The Recommender System for Smart E-Learning System Using Big Data: A Survey

M. Murugeswari[1*], S. Vimala[2]

[1,2]Department of Computer Science, Mother Teresa Women's University, Kodaikanal, Tamilnadu, India

Corresponding Author: murugi1310@gmail.com, Tel.: +91-98945-20739

Abstract— Recommender systems utilize the opinions of a residential district of users. It assists individuals for the reason that the community more effectively identifies the content of great interest from a set that is potentially overwhelming. The instructor provides an online course which consists of the learning materials, self-quiz, and learning path in a virtual classroom. Typical learners study course material and do self-quiz so that you can evaluate their knowledge. The essential thing that is important to the success learners relates to the standard of the educational materials that are not only be determined by given materials by the instructor but additionally be determined by other learners' recommendations. Recommender systems have now been a helpful tool to recommend items in a lot of online systems, including e-learning. However, not much research has been done to gauge the learning effects regarding the learners if they use e-learning with a recommender system. Instead, most of the researchers were concentrating on the recommender system precision in forecasting the learner's recommendation as opposed to the knowledge gain. The detailed literature review is presented by the various researchers in the Recommender Systems for the Smart E-Learning environment in this survey article.

I. INTRODUCTION

Big data technologies describe a different generation of architecture and technologies built to extract information that is useful large volume data and allowing a proper time analysis and collection [1]. The education systems have utilized Big Data to develop various applications in order to explore massive educational data in previous years. Both basic and higher education sects have used Big Data for education data mining, which aims to analyze the data generated in education or his entourage in order to solve educational research problems [2] in recent years. The real question is how from most of the large data that is retrieved, learners (such as their traces and productions on educational systems) can both measure the qualities and weaknesses of educational resources and see the difficulties, improvements of every learner. These data that are massive their analysis pose numerous problems of protection of personal data, and many more questioned the uses that would be manufactured from these data. The learning analytics [3] gives along with the methods of Artificial Intelligence (machine learning, analysis of social networking sites) so that you can apply them in education [4]. It is constituted because of the records of action involved in events because of the automated analysis. A conference may be a link, download videos, documents, communication in a forum, the processing of exercises, etc. The number of these data assists you in reflecting on the in-patient or activity that is collective. The task of Big Data

analytics in education helps to measure the impact of collaborative exchange, and this research domain is dependent on predictive analysis, data mining, my space, and Facebook analysis adaptive algorithms [2]. It gives the approach called education that is data-driven, objectifying the selection predicated on performance in addition to the adapta- tion of pedagogical articles. Additionally, it is a potential that is economic in this regard the ongoing future of education systems, these strategies offering research possibilities for growing present e-learning systems to modify their educational content. This survey paper shall explain how education data allows us to fix education problems making use of Big Data applications in education systems.

II. BIG DATA

Big Data [5] comprises of a group of just about structured data that gets so bulky, for a good cause it is hard to work well with conventional database management tools. Big Data is described as the rule that is"3V" Volume (data implies huge volumes of raw data), Variety (data set normally consists of heterogeneous data, structured or perhaps not) and Velocity (information is generated at "high speed" and sometimes even continuous, that also involves processing them quickly, and sometimes even in real-time). Some definitions add two more V, the very first is the Veracity, which is the notion of imperfect quality of information. In essence, data contain "noise", incomplete

information, duplicate or confused information (two different versions of the identical data). The second reason is worth it: it is about to be able to pay attention to the info having a value that is real to be able to be actuated. At the moment, Hadoop could be the platform that is main of Data—utilized for storing and processing large volumes of information, this framework and its components that are variously employed by numerous many e-learning platforms with their Big Data projects. The core that is key of is composed of a storage part: HDFS (Hadoop Distributed File System) [6], and a processing part called MapReduce.

III. SMART E-LEARNING RECOMMENDER SYSTEM

The recommender method is a computer software representative that collects the score information from each user so that you can recommend or predict the essential materials that are suitable for every user. Two approaches are traditional, building recommender systems. Smart web is e-learning that allows all learners to collaborate their expertise so that you can predict the essential appropriate learning content for every learner.

IV. LITERATURE REVIEW

Udupi, Prakash Kumar, Puttaswamy Malali, and Herald Noronha [7] suggested a evaluates and identifies one of the key guidelines of the e-learning prototype and offers the alternative creating framework that is new big data integration. More examines the research that is possible, an integral system, and suggests a unique system for consolidation of big data and smart technology framework inside the e-learning prototype that guides in the direction of smart learning. New learning methods like flipped learning, project-based learning, assignment-based learning, skill-based learning, adaptive learning are discovered to date, to give you improve learning techniques. But none among these are receiving the pattern that is in-depth of learner's information or clustering of creating patterns, which will probably be the most needed aspects for upcoming generation augmented learning. In this direction, a smart learning system combines different dimensions of learning framework, such as big data integration, which necessary for next-generation learning.

Chaffai, Abdelmajid, Larbi Hassouni, and Houda Anoun [8] concentrates in the design and utilization of hybrid and modern time that is real pipeline structure using Apache Flume to get data, Apache Spark as a coordinated engine calculation for executing analysis on students' tasks data and Apache Hive as a data warehouse for keeping the processed data as well as utilized by different reporting tools. This has dealt with the task of making usage of an event-driven system around a learning platform; this is certainly web-based.

Anshari, Muhammad, et al. [9] examined present Internet literacy and activities in Brunei Darussalam (Brunei) regarding Internet behavior and learning that is online. It has outlined milestones with regards to current problems, trends, and challenges in Internet tasks, centering on online learning and its perspective when you look at the data that are big. A thorough survey that is nationwide performed so that the reliability regarding the data: the survey presented that folks in Brunei demand and expect improved learning experiences and services through an online learning system. It revealed information that is important Internet-based strategies and online learning in Brunei.

Birjali, Marouane, Abderrahim Beni-Hssane, and Mohammed Erritali [10] introduced the educational applications of Big Data technology so that you can describe why education data mining makes it possible to fix education problems. The architecture that is technical to be a relevant academic model for operating Big Data techniques and educational organizations. Later, an advantage that is essential to this task is the detection of Big Data sources from or generated by the educational environment. Also presented, how Big Data helps solve education issues through attaining the objective of learning. They introduce some opportunities for Big Data analytics to produce the effectiveness and efficiency of students learning and maximize their knowledge retention.

Huda, Miftachul, et al., [11] implemented thematic analysis to make a design for higher organizations to modify their situation on big data application. The conclusions show that it could be employed to enhance decision-making, give ideas, knowledge findings, and optimize processes that are learning. A higher learning organization can choose big data teaching that is an analytics-based learning technique to maintain in offering innovative learning and teaching experience into the students with possibilities to enhance learning activities in their minds with big data analytics. The utilization of big data in innovative teaching happens to be a promising experience that is new to the students, education providers, instructors, in addition to the community.

Khan, Samiya, Kashish A. Shakil, and Mansaf Alam [12] examined how cloud-based big data analytics may be placed on education that is Indian reviews and research the difficulties that have to be addressed ahead of the true advantages of this technology are available. It also provides a statistical view regarding the education that is Indian and investigated the stability of cloud-based big data analytics possibilities for fine-tuning education at various levels and enhancing research regarding the country. The non-technical and technical difficulties when you look at the adoption that is widespread in data technology for research and education in India have now been addressed and identified.

Santos, Olga C. [13] mentioned open issues about affective computing when you look at the educational domain. As

well as those open issues, plus in line with affective computing research, generally speaking, further efforts may also be needed in this domain to support learning that is naturalistic in the wild by ensuring generalization, utility, and authenticity.

Huba, Mikuláš, and Štefan Kozák [14] introduced a trend that is newest on the market 4.0 initiative and their effect on education in your community of Mechatronics. Also, it pointed out different common points using the learning that is flexible emerging when you look at the 2nd 1 / 2 of the earlier century along with the growth of the e-learning. A quick comparison with top world universities suggests that the HE sectors in Slovakia will not be in a position to master flexible learning development sustainably.

Jin, Sung-Hee, and Y. Kim. [15] introduced learning behavior, and activity in e-learning environments was taken as a foundation that is basic research in mastering analytics. By the modeling of learning data, the behavior is recognized, and learning analytics offer information that is meaningful to the educational process for instructors and learners. The target is to enable learning and also to support teaching tasks. So, by the model of learning information, behavior data should always be offered to both instructors and learners.

Gudivada, Venkat N., et al. [16] suggested a reference structure for cognitive analysis and suggested techniques to employ the architecture and presented learning varieties and mentioned many classes of machine algorithms that are learning. A couple of analytics that are cognitive are briefly described. Cognitive analytics is pursued from two perspectives that are complementary computer science, and cognitive and neurosciences.

Ducange, Pietro, et al. [17] analyzed the key attributes of current e-learning systems, pointing out their types of data in addition to the amount that is huge that can be retrieved from their website. Moreover, it assessed the thought of educational Big Data, indicating a logical and functional model that is layered and can turn to be very helpful in true to life. It is light that is a lean-to on various ranges where Big Data mining usually takes place as well as on the respective benefits.

Mahboob, Tahira, Sadaf Irfan, and Aysha Karamat [18] recommended a model for fair/transparent student evaluation/performance has additionally been displayed. Particular variables have now been described, which can then be efficaciously tested by applying machine learning algorithms. In this research, classifiers such as Random Forest, Naive Bayes, and Decision Trees-J48 are widely used to develop the superiority of student data by at first eliminating noisy data, and therefore getting better accuracy that is prognostic. The scope regarding the paper happens to be ready for undergraduate programs. It explored the alternative to predict the rate of success of students signed up for a program machine that is using algorithms.

Murugananthan, V., and B. L. ShivaKumar [19] projected a data model utilizing a classifier that is advanced a new agent-based intelligent method, which builds the info model in as a type of 3D cubes for almost any classification. A methodology that is optimally recommended for mining the info models. The suggested hypothesis with a 3D view that is cubical be adoptable at any example.

Authors Umbleja, Kadri, and Manabu Ichino. [20] presented an ongoing process to build a forecast model for student's grades that is the final period of finishing, predicated on students' earlier behavior. Forecast model was created utilizing data mining with principal component analysis, regression analysis, and clustering that is hierarchical of histogram valued data.

Dwivedi, Surabhi, and VS Kumari Roshni [21] projected a Recommendation system utilized filtering that is a collaborative item-based recommendation system to advise elective courses to a student. These recommendations are derived from his grade obtained can be turned out to be very useful to students to pick the elective courses. The institute that is educational design their syllabus to offer more choices to the students to decide on subjects by the specific skills and expertise regarding the students. Big data pops up along with its challenge to undertake the info, but it can be beneficial to improve the quality of the current education system and process if it is appropriately managed. The suggestions generated by such systems can be handy into the institute that is educational increases the performance of students, schools, and teachers, etc.

Appalla, Padmaja, Venu Madhav Kuthadi, and Tshilidzi Marwala [22] introduced a simple yet effective data that are an educational approach to aid the e-learning. The proposed approach is made from two modules, for instance, the server module, in addition to the client module. The documents are read from the database, and the equivalent knowledge representation is made in the server module. The information is retrieved based on the user requirements in the client module. This process is evaluated using parameters that are various given that precision, recall, in addition to F-measure. The results that are comprehensively obtained by varying the keywords, a wide range of documents together with K-size.

Vyas, Madhav S., and Reshma Gulwani [23] endorsed C4.5 Decision Tree methodology execution in Big Data's Map-Reduce framework for E-learning Systems to higher estimate students performance that is. This has analyzed data that are different algorithms utilized for forecasting students' performance in an E-Learning System. Decision Tree Algorithm is way better when students that are predicting performance were found better along with forms of data mining methods, input values, and accuracy. C4.5

Decision Tree Algorithm for analyzing students that are large data was presented. Map-Reduce Framework to its Integration is proposed in comparison with conventional Decision Tree implementation. Thus, it will probably result in more effective predictive analytics of student's performance that is. It helps us identify the learning students susceptible to dropping or failing far away from the program.

Tashtoush, Yahya M., et al. [24] proposed a brand-new e-learning system that is adaptive. The system incorporates a well-known intelligent web-based e-learning that is English with data mining techniques. Also, the info mining techniques are employed to cluster students' learning styles relating to Jackson's learning styles. The greatest aim of the proposed system is always to determine the teaching pattern that is best for every single learner. The system that is proposed be manufactured available through the internet everywhere in addition to every time (EWET). Also, it offers facilities that are adaptive as learning videos, adaptive presentations, and quizzes when it comes to students.

El Fouki, Mohammed, Noura Aknin, and K. Ed El. Kadiri. [25] examined the deep component that is principal and reinforcement learning could enhance classification models and also raise the forecast performance of a deep neural network algorithm by decreasing the wide dimensionality range of variables of information sets, without having to sacrifice the classifier reliability and accuracy. The simple fact to divide the educational phase into many agents will reduce the information that is correlated by overlapping input instances, which may lower the network training time and increase the efficiency of DNN systems, especially in solving complex problems involving a lot of input data. However, more tasks are necessary as to how we could implement Deep Learning algorithms in problems about Big Data, including dimensionality that is high scalability of Deep Learning models, criteria for extracting good data representations, and domain adaptation.

Rao, NV Krishna, N. Mangathayaru, and M. Sreenivasa Rao [26] proposed to gauge student performance to analyze student achievement and performance track that is future. A Learning algorithm is implemented by permitting student data to keep in clustered memory for repeated querying. Clustering is done to get into the student that is vital from big data and process it with amazing speed. The classification technique with a decision tree algorithm is implemented to get the learning student performance.

Tian, Feng, Qinghua Zheng, and Kuo-Ming Chao [27] examined the past reputation for the e-learning evolution process, and it may be categorized into three stages: networked, digitalized, and intelligence. The authors determined, presently, it was still when you look at the intelligence stage, named smart e-learning, showing intelligent, ubiquitous, personalized, open, and collaborative trends and demanding for in-depth integration of artificial intelligence and big data.

Cerezo, Rebeca, et al. [28] proposed an algorithm that is new to the educational field called Inductive Miner throughout the interaction traces from 101 university students in a program given over one semester from the Moodle 2.0 platform. Data were obtained from the platform's event logs with 21,629 traces in order to find out students' self-regulation models that donate to enhancing the instructional process. The Inductive Miner algorithm discovered optimal models with regards to fitness both for Pass and Fail students in this dataset, in addition to models at a specific standard of granularity, which can be interpreted in educational terms,that are the essential achievement that is important model discovery. The authors determined that although students who passed would not proceed with the instructors' suggestions exactly, they did proceed with the logic of fruitful learning that is self-regulated in the place of their failing classmates.

Moubayed, Abdallah, et al. [29] proposed the usage k-means algorithm to cluster students predicated on 12 engagement metrics divided into two categories: interaction-related and energy-related. The goal of the model would be to explore how data that are predictive models can help identify weak students at the beginning of an endeavor to boost their performance.

Razzaque, Anjum, and Allam Hamdan [30] discussed literature to investigate what character students' Learning Style (LS) and the Internet of Things (IoT) has on learner expectations, through the outcomes that are learning (LOs) of the students.

Vinutha K.N and K.S.Sampada [31] proposed a Automation Testing for the recommender system, with Feature Vector Algorithm .It performs a automation on each modules of the Feature Vector algorithm and also checks the Cross-Browser compatibility across the browser and also collects the online reviews from by using Web Crawling Technique.

V. FUTURE RESEARCH DIRECTION AND CONCLUSION

Since the future research direction, we focused on suggesting a recommender system for e-learning. This is certainly in a position to recommend similar what to the viewing item also to recommend items predicated on good learners' ratings that are. We're going to set accuracy given that objective regarding the extensive research work considering that the system will impact on student's performance happens to be put to evaluate thoroughly.

There are lots of works which can be done as time goes by justifying and enhance our work. To begin with, the operating system was tested on Software Engineering students whose computer literacy is anticipated to be high.

Thus, the user ratings as a recommendation tool are very easily understandable because of the students. It could be appropriate to evaluate the device along with other categories of students, which have another type of standard of computer literacy (e.g., sociology students) or another type of standard of education (e.g., primary school students and secondary school students). It is essential to keep in mind that the educational process must certainly be followed closely by assessments (post-tests to categorize the students when the learners that are good ratings will soon be used as a recommendation). Also, the testing was done on a set that is predefined of materials and just one type of learning material (lecture slides), which were decided by the program lecturer. It could be interesting to look at accuracy regarding the system; additionally, the student's performance when sets that are different kinds of learning materials are used.

REFERENCES

[1] Anshari, Muhammad, Yabit Alas, and Lim Sei Guan. "*Developing online learning resources: Big data, social networks, and cloud computing to support pervasive knowledge*." Education and Information Technologies Vol.**21**, Issue **6** ,pp. **1663-1677** , 2016

[2] Shen, Liping, Minjuan Wang, and Ruimin Shen. "*Affective e-learning: Using "emotional" data to improve learning in pervasive learning environment.*" Journal of Educational Technology & Society Vol.**12,** Issue **2** ,pp. **176-189, 2009**

[3] Picciano, Anthony G. "*The evolution of big data and learning analytics in American higher education.*" Journal of asynchronous learning networks Vol.**16**, Issue.**3** ,pp. **9-20,2012**

[4] Reyes, Jacqueleen A. "*The skinny on big data in education: Learning analytics simplified.*" TechTrends Vol.**59**, Issue .**2** , pp,**75-80, 2015.**

[5] Macfadyen, Leah P., et al. "*Embracing big data in complex educational systems: The learning analytics imperative and the policy challenge.*" Research & Practice in Assessment Vol.**9** , pp.**17-28** ,**2014**

[6] Prinsloo, Paul, and Sharon Slade. "*Big data, higher education and learning analytics: Beyond justice, towards an ethics of care.*" Big data and learning analytics in higher education. Springer, Cham, pp. **109-124. 2017**

[7] Udupi, Prakash Kumar, Puttaswamy Malali, and Herald Noronha. "*Big data integration for transition from e-learning to smart learning framework.*" 2016 3rd MEC International Conference on Big Data and Smart City (ICBDSC). IEEE, **2016**.

[8] Chaffai, Abdelmajid, Larbi Hassouni, and Houda Anoun. "*Real-time analysis of students activities on an e-learning platform based on apache spark.*" (IJACSA) International Journal of Advanced Computer Science and Applications Vol.**8**,Issue. **7** ,pp. **101-109** , **2017**

[9] Anshari, Muhammad, et al. "*Online Learning: trends, issues and challenges in the Big Data Era."* Journal of0 e-Learning and Knowledge Society *Vol. 12.Issue .1*, 2016.

[10] Birjali, Marouane, Abderrahim Beni-Hssane, and Mohammed Erritali. "*Learning with big data technology: The future of education."* International Afro-European Conference for Industrial Advancement. Springer, Cham, **2016.**

[11] Huda, Miftachul, et al. "*Innovative teaching in higher education: the big data approach*." TOJET **2016.**

[12] Khan, Samiya, Kashish A. Shakil, and Mansaf Alam. "*Educational intelligence: applying cloud-based big data analytics to the Indian education sector*." 2016 2nd international conference on contemporary computing and informatics (IC3I). IEEE, **2016.**

[13] Santos, Olga C. "*Emotions and personality in adaptive e-learning systems: an affective computing perspective*." Emotions and personality in personalized services. Springer, Cham, **pp.263-285** , 2016

[14] Huba, Mikuláš, and Štefan Kozák. "*From E-learning to Industry 4.0."* 2016 International Conference on Emerging eLearning Technologies and Applications (ICETA). IEEE, **2016.**

[15] Jin, Sung-Hee, and Y. Kim. "*Learning activities and learning behaviors for learning analytics in e-learning environments.*" Educational Technology International Vol.**17**.,Issue **2** pp.**175-202, 2016**

[16] Gudivada, Venkat N., et al. "*Cognitive analytics: Going beyond big data analytics and machine learning.*" Handbook of statistics. Vol. **35**. Elsevier, pp.**169-205** , 2016

[17] Ducange, Pietro, et al. "*Educational big data mining: how to enhance virtual learning environments.*" International Joint Conference SOCO'16-CISIS'16-ICEUTE'16. Springer, Cham, **2016.**

[18] Mahboob, Tahira, Sadaf Irfan, and Aysha Karamat. "*A machine learning approach for student assessment in E-learning using Quinlan's C4. 5, Naive Bayes and Random Forest algorithms.*" 2016 19th International Multi-Topic Conference (INMIC). IEEE, **2016.**

[19] Murugananthan, V., and B. L. ShivaKumar. "*An adaptive educational data mining technique for mining educational data models in elearning systems.*" Indian Journal of Science and Technology Vol.**9**, Issue. **3** pp. **1-5** , 2016

[20] Umbleja, Kadri, and Manabu Ichino. "*Predicting Students' Behavior During an E-Learning Course Using Data Mining.*" International Conference on Interactive Collaborative Learning. Springer, Cham, **2016.**

[21] Dwivedi, Surabhi, and VS Kumari Roshni. "*Recommender system for big data in education.*" 2017 5th National Conference on E-Learning & E-Learning Technologies (ELELTECH). IEEE, **2017.**

[22] Appalla, Padmaja, Venu Madhav Kuthadi, and Tshilidzi Marwala. "*An efficient educational data mining approach to support e-learning.*" Wireless Networks Vol.**23** , Issue .**4** , pp.**1011-1024. 2017**

[23] Vyas, Madhav S., and Reshma Gulwani. "*Predictive analytics for E learning system."* 2017 International Conference on Inventive Systems and Control (ICISC). IEEE, **2017.**

[24] Tashtoush, Yahya M., et al. "*Adaptive e-learning web-based English tutor using data mining techniques and Jackson's learning styles.*" 2017 8th International Conference on Information and Communication Systems (ICICS). IEEE, **2017.**

[25] El Fouki, Mohammed, Noura Aknin, and K. Ed El. Kadiri. "*Intelligent Adapted e-Learning System based on Deep Reinforcement Learning.*" Proceedings of the 2nd International Conference on Computing and Wireless Communication Systems. **2017.**

[26] Rao, NV Krishna, N. Mangathayaru, and M. Sreenivasa Rao. "*Evolution and prediction of radical multi-dimensional e-learning system with cluster-based data mining techniques.*" 2017 International Conference on Trends in Electronics and Informatics (ICEI). IEEE, **2017.**

[27] Tian, Feng, Qinghua Zheng, and Kuo-Ming Chao. "Current and future of technologies and services in smart e-learning." ,Springer pp. **1-3. 2020**

[28] Cerezo, Rebeca, et al. *"Process mining for self-regulated learning assessment in e-learning.*" Journal of Computing in Higher Education Vol. **32**.Issue.**1**,pp.. **74-88.**2020

[29] Moubayed, Abdallah, et al. "*Student Engagement Level in e-learning Environment: Clustering Using K-means.*" American Journal of Distance Education pp. **1-20.**2020

[30] Razzaque, Anjum, and Allam Hamdan. "*Internet of Things for Learning Styles and Learning Outcomes Improve e-Learning: A Review of Literature.*" Joint European-US Workshop on

Applications of Invariance in Computer Vision. Springer, Cham, **2020**

[31] Vinutha K.N and K.S.Sampada "*Survey of Automated Recommender System for Web Application",s* International Journal of Computer Science and Engineering Vol.**4** , Issue **3**, pp. **36-39 , 2016.**

AUTHORS PROFILE

*Mrs.M.Murugeswari i*s a Research Scholar in the field of Big Data , currently working as an Assistant Professor in Department of Computer Applications at EMGYWC , Madurai, India. She completed her MCA from MKU(TamilNadu) and undergoing Ph.D at Mother Teresa Women;s University (Kodaikanal),India .
Email : murugi1310@gmail.com

Dr.S.Vimala , Associate Professor, Department of Computer Science, Mother Teresa Women's University, Kodaikanal, Tamil Nadu. She was a Liaison Officer for the Search Committee for the Appointment of Vice –Chancellor during Feb 2019. NMEICT Nodal Officer. Coordinator, Incubation and Technology Transfer Centre (ITCC). Coordinator, Internal Quality Assurance Cell (IQAC).
Email: vimalamtwu70@gmail.com