database and the whole flow of data influence the creation and distribution of aligned data mining development.

Incorporable of framework knowledge

In the matter that framework knowledge may be combined, more actual, exact and well-founded data mining positioning might be established. Forecasting tasks might create more exact predictions while illustrative tasks might come up with more fruitful and effective searching.

Data security and privacy

Data mining commonly gives rise to importance governance, data security, privacy, and data challenges related issues. For example, when a dealer examines the details of the purchase, it reveals every detailed information about purchasing tendencies and consumer's exact choice without their consent.

In every educational institution, it is necessary to predict the performance of students; right prediction plays an important role in classifying performance levels of students. Data mining has a great necessity in aiding educational sector to get an understanding of learners' performance [9]. In case teachers of educational institutions can be capable of comprehending progression of students, it becomes easy for these individuals to ameliorate their overall teaching process. A number of students are enrolled into several programs and courses by private and public universities and colleges each year. These institutions gather essential information of students during admissions and this data is stored in computers. Universities and colleges are helped by data mining in the field of keeping a track of students' behavioural patterns and teaching procedures of teachers. Data mining consists of a huge importance in aiding educational sector to recognise different choices of students [10]. Students' requirements towards course selection of specialisation can be identified by universities based on data mining techniques.

In many cases, different pattern trends can be noticed in students which are recognised with the help of data mining. Mining technique of data is a necessity in making universities

efficient in the field of assuming final results, grades, and knowledge of students. In recent days, competition between different educational institutions is increasing and each institution is endeavouring to provide the best learning facilities to students. In this case, the role of data mining in helping the educational sector to bring effectiveness in both learning and teaching procedures of learners and educators respectively cannot be denied. Profiles of students can be predicted through following data mining procedure [11]. Psychographic, geographic, and demographic traits of students are comprehended by educational institutions by including data mining techniques in all sorts of work.

Data mining helps teachers to keep the record of attendance of learners; apart from this feedback of students can be understood by teachers through data mining procedures. Students' feedback is necessary for teachers in the field of getting an idea about issues that are faced by learners [12]. Subsequently, educators can modify their teaching methods and bring improvement in the overall learning procedure of students. However, institutions are capable of observing the attitude of instructors with the help of several procedures related to data mining. Understanding of students' requirements based on data mining provides a huge opportunity to educators in the field of establishing a robust relationship with students [13]. Good relationships motivate learners to strengthen their learning procedures and achieve outstanding results. Hence, it is clear that data mining not only helps students and teachers to ameliorate learning and teaching methods but also it plays an essential part in intensifying the reputation of educational institutions.

Data usage is important because this system enables more fruitful assessment of programs, interference to ease students resources and the success of the entire district. Integrated and state systems are also taking help from data to make codification which concentrates on the achievements of every student, progress and connecting the particular needs of the district and state [14]. An amalgamation of several types of information or data is more fruitful in causing strong proof to evaluate the performance of the school and develop practice.

Types of data

Data is generally the information, specifically the numbers and facts that are gathered to be considered, inspected, and used to help the process of making decisions, or information in electrical way, which can be kept and utilised by a computer. Data mining has some different types such as text mining, social media mining, video and audio mining, pictorial data mining and web mining. In recent days, universities work in a hugely complex and competitive environment [15]. With highly frequent technology improvement and low cost IT tools, the huge amount of data kept in educational databases grows frequently, in case this data is not examined further, it endures a high amount of data only. Data mining equipment, techniques and methods permit users to examine data and discover information and concealed patterns. Data mining is utilised to find

relationships and patterns in data to develop the process of making decisions. This is an integrative place which gathers techniques from demography, fake intelligence, audio-visual networks, systems of database, process of learning, pattern identification, data measurements, knowledge accession and the theory of information.

The quantity of data gathered and kept in several educational institutions increased huge and educational data examination could not be functioned dynamically. Educational data mining is a proportionately new regulation which appeared from the data mining application on educational data [16]. Educational data mining keeps increasing from several research locations such as machine learning and data mining, pattern identification, neurometrics and other places of statistics, and fake intelligence, data contrivance and mathematical modelling. The main motto of educational data mining is to develop the process of education and to clarify the strategies of education for making decisions better [17]. There are several descriptions of educational data mining. This is an incorporative research place that utilises several techniques and methods from data analysis, data mining, machine learning and statistics, to examine data gathered while learning and teaching in order to find earlier disclosed data, patterns in huge data depositories and relationships.

The educational data mining process has four most important stages. The first phase is known as "problem definition" in which a particular problem is interpreted into the problem of data mining. In this stage, the project objectives and aims are put together, and also the questions of research are also put together. The most time devouring stage is the second stage, data collecting and preparation phase. It may take most of the examination time. A major issue in data mining is data quality. Source data must be recognised, formatted and cleaned in pre-defined format in this stage. Next, there is an evaluating and modelling stage in which the parameters are settled to proper worth and several modelling methods are chosen and applied. The last stage or phase is the "deployment phase" in which the final outcomes of data mining are arranged and introduced by reports and graphs [18]. It is extremely important to indicate that the data mining method is a repeated method that means the method never stops while a specific solution is placed. This may be a new insert for the latest data mining method.

Educational data mining is a new age research place that is being hugely known due to its prospective. Educational data mining data might be used to help teachers or any other instructors, to develop modules and curriculums. Apart from this to know the behaviour of students, to develop the process of teaching and to develop the process of e-learning, to find out the actual reasons behind drop-outs which is excessively important, to support the procedure of making decisions [19]. Educational data mining research is divided in two key categories; first one regards the examination of learning the proper behaviour and assigns which work on a successful study, and the second one of the category works on searching

an anticipating model for making better performance of students.

Educator, academic, details of students from childhood to human resources, demographic, these are known as educational data. Educational data are assembled from several sources in different formats, in spite of the fact that the data type and the person that can access it, varies. Applications and advantages of data mining in education are countless [20]. The most general use of educational data mining systems is developing the procedure of study, upgrading completion of every course, helping learners to find the perfect course, making the appropriate profile of students, finding out the actual reason behind dropouts, target of students, curriculum improvement and many more.

On the other hand, data mining equipment is extremely complicated and needs proper training and lessons to use. Data logic is an extremely complex procedure and rapidly needs users with appropriate training to use the equipment. Data mining processes are not an error free and perfect process all the time as this system does not give correct, accurate and faultless information every time. One of the worst disadvantages of data mining are privacy and data issues [21]. Generally, organisations only share personal information while providing a service.

In recent days, people are getting tensed that the personal data is going to be sold to third parties even without their concern. Some users will not feel good and comfortable to know that the government of the country can always trace some specific information about the users and know everything about how a user is using the device. Data mining is one of the strongest instruments in the market; on the other hand it has its own flaws. One major fault is, data mining always needs databases that are large to be more successful. It usually does not accept small databases. Data mining does not provide genuine and correct information every time [22]. There are several methods to examine data and few of them are more correct than the others. Another major problem is when there is some lost data in a database that requires to be considered to develop a complete investigation. Data mining systems can be immensely high cost procedures. It is obvious that schools, colleges or universities have to hire extra technology professionals and workers to make sure that the data mining is done accurately. Many businesses need to put money into developed and advanced data mining software that can be extremely expensive.

DISCUSSION

Educational data mining can be described as the process for searching the particular types of information or data which come from the system of education and implements those methods to understand the system and the students as well in a better way. Educational data mining is a method of converting raw information or data acquired from educational systems into functional data which can be utilised to create data related decisions. The improvement of data mining and logical in the field of education was

knowingly late as compared to any other fields. Though, it is quite challenging for data that are related to the educational field by the web because of the data mining system's particular functions on data. At the time when several types of data have consecutive characteristics, the dispensation of information that is related to education over time has been exceptionally ascribed. Educational data mining methods have been effectual and make a scale of occurrence concerning learner absorbing on the online area, and fulfilling better correctness continuously. There are notable features which need to be surveyed to explain the extreme advancement for educational data that is improving acceptance which is not all the major data kept in one data brook. Educational research has consequences in several new educational increments. Innovations that are computer based have transformed extremely that everyone knows. Data mining utilises its tools to search the previously disclosed relationships and patterns in a large data set. These tools can include analytical models, numerical algorithms and techniques that can be learned by machines. These methods are capable of finding the information within the data which reports and queries cannot productively disclose.

Data mining system makes sure that the information or data are properly secured and cannot get hacked by hackers easily. Proper security is an extremely important factor in every place. Every system or technology gets considered as a good one by seeing the security system first and a system such as data mining where several data or information of many users has been kept has to be properly secured. This technology helps to work with the events that happened in schools and calendars as well. Data mining process assists in automating the generation of timetables, this is an important part as it will be extremely problematic if one has to change the timetable every time or every day. This process permits a well organised management of every assignment. Data mining process guides one to observe the performance of students. Observing the performance of every student manually is quite difficult for a person. When the observation is done by a system, it gets easier. This technology records each and every data on exams, classes, library and demographic data. This type of data management system gives important inputs from which school management and teachers can get help to observe a student's everyday performance. Schools, colleges and universities have a huge number of resources at the ejection. To make sure resources such as classrooms, libraries, laboratories and every workspace are kept in use, a fruitful database management process is immensely important. This system keeps updated data and information on every resource that is available and assists the school authority to make sure the allotment is accurate. System workers need to frequently back up the data with normal and regular file processing systems to make sure that the data are accessible and safe. It is a quiet time taking process as every educational place has lots of big files in the system. In case the organisations have a good database management system such as a data mining system, workers

will not have to do the same job or back up every data every day.

CONCLUSION

Educational data mining method is a new age regulation with a big future for each and every candidate of the educational process. Data mining methods were improved to find the disclosed knowledge and identify patterns from data on its own. Educational data mining might be utilised for distinguishing and forecasting the performance of each and every student, teachers and the dropouts as well. This system can help teachers to trace the progress of academics to develop the method of teaching. It can also assist learners to choose the perfect and correct career and educational management as well to be more successful and efficient. Educational data mining can be utilised to maintain, retain and attract the learners to accomplish the revenue of university.

Inspecting learners' data is important for detecting, finding and knowing about the perfect and genuine instructional practices that are fruitful. It is obvious that people always want new methods and love to work with new processes, data mining systems are one of them. Data mining system is really important in education because it helps students and the teachers to find the data in new ways or find patterns the users even did not know existed. This process has several plus points such as decreasing the chances of fraud and making businesses more successful. On the other hand, it is quite obvious that it will have some issues like any other technical process would have such as inaccurate and unreal information or data and false inputs.

REFERENCE

- [1] Hasan, Raza, et al. "Predicting student performance in higher educational institutions using video learning analytics and data mining techniques." *Applied Sciences* 10.11 (2020): 3894.
- [2] Sathishkumar, V. E., Jangwoo Park, and Yongyun Cho. "Using data mining techniques for bike sharing demand prediction in metropolitan city." *Computer Communications* 153 (2020): 353-366.
- [3] Yang, Jin, et al. "Brief introduction of medical database and data mining technology in big data era." *Journal of Evidence-Based Medicine* 13.1 (2020): 57-69.
- [4] Romero, Cristobal, and Sebastian Ventura. "Educational data mining and learning analytics: An updated survey." *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 10.3 (2020): e1355.
- [5] Jones, Kyle ML, Alan Rubel, and Ellen LeClere. "A matter of trust: Higher education institutions as information fiduciaries in an age of educational data mining and learning analytics." *Journal of the Association for Information Science and Technology* 71.10 (2020): 1227-1241.
- [6] Janssen, Marijn, et al. "Data governance: Organizing data for trustworthy Artificial Intelligence." *Government Information Quarterly* 37.3 (2020): 101493.
- [7] Kiger, Michelle E., and Lara Varpio. "Thematic analysis of qualitative data: AMEE Guide No. 131." *Medical teacher* 42.8 (2020): 846-854.
- [8] Guo, Yuan, et al. "The internet of things-based decision support system for information processing in intelligent manufacturing using data mining technology." *Mechanical Systems and Signal Processing* 142 (2020): 106630.
- [9] Shin, Dongjo, and Jaekwoun Shim. "A systematic review on data mining for mathematics and science education." *International Journal of Science and Mathematics Education* 19.4 (2021): 639-659.
- [10] Xiao, Wen, Ping Ji, and Juan Hu. "A survey on educational data mining methods used for predicting students' performance." *Engineering Reports* 4.5 (2022): e12482.
- [11] Tomasevic, Nikola, Nikola Gvozdenovic, and Sanja Vranes. "An overview and comparison of supervised data mining techniques for student exam performance prediction." *Computers & education* 143 (2020): 103676.
- [12] Muñoz, Jorge Leoncio Rivera, et al. "Systematic Review of Adaptive Learning Technology for Learning in Higher Education." *Eurasian Journal of Educational Research* 98.98 (2022): 221-233.
- [13] Khan, Anupam, and Soumya K. Ghosh. "Student performance analysis and prediction in classroom learning: A review of educational data mining studies." *Education and information technologies* 26.1 (2021): 205-240.
- [14] Barksdale, Christopher, Michelle L. Peters, and Antonio Corrales. "Middle school students' perceptions of classroom climate and its relationship to achievement." *Educational Studies* 47.1 (2021): 84-107.
- [15] Hamadamin, Halbast Hussein, and Tarik Atan. "The impact of strategic human resource management practices on competitive advantage sustainability: The mediation of human capital development and employee commitment." *Sustainability* 11.20 (2019): 5782.
- [16] Ashraf, Mudasir, Majid Zaman, and Muheet Ahmed. "An intelligent prediction system for educational data mining based on ensemble and filtering approaches." *Procedia Computer Science* 167 (2020): 1471-1483.
- [17] Abu Saa, Amjed, Mostafa Al-Emran, and Khaled Shaalan. "Factors affecting students' performance in higher education: a systematic review of predictive data mining techniques." *Technology, Knowledge and Learning* 24.4 (2019): 567-598.
- [18] Irawan, Yuda. "Implementation Of Data Mining For Determining Majors Using K-Means Algorithm In Students Of SMA Negeri 1 Pangkalan Kerinci." *Journal of Applied Engineering and Technological Science (JAETS)* 1.1 (2019): 17-29.
- [19] Marlina, Evi, Bambang Tjahjadi, and Sri Ningsih. "Factors affecting student performance in e-learning: A case study of higher educational institutions in Indonesia." *The Journal of Asian Finance, Economics and Business* 8.4 (2021): 993-1001.
- [20] Jones, Kyle ML, Alan Rubel, and Ellen LeClere. "A matter of trust: Higher education institutions as information fiduciaries in an age of educational data mining and learning analytics." *Journal of the Association for Information Science and Technology* 71.10 (2020): 1227-1241.
- [21] Zhao, Ping, et al. "A survey of local differential privacy for securing internet of vehicles." *The Journal of Supercomputing* 76.11 (2020): 8391-8412.
- [22] Francis, Bindhia K., and Suvanam Sasidhar Babu. "Predicting academic performance of students using a hybrid data mining approach." *Journal of medical systems* 43.6 (2019): 1-15..