International Journal of Scientific Research in Computer Science, Engineering and Information Technology
ISSN : 2456-3307 (www.ijsrcseit.com)
doi : https://doi.org/10.32628/IJSRCSEIT

# An Improved Deep Learning Algorithm for Student Achievement Prediction

**Sayed Noorusabah[1], Dr. Sweta Ramendra Kumar[2], Mohamed Hasan Phudinawala[3]**
[1]Lecturer in I.T Department, M. H. Saboo Siddik Polytechnic, Mumbai University, Mumbai, Maharashtra, India
[2]Assistant Professor, B.Sc-IT, Niranjana Majithia College of Commerce, Mumbai University, Mumbai, Maharashtra, India
[3]Assistant Professor, B.Sc.-IT Department, Niranjana Majithia College of Commerce, Mumbai University, Mumbai, Maharashtra, India

**Article Info**
Volume 8, Issue 1
Page Number : 187-192

**Publication Issue :**
January-February-2022

**Article History**
Accepted : 01 Feb 2022
Published : 10 Feb 2022

**ABSTRACT**

Educational Data Mining (EDM) research has risen to prominence because it aids in the discovery of relevant knowledge from educational data sets that can be used for a variety of reasons, such as forecasting students' academic performance and outcomes. Predicting student accomplishment may be beneficial in the development and implementation of a variety of improvements in education settings as a response to current educational systems. Machine learning has been used to predict students' achievement in a huge amount of existing research, which has taken a variety of factors into account, including family income, students' gender, students' absence, and stage-by-stage characteristics. In this proposal, an attempt is made to investigate the usefulness of applying the Deep Learning Algorithm (DLA), more specifically the Optimal Deep Neural Network (ODNN), to forecast students' progress, which could aid in determining whether or not students would be able to complete their degree. Using experimental data, it was discovered that the suggested framework outperformed the current one that was within reach in terms of accuracy of prediction.
**Keywords :** Deep Learning, Optimal Deep learning Algorithms (ODNN), Classification, Students performance Prediction

## I. INTRODUCTION

Educational institutions are operating in an increasingly competitive and multifaceted environment, and they are under pressure to deal with global and national economic issues. As a result, these institutions are attempting to bring about change, such as the growing need to increase the proportion of students in explicit disciplines, and to ensure that the quality of learning programmes is both internationally and nationally comparable [1]. Students are the most essential stakeholders in educational institutions, and their participation and presentation play a significant role in the social and economic growth of a country by producing innovative former students. There is a fundamental interest for scholarly institution to keep up with and slot in vast datasets of pupils for versatile basic leadership in order to maintain their position.

Internet-based learning has also become an important aspect of the contemporary period of education in many colleges, increasing the amount of accurate information available about students, instructors, and their interactions with learning and informational frameworks and resources. The advancement and development of a society are directly related to the level of education received. The meadow contains information about members such as students, teachers, offices, and educational programmes, and it is divided into several sections. The performance of kids is a primary concern for a variety of stakeholders, including instructors, administrators, and the organisation. As a result, students must put forth significant effort to achieve high academic status in order to maximise their potential as educational partners. Educational institutions place great emphasis on student achievement since the remarkable educational achievements of a specific institution of higher education will result in the improvement of the overall quality of that university. It is possible to determine a student's performance by measuring their co-curriculum and earning an assessment. Despite this, the majority of studies have concluded that graduation is the best indicator of a student's performance [3].

The following is a succinct summary of the primary contribution of this paper: Machine Learning (ML) is a technique that is commonly utilised in academic institutions to predict the performance of students. Several machine learning algorithms are used in this study, with a variety of distinct attributes being addressed and taken into account for the prediction of students' performance. It is the purpose of this paper to address and examine the capability of employing the Optimal Deep Neural Network (ODNN) and Deep Learning Algorithm (DLA) in the neighbourhood of predicting students' achievements, where several parameters are manipulated in order to produce the best performance result in terms of recall, precision, and F-measures. Because of the limited number of similar works, the work provided in this paper produces promising outcomes in assessment. In [8] and [18] developed a theoretical framework that distinguishes the different forms of analytics and their engagement in each of the academic areas. Many prior studies showed an array of different types of analytics and how it may be applied in the academic fields. In the end, they offered a synthesized list of analytics-related phrases that are regularly seen in the academic community. Data mining is a concept that is based on the extraction of hidden patterns and the discovery of correlations between parameters in a large amount of information. With the help of Nave Bayes, Decision Tree, and Rule Based categorization techniques, the authors in [9] and [19] developed a framework for predicting students' presentation in order to construct the best students' performance prediction model. The results demonstrate that the imperative base is indeed the preferred product among the other sorting techniques by achieving the highest accuracy value.

Two experiments are carried out on the basis of data collected from Mexican high school pupils. The majority of the currently underway research on EDM application to resolve the problems of students dropping out of courses is directed at clear-thinking scenarios in higher education institutions. It has been discovered through study into vital education dropout rates that the suggested work develops upright classification models to assist students in making an early decision to dropout, before the middle of the course [10] and [18]. Several scholars discuss the accomplishments of students as they progress from one stage of their educational journey to the next.

## II. RESEARCH METHOD

**Dataset:**

The data set used in this work is derived from a learner activity follower tool [15, 17] experience API [Application Programming Interface] (xAPI). Using the experience API, you may create a Total Learning Architecture (TLA) that tracks your learning progress and keeps track of your tricks, such as writing an article, watching a video, or reading an article, among other things. The experience API enables the learning platform facilitator to identify the student, as well as their behaviours and all other associated fundamentals that may be useful in addressing the learning live out scenario. 480 student records and 16 attributes make up the dataset in this study. Generally speaking, these considerations can be divided into three categories: (1) Demographic characteristics such as gender and ethnicity are taken into consideration. (2) Academic-related characteristics such as educational stage, grade level, and section are included. (3) Behavioral characteristics such as raising one's hand in class, opening resources, doing surveys with peers, and being satisfied with one's school. The dataset is divided into two genders, with 305 males and 175 females, respectively, represented. This data set contains all of the information about the participating students, who come from a variety of countries. For example, there were 179 participating students from Kuwait, 172 participating students from Jordan, 28 participating students from Palestine, 22 participating students from Iraq, 17 participating students from Lebanon, 12 participating students from Tunis, 11 participating students from Saudi Arabia, 9 participating students from Egypt, 7 participating students from Syria, 6 participating students from the United States, Iran, and Libya, and 4 participating students from the Middle East. For the dataset, two academic semesters are required: in the first semester, total numbers of gathered data are registered for 245 students; in the second semester, overall statistics of collected data are documented for 235 students. The dataset is collected over the course of two academic semesters. The information gathered includes student attendance data, which revealed that in total, 191 of the participating students were missing for more than 7 days and 289 of the participating students were absent for less than 7 days, resulting in a total of 191 absent students. This dataset does not overlook the critical impact that parents have in their children's learning and achievement as they progress through the learning and achievement process. The parents of the kids are also asked to participate in this data collection by completing a survey about their satisfaction with the educational process and all other elements related to it. As indicated in Table 1, the data set properties are as follows:

| Attribute Name | Description |
|---|---|
| Gender | Students Gender |
| Nationality | All Participated Students Country |
| Place of Birth (POB) | All Participated Students Place of Birth Country |
| Educational Stages | The Educational Levels of the Participated Students |
| Grade Levels | The Grade Levels of the participated students |
| Section ID | Participated Students Section number |
| Topic | The Course Area |
| Semester | The Academic Term |
| Parents Responsible for Student | The Participated students associated parent gender |
| Raised Hand | Total No. of raising hand per semester ( 0-100 ) |
| Visited Resources | Total No. of visiting course materials |
| Viewing Announcements | Total No. of reading announcements |
| Discussion Groups | Total No. of discussion participation |
| Parents Answering Survey | Total No. of surveys that have been answered by parents |
| Parents School Satisfaction | School Satisfaction from parents perspectives |
| Student Absence Days | Total No. of participated students Absences |

**Fig.1** Data Set Attributes

## III. RESULTS AND ANALYSIS

The Optimal Deep Neural Network (ODNN) is comprised of Convolution 1D composition as a baseline formation for unique findings. The next experiment is performed for the above-mentioned network topologies as a follow-up to the previous experiment. A large number of settings have to be defined and manipulated in order to achieve the most favourable recital precision. For the most part, the experiment was well thought-out, involving a large number of epochs and an unusual number of layers to ensure Optimal Deep Neural Network (ODNN) is extensively used for prediction purposes, as Optimal Deep Neural Network (ODNN) provide a large number of settings that have to be defined and manipulated in order to achieve the most advantageous recital precision. There is a strong collision between the performance of the Optimal Deep Neural Network (ODNN) model and these criteria (i.e. activation function, number of layers, number of neuron in each layer and hyper-parameter). The essential steps of the research technique are depicted in Fig. 2. Following the collection of the data set, the suggested research technique will undertake a series of experiments to train the model, with the remaining experiments being utilised to test the proposed hypotheses after the data set has been gathered. This research focuses on the number of epochs and the number of layers as the primary elements to be manipulated in order to attain the highest possible exactness in the calculations. After the trials are completed, the accuracy measures obtained will be compared with those obtained from a few other research in the dataset that have addressed the prediction of students' achievement. The primary step of the planned research is depicted in the diagram.

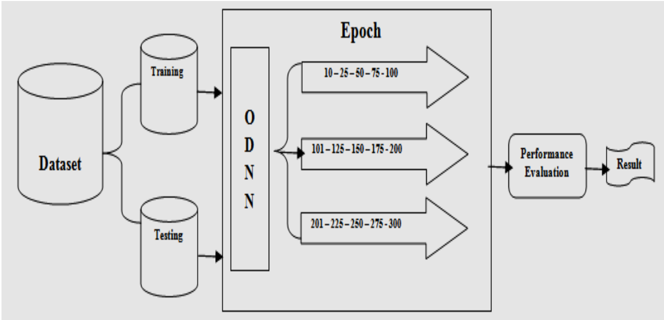


**Fig.2** Methodology Steps

The best result in terms of precision and F-Measures was achieved here. Furthermore, the precision and recall measures have been planned in advance. Tables 2 to 6 demonstrate how increasing the number of epochs in the experiments produces better results in terms of accuracy. As a comparison, three levels of layer are used in this article, and for each experiment, the number of epochs is gradually increased from 10-200, which results in the highest accuracy. As a result, only five primary research are carried out, but in reality 15 are carried out in total. Tests are carried out with the number of layers being manipulated varying from 1-3 for the purpose of carrying out each test. When the number of epoch is 200 and the number of layers 2 and 3 are used, it is clear that testing number five achieves the highest level of prediction precision.

| Table 2. Number of Epoch (10) | | | | |
|---|---|---|---|---|
| ODNN's Results | | | | |
| No. of Layer Accuracy Precision Recall | | | | |
| F- Measure | | | | |
| ONE | .594 | .468 | .594 | .509 |
| TWO | .620 | .620 | .620 | .613 |
| THREE | .575 | .687 | .575 | .499 |

| Table 3. Number of Epoch (25) | | | | |
|---|---|---|---|---|
| ODNN's Results | | | | |
| No. of Layer Accuracy Precision Recall | | | | |
| F- Measure | | | | |
| ONE | .626 | .629 | .690 | .579 |
| TWO | .829 | .832 | .829 | .827 |
| THREE | .892 | .909 | .892 | .885 |

| Table 4. Number of Epoch (50) | | | | |
|---|---|---|---|---|
| ODNN's Results | | | | |
| No. of Layer Accuracy Precision Recall | | | | |
| F- Measure | | | | |
| ONE | .639 | .675 | .639 | .641 |
| TWO | .981 | .981 | .981 | .980 |
| THREE | 1.00 | 1.00 | 1.00 | 1.00 |

| Table 5. Number of Epoch (100) | | | | |
|---|---|---|---|---|
| ODNN's Results | | | | |
| No. of Layer Accuracy Precision Recall | | | | |
| F- Measure | | | | |
| ONE | .816 | .819 | .816 | .817 |
| TWO | .993 | .993 | .993 | .993 |
| THREE | 1.00 | 1.00 | 1.00 | 1.00 |

| Table 6. Number of Epoch (200) | | | | |
|---|---|---|---|---|
| ODNN's Results | | | | |
| No. of Layer Accuracy Precision Recall | | | | |
| F- Measure | | | | |
| ONE | .955 | .956 | .955 | .955 |
| TWO | 1.00 | 1.00 | 1.00 | 1.00 |
| THREE | 1.00 | 1.00 | 1.00 | 1.00 |

Table 7 summarizes a limited assessment based on a small number of studies that have been conducted recently in the subject area under consideration. Based on the results provided in Table 7, the highest result was obtained by [10], who used the Naive Bayes Algorithm to reach an accuracy of 89 percent. Compared to the forecasted performance of the related works, the recommended move toward in proposal outperforms the forecasted performance. Furthermore, upwards of a few dimensions, not simply recall, were included in the design in order to offer the reader with an all-encompassing perspective on computation performance.

| Table 7. Related Work Results | | |
|---|---|---|
| Reference | Method Used | Accuracy |
| [9] | Rule Based (RB) | 0.713 |
| [9] | Native Bayes (NB) | 0.67 |
| [9] | DT Decision Tree (RB) | 0.688 |
| [2] | SVM | 0.867 |
| [10] | Support Vector Machines | 0.872 |
| [10] | Native Bayes | 0.89 |
| [10] | Class Association Rules (CAR) | 0.80 |

## IV. CONCLUSION

According to this article, one's ability to use the Optimal Deep Neural Network (ODNN) as well as the Deep Learning Algorithm (DLA) in the domain of forecasting students' accomplishment is addressed. A large number of parameters were modified in order to get the desired performance in terms of recall, precision, and F-measures, among other things. When compared to the results of a few relevant studies, the work provided in this research yields promising results, which is encouraging. As part of future research, we will strive to get a larger data set, develop and deploy more Deep Learning Algorithms, and assess their performance in the region under investigation.

## V. REFERENCES


[1]. Daniel, Ben, "Big Data and analytics in higher education: Opportunities and challenges," British journal of educational technology, vol. 46, no. 5, pp. 904-920, 2015.

[2]. Daud, Ali, et al., "Predicting student performance using advanced learning analytics," Proceedings of the 26th international conference on world wide web companion. International World Wide Web Conferences Steering Committee, 2017.

[3]. Shahiri, Amirah Mohamed, and Wahidah Husain, "A review on predicting student's performance using data mining techniques," Procedia Computer Science vol. 72, pp.: 414-422, 2015.

[4]. Van Barneveld, Angela, Kimberly E. Arnold, and John P. Campbell, "Analytics in higher education:

Establishing a common language," EDUCAUSE learning initiative vol. 1, no. 1, pp. l-ll., 2012.

[5]. Márquez-Vera, Carlos, et al., "Early dropout prediction using data mining: a case study with high school students," Expert Systems, vol. 33, no. 1, pp. 107-124, 2016.

[6]. Mansur, Andi Besse Firdausiah, Norazah Yusof, and Ahmad Hoirul Basori, "Personalized Learning Model based on Deep Learning Algorithm for Student Behaviour Analytic," Procedia Computer Science vol. 163, pp. 125-133, 2019.

[7]. Patil, Akhilesh P., Karthik Ganesan, and Anita Kanavalli, "Effective deep learning model to predict student grade point averages," IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), pp. 1-6, 2017.

[8]. Waheed, Hajra, et al., "Predicting academic performance of students from VLE big data using deep learning models," Computers in Human Behavior, vol. 104, pp. 106189, 2020.

[9]. Ahmad, Fadhilah, Nur Hafieza Ismail, and Azwa Abdul Aziz, "The prediction of students' academic performance using classification data mining techniques," Applied Mathematical Sciences, vol. 9, no. 129, pp. 6415-6426, 2015.

[10]. Márquez-Vera, Carlos, et al., "Early dropout prediction using data mining: a case study with high school students," Expert Systems, vol. 33, no. 1, pp. 107-124, 2016.

[11]. Kim, Byung-Hak, Ethan Vizitei, and Varun Ganapathi, "GritNet: Student performance prediction with deep learning," arXiv preprint arXiv:1804.07405, 2018.

[12]. Tu, Yuwei, Weiyu Chen, and Christopher G. Brinton, "A Deep Learning Approach to Behavior-Based Learner Modeling," arXiv preprint arXiv:2001.08328, 2020.

[13]. Hussain, Sadiq, et al., "Prediction model on student performance based on internal assessment using deep learning," International Journal of Emerging Technologies in Learning (iJET), vol. 14, no. 08 pp. 4-22, 2019.

[14]. Mansur, Andi Besse Firdausiah, Norazah Yusof, and Ahmad Hoirul Basori, "Personalized Learning Model based on Deep Learning Algorithm for Student Behaviour Analytic," Procedia Computer Science vol. 163, pp. 125-133, 2019.

[15]. Patil, Akhilesh P., Karthik Ganesan, and Anita Kanavalli, "Effective deep learning model to predict student grade point averages," IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), pp. 1-6, 2017.

[16]. Waheed, Hajra, et al., "Predicting academic performance of students from VLE big data using deep learning models," Computers in Human Behavior, vol. 104, pp. 106189, 2020.

[17]. Students' Academic Performance Dataset, Online], available https://www.kaggle.com/aljarah/xAPI-Edu-Data, Access on August 2019.

[18]. R. Aruchamy, and V.K. David. "A comparative performance study on hybrid swarm model for micro array data," International Journal of Computer Applications, vol. 975, 2011.

[19]. R. Aruchamy. "A hybrid metaheuristic algorithm for classification using micro array data," International Journal of Scientific and Engineering Research, vol. 3, pp.520-528, 2012.

**Cite this article as :**

Sayed Noorusabah, Dr. Sweta Ramendra Kumar, Mohamed Hasan Phudinawala, "An Improved Deep Learning Algorithm for Student Achievement Prediction", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 8 Issue 1, pp. 187-192, January-February 2022. Journal URL : https://ijsrcseit.com/CSEIT228138