

# Deep Learning in the field of Education

Dr. Divya Kashyap

Assistant Professor,

Indian Academy Degree College Autonomous, Bangalore.

**Abstract:** Educational Data Mining could be a analysis field that focuses on the applying of machine learning, information mining and statistical methods to detect patterns in massive collections of instructional information. Various machine learning techniques area unit applied throughout this field over the years, but it's been recently that Deep Learning has gained increasing attention among the academic domain. Deep Learning may well be a machine learning technique supported neural network architectures with multiple layers of method units, that has been successfully applied to a broad set of problems among the areas of image recognition and language method. This paper surveys the analysis dole out in Deep Learning techniques applied to [EDM], from its origins to the current day. The most goals of this study to square measure to spot the [EDM] tasks that have benefited from Deep Learning and people that square measure unfinished to be explored, to explain the main data sets used, to produce an outline of the key ideas, main architectures, and configurations of Deep Learning and its applications to EDM, and to debate current progressive and future directions on this space of analysis.

## 1. Introduction

The analysis field of Educational Data Mining focuses on the applying of techniques and strategic of data mining in educational environments. EDM is concerned with researching, developing, and applying machine learning, data processing, and applied math strategies to notice patterns in giant collections of educational data that will preferable be not possible to research [1].

EDM leverages e-learning platforms like Learning Management System (LMS), Intelligent Tutoring Systems (ITS), and, within the last years, huge Open on-line Courses (MOOC), to

get made and multi modal data from student's learning activities in academic settings. For instance, these platforms record when the students access a learning object, how many times they accessed it, whether the answer provided to an exercise is correct or not, or the amount of time spent reading a text or watching a video.

All this data will be analyzed to handle completely different educational problem, like generating recommendations, developing adaptive systems, and providing automatic grading for the students' assignments. completely Different machine learning techniques have been applied over time to research this information, however it's been in recent years that the utilization of Deep Learning techniques has emerged within the field of EDM.

The topic of Deep Learning (DL) has gained increasing attention within the industry and research areas within the last decade, revolutionizing the sector of machine learning by getting progressive ends up in perception tasks like image and speech recognition[2]. Major companies such as Google, Facebook, Microsoft, Amazon, and Apple are heavily investing in the development of software and hardware innovations in this field, trying to leverage DL potential in the production of smart products.

DL is based on neural network architectures with multiple layers of processing units that apply linear and nonlinear transformations to the input data. These architectures can be applied to all type of data: image, audio, text, numerical, or some combination of them. Many research fields have benefited from applying these technologies, and EDM is not an exception.

In the previous few years there has been a proliferation of analysis within the Educational

Data Mining field using DL architectures. This article presents a review on the literature of DL methods applied to Educational Data Mining, from its start in 2015 to the current day. The primary contributions of this article are as follows:

1. Summarize the most EDM tasks and classify the present works that have applied DL on every of those tasks.
2. Identify the tasks that have gained major attention and those that are still unexplored.
3. Describe and categorise the most public and personal datasets employed to train and test DL models in Educational Data Mining tasks.
4. Introduce key DL ideas and technologies, describing the techniques and configurations most generally utilized used in Educational Data Mining and its specific tasks.
5. Discuss future directions for analysis in DL applied to Educational Data Mining based on the data gathered during this study.

## 2. Machine learning and its application in education area

Machine learning can be considered a part of Artificial Intelligence [AI]. Machine learning is the method of granting a machine or model access to information and property it learn for itself. In 1959, Arthur Samuel came up with the good concept we must always not need to teach computers, but rather, we have a tendency to might allow them to their own. He coined the term “machine learning” to describe his theory, which is now a standard definition for the ability of computers to learn autonomously [3].

Machine learning is programming computers to optimize a performance criterion using example data or experience [4]. Implementing a machine learning algorithm means implementing a model that outputs correct information given that we have provided input data. You can think of a model as a black box: data go in at the beginning, and some other data go out at the end — but the processes in between are

complex. For example, if we want to create a model that predicts what the house price in some region next year will be, based on situation on the market for the last three years, we would feed the model metrics such as house prices on the market in last three years, interest rates and salary rates.

The output would be the house price prediction for the next year. The process through which a model learns how to make sense of input data is called “model training”. Training is a key concept in machine learning.

The best way to describe the potential of machine learning is to explore how people and companies are currently taking advantage of it. Some examples could be:

- Natural language processing: - Google Translate is created from a set of machine learning algorithms that updates the service over time based on input from users, like new words and syntax. Siri, Alexa, and, last, Cortana, and, Google Assistant all rely on natural language processing to recognize speech and synthesis, permitting them to know or pronounce words they have never encountered before.
- Recommendation systems: - On Amazon, Netflix, Google, etc., everything that's suggested to you depends on your search activity. These websites deliver recommendations across platforms, devices, and apps. Machines match buyers with sellers and their products, digital content with viewers who want to see them - all of which improves our online experiences significantly. The company even owns a patent for Anticipatory Shipping, a system that ships a product to the closest warehouse thus you'll be able to order and receive your item on the constant day although it's unclear whether or not they have enforces it yet.

- Algorithmic trading: - Algorithmic trading is a process that involves random behaviour, ever-changing data, and a variety of factors. While financiers cannot predict all of that behaviour, machine learning algorithms will — and that they reply to changes within the market quicker than somebody's.

There are many alternative business implementations of machine learning, heap of them area unit in education area . a number of fascinationg areas are:

- Predict Student Performance (A nice and great application of machine learning is predicting student performance. By learning concerning every student, the machine learning model will conclude weaknesses and suggests ways that to enhance like extra lectures or study extra literature ).
- Test Students & Grade Students Fairly (Machine learning can help creating computerized adaptive assessments . The machine learning based assessment provides constant feedback to faculty members and students regarding how the student learns, the support they have and also the progress towards their learning goals ).
- Improve Retention (Machine learning, like learning analytic, will also help improve retention rates. By identifying “at risk” students, facilities will reach intend on those students and obtain them the assistance they have to achieve successful ).
- Support academics and institution stuff (Machine learning based mostly algorithms will facilities with classification of students written assessment papers)

### 3. Research

The above-mentioned research questions were extracted when we studied the aims of this research. To answer these research questions, we performed a literature study. We decided to use the SLR method to collect the relevant primary studies and followed the guidelines given by[5].

In the next phase, we created four categories and classified the papers into those. The reason for this classification was that most of the papers published were relatively distinctive in terms of the research objective, methodology, and application. To be simplest as possible, we created four categories of studies without ignoring the variations of themes. This way, we examined the research papers that fell under the same. The final number of papers collected was 67. Some studies was included in more than just one category.

The research categories and the total number of selected studies are as follows:

- (A) Grading students (12 studies)
- (B) Improving student retention (17 studies)
- (C) Predicting student performance (42 studies)
- (D) Testing students (6 studies)

Letters A, B, C and D are just used as tags and serve as a mark for further referencing in the rest of this study.

### 4. Result

This section describes the main characteristics of each category. We present all four main categories in subsections, supported with relevant papers.

## 1. Grading students

Machine learning can grade students by removing human biases. Some recent examples are use of the supervised Machine Learning for text classification to predict students' final course grades in some course and exhibited the potential of using ML classified messages to identify students at risk of course failure [9]. In addition, there are aims to improve the assessment of problem solving in education by employing language technologies and computational-statistical machine learning methods to grade students' natural language responses automatically [10]. Great example of use machine learning for grading students is by comparing their actions to a model of expert behaviour [11].

## 2. Improving student retention

As we said before, by identifying at risk students early, faculties will find and call those students and facilitate them to be more successful. Student retention is an essential part of many enrollment systems. It affects most all segments of university school or college metrics: name, financials, reputation, ranking. Specially, student retention has become one of the most important things for managers in higher education institutions. There are few studies, which developed models to predict and to introduce the reasons behind student's number decreasing.

## 3. Predicting student performance

Probably a serious good thing about machine learning (regarding range of studies in scientific databases) is its ability to predict student performance. By "learning" regarding every student, the technology can identify weaknesses and suggests ways in which to enhance, like additional practice tests. This appears to be very popular analysis trend; there are lot of studies in

recent years in this area, as we said before. For example, study[6] employs the machine learning approach called the Recursive Clustering method to group the students of the programming course into groups based on their performance within the necessity courses, co-requisite and current course work result. Students present in the lower groups will be taken into consideration since they are highly prone to fail. In another interesting study in this category, authors have proposed a new model to categorize students into three categories to determine their learning capabilities and to help them to improve their studying techniques. They have chosen the state of the art of machine learning approach to classify student's nature of study by choosing distinguished options of their activity in their academic field. They have chosen a data driven approach where key factors that determines the base of student and classify them into high, medium and low ranks. Yet another study in this category, but from other perspective brings.

## 4. Testing students

The machine learning based assessment provides constant feedback to lecturers, students and parents regarding however the student learns, the support they need and the progress they are making towards their learning goals. Authors in study [10] introduced a system for training of students' ability to construct correct proofs in propositional or predicate logic. In addition to common techniques, as well as shows supported by slides and exercises they used animations, which were based on carefully selected demonstrative examples and their step-by step solutions. as to check students' knowledge, they prepared a questionnaire that captured the whole method of a logic proof construction. A student constructed a proof and then answered queries from the questionnaire. They described the design of the questionnaire and discussed its dis/advantages. At the end, they then applied frequent sub graph mining along with supervised



machine learning algorithms to perform an automatic evaluation of correctness of the proofs.

## Conclusion

This study has reviewed the emergence of DL applications to EDM, a trend that started in 2015 with three papers published, increasing its presence once a year so far with 19 papers published in 2019. After a systematic search, 41 works were retrieved in this area. It is worth mentioning the presence of these approaches in relevant EDM forums such as the annual International Conference in Educational Data Mining, with 7 papers published in the last edition (for a total of 16 in the last three years). Based on the taxonomy of EDM applications defined by [11], only 4 of the 13 tasks proposed in that study have been addressed by DL techniques. This reveals that there are several open opportunities for the use of DL in unexplored EDM tasks, moreover taking into account the promising results obtained by these models in the works reviewed (67% of them reported that DL outperformed the “traditional” machine learning baselines in all their experiments). The aim of this study was to judge the present state of the art within the application of machine learning in education space. The amount of studies (papers and articles) was large, so only some of studies, which we found as good representatives, were mentioned in results this study. This study shows that there a square measure important other ways to profit from machine learning application in education area. As we tend to expressed in introduction section, one of our goals was attempt to classify studies within the field of machine learning application in education space. Based on our survey, the papers reviewed under category marked as A research a ways how machine learning can grade students by removing human biases (fairly grading). Reviewing studies under category marked as B, showed how machine-learning algorithms will facilitate faculties or schools or faculties to reach out to students and get them the help they need to be successful as

early as possible. It affects university rankings, faculty name, school reputation, and financial wellbeing. Student retention has become one in all the foremost vital priorities for decision makers in higher education institutions, so there are lot of studies in that category. Reviewing studies under category marked as C, showed us how major good thing of machine learning regarding number of studies in scientific databases is its ability to predict student performance. By “learning” about each student, the technology can identify weaknesses and suggests ways to improve. According to our survey, this is most interesting area of machine learning application to researchers. There are lot of studies in recent years in that category, and lot of machine learning models were provided to predict student performance on totally different parameters. We might say that this category is definitely the trend. Reviewing studies under category marked as D, showed some models however machine learning will facilitate move far from standardized testing. Machine learning based assessment provides constant feedback to lecturers, students and parents regarding how the student learns, the support they need and the progress they're creating towards their learning goals. As we have found earlier, research has been made over several relevant databases, but of course, not all were involved, so this can be considered as limitation of study. In addition, there is a possibility that some of the relevant studies may be skipped by chance.

In the future, we have a tendency to arrange to implement own machine learning model for suggesting potential student to enroll or not to enroll on University College algebra, Study of Software Engineering, based on different parameters. As we have rich database with lot of information of students on previous years, we believe that study would be of help to support our admission office as help in student enrolment process.

## References

1. C. Romero, S. Ventura, M. Pechenizkiy, and R. Baker, *Handbook of educational data mining*, CRC Press, 2010.
2. I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, MIT Press, Cambridge, Mass, USA, 2016.
3. Samuel, A.L. (1959) Some Studies in Machine Learning Using the Game of Checkers, *IBM Journal of Research and Development*, vol. 3, no. 3, pp. 210-229, doi: 10.1147/rd.33.0210
4. <https://www.thetechedvocate.org/8-ways-machine-learning-will-improve-education>, (2018), Accessed on: 2018- 09-25
5. Kitchenham. B.A., Charters,S. (2007) Guidelines for performing Systematic Literature Reviews in Software Engineering. Version 2.3, EBSE Technical Report EBSE-2007-01, Keele University, Keele, Staffordshire, United Kingdom
6. Wu, J. Y., Hsiao, Y. C., & Nian, M. W. (2018). Using supervised machine learning on large-scale online forums to classify course-related Facebook messages in predicting learning achievement within the personal learning environment. *Interactive Learning Environments*, 1-16.
7. Wang, H. C., Chang, C. Y., & Li, T. Y. (2008). Assessing creative problem-solving with automated text grading. *Computers & Education*, 51(4), 1450-1466.
8. Abbott, R. G. (2006, June). Automated expert modeling for automated student evaluation. In *International Conference on Intelligent Tutoring Systems* (pp. 1-10). Springer, Berlin, Heidelberg.
9. Anand, V. K., Rahiman, S. A., George, E. B., & Huda, A. S. (2018, March). Recursive clustering technique for students' performance evaluation in programming courses. In *Majan International Conference (MIC)*, 2018 (pp. 1- 5). IEEE.
10. Vaculík, K., Popelínský, L., Mráková, E., & Jurco, J. (2013, October). Tutoring and automatic evaluation of logic proofs. In *European Conference on e-Learning* (p. 495). Academic Conferences International Limited.
11. B. Bakhshinategh, O. R. Zaiane, S. ElAtia, and D. Ipperciel, "Educational data mining applications and tasks: A survey of the last 10 years," *Education and Information Technologies*, vol. 23, no. 1, pp. 537–553, 2018.