Recent systematic review on student performance prediction using backpropagation algorithms

Edi Ismanto[1], Hadhrami Ab Ghani[2], Nurul Izrin Md Saleh[2], Januar Al Amien[3], Rahmad Gunawan[3]
[1]Department of Informatics Education, Faculty of Teacher Training and Education, Universitas Muhammadiyah Riau, Pekanbaru, Indonesia
[2]Department of Data Science, Faculty of Bioengineering and Technology, Universiti Malaysia Kelantan, Kota Bharu, Malaysia
[3]Department of Informatics Engineering, Faculty of Computer Sciences, Universitas Muhammadiyah Riau, Pekanbaru, Indonesia

Article Info

Article history:

Received Oct 15, 2021
Revised Mar 29, 2022
Accepted Apr 06, 2022

Keywords:

Data preprocessing
Deep learning
Deep neural network
Students' performance
Systematic review

ABSTRACT

A comprehensive systematic study was carried out in order to identify various deep learning methods developed and used for predicting student academic performance. Predicting academic performance allows for the implementation of various preventive and supportive measures earlier in order to improve academic performance and reduce failure and dropout rates. Although machine learning schemes were once popular, deep learning algorithms are now being investigated to solve difficult predictions of student performance in larger datasets with more data attributes. Deep neural network prediction methods with clear modelling and parameter measurements formulated on publicly available and recognised datasets are the focus of the research. Widely used for academic performance prediction, backpropagation algorithms have been trained and tested with various datasets, especially those related to learning management systems (LMS) and massive open online courses (MOOC). The most widely used prediction method appears to be the standard artificial neural network approach. The long-short-term memory (LSTM) approach has been reported to achieve an accuracy of around 87 percent for temporal student performance data. The number of papers that study and improve this method shows that there is a clear rise in deep learning-based academic performance prediction over the last few years.

Corresponding Author:

Edi Ismanto
Department of Informatics Education, Faculty of Teacher Training and Education
Universitas Muhammadiyah Riau
Jalan Tuanku Tambusai, Kota Pekanbaru, Provinsi Riau, Indonesia
Email: edi.ismanto@umri.ac.id

1. INTRODUCTION

In recent years, analysis and evaluation of student performance have become the essential indicators for academic quality assurance. Management and maintenance of an excellent academic atmosphere will support any sound effort to improve education. One of them is maintaining student performance to help them to complete their studies on time whilst reducing the dropout rate. The large number of students in Indonesia who drop out of college is found to cause a lot of social problems. Based on a report issued by the ministry of research, technology and higher education in 2018/2019, the number of students dropping out of college is large in Indonesia, amounting to 27.86%. With better academic management, preventative measures are conceivable to reduce failure and dropout rates amongst students based on student performance predictive analysis. One possible way to design an intelligent predictive analysis of student performance is via machine learning (ML), through which various algorithms and models can be developed and applied, such as neural network architectural models and deep learning (DL) algorithms to predict students' academic performance. Unlike conventional ML techniques, DL is able to perform multi-level representation of the raw input data via neural network architectures, thus rendering more advanced learning activities. By analysing the predicted student performance, more insights will be discovered much earlier before the students are expected to complete their study programs. Therefore, the paper's main objectives are to conduct a systematic review on the current deep neural network (DNN) models proposed in the literature for predicting students' academic performance and propose a new framework for the prediction task to achieve better performance. The multi-dimensional nature of the current education data requires advanced prediction techniques [1], [2] based on the artificial intelligence technologies such as ML and DL approaches [3], [4]. The ML models for predicting the performance of students have been widely studied in the past [5]. An ML approach proposed is designed based on a bi-layered structure to track and predict student performance.

Nowadays, DL constitutes one of the most promising fields in ML for solving various problems, including education management. It is both time and cost-efficient, especially in exploring high-dimensional data by employing a backpropagation algorithm. Significant advancements in DL with tremendous performance in numerous applications have been carried out in various fields and problem domains including computer vision, intelligent transportation, financial and educational analysis. DL, which comes with various DNN architectures including recurrent neural network (RNN), convolutional neural network (CNN) and long short-term memory (LSTM) network, offers various advantages for classification and prediction problems over the traditional ML models [6]. With advanced learning ability and higher prediction accuracy, DL is suitable to be employed for exploring datasets containing a rich set of features, such as in the education domain.

This paper studies the recent DL algorithms and models for student performance prediction. The main contributions of this paper are a recent systematic review of the current DL mechanisms in the literature for predicting student performance and a new DL framework proposed for performance prediction with clear modelling and parameter measurements based on public and recognised datasets. Hence, three research questions are addressed in this paper: 1) what is the best DNN architecture for predicting student performance? 2) which of the many DNN methods is most suitable and widely applied for predicting student performance? 3) what is the best dataset for predicting student performance? 4) how to improve the prediction accuracy by adding more dimensions of data? and 5) what are the future directions in student performance prediction research? all in all, this research was conducted to carry out a comprehensive review of the existing student performance prediction algorithms based on DL methods to help students graduate on time while simultaneously reducing the dropout rate.

2. LITERATURE REVIEW

2.1. Students' academic performance

As the volume of data in education keeps increasing, better education management is paramount. Therefore, analysing this increasingly large educational dataset is of great importance for predicting student performance [7] and improving the quality of higher education institutions. Good performing students are expected to achieve the learning outcomes set in their study programmes taken at their respective academic institutions [8]. Although it is near impossible to achieve a zero-dropout situation in developing countries [9], an attempt to reduce the dropout rate is feasible. On the other hand, the reduced dropout rates will also boost the reputation of the academic institutions [10]. If the performance of students is predicted earlier, precautionary measures can be carried out to improve the current performance and achievement of the students [11]. The educational system always needs to be consistently improved to achieve the best results and reduce the percentage of failure [12], which are essential in evaluating the quality of the graduating students [13].

2.2. ML and DL for predictive analysis

In recent years, one of the most effective ways of predicting student performance has been through machine learning mechanisms [14]. In general, research in the educational field that involves ML techniques is rapidly increasing. Applying ML techniques to the related education datasets aims to discover the hidden patterns of student performance [15]. The ML models that have been developed to solve various data science problems, including those in the education sector, have been proven very efficient and decisive over the past few years [16].

However, due to the increase in the nature and complexity of the datasets in education, student performance prediction has been designed using DL, which is found to render better performance [17], [18]. The good track records of DL-based predictive mechanisms in many areas, such as stock market analysis and intelligent transport systems, further demonstrate the advantage of employing DL in predicting student performance. Besides prediction, deep learning is also useful for detecting, classifying, and predicting over a large number of data points with better accuracy [19]. One of the main points of predicting student performance is to find and identify under-performing students as early as possible so that suitable constructive interventions can be made to help the students [17]. The deep-learning model has also been developed for predicting online performance to help at-risk students undertake courses [18].

2.3. Artificial neural network and DL techniques

It has been reported that neural network algorithms, which are closely related to DL, perform well with good detection and classification accuracy values [20]-[22] artificial neural networks (ANNs) are mainly characterised by their topologies and learning algorithms. One of the main differences between the many variants of ANN is the way the connections are made between the neurons in the network architecture. If the connections form cycles, then the ANN architecture has loops and examples of such ANN variants are the recurrent neural network and the long short-term memory architecture, which will be further explained in the following subsections. When no cycles are formed, then there are no loops in the architecture, and one of the simplest examples is the feedforward neural network, which will be briefly described next. One of the most widely used types of neural network architecture is known as the feedforward neural network (FNN), which contains no cycles or feedback loops. The simplest variant of FNN is the single-layer perceptron (SLP), which allows the input data to go through several layers before reaching the output or exit node [23].

Another variant of neural networks is called multi-layer perceptrons (MLP), which inspires another type of DL architecture known as convolutional neural networks. CNN is generally applied in image processing to learn and extract the spatial elements and structures from image datasets for learning and classification purposes. The convolutional interactions between the neighbouring neurons are essential for running the classification process [24]. In the RNN architecture, the neurons from one layer are sequentially connected to the neurons in the next layer. Therefore, the outputs of the neurons in the layer become the inputs to the neurons in the next layer. In addition, the hidden layers, which are typically present in the middle of the network, are useful in enhancing the prediction performance of the RNN architecture [25].

LSTM network is considered as a variant of RNN. One of the main features of LSTM is its ability to learn from time-series data, by identifying the data patterns which are useful for prediction purposes. Each neuron in the network is given an authority to control the incoming inputs by using special units known as gates. In this way, any possible errors from the previous neurons or layers will not be escalated to the next layer. Thus, error reduction is performed better in LSTM as only the selected input neurons are authorised to take part in generating the output [26]. In addition, LSTM architecture is also equipped with memory blocks which further enhance its ability in learning the data patterns by recording the time status during the learning process. Some unnecessary information learned during the process will be removed via a special gate called forget gate [27]. Another type of DNN is deep belief network (DBN), which is designed to work with unlabeled data, although determining the suitable structure of the network with its corresponding gradient dispersion is challenging [28]. As a DL network, DBN is typically stacked with restricted boltzmann machines (RBM) such that the hidden layer in RBM can be transformed as the visible layer between two consecutive layers. DBN has also been demonstrated to be effective in reducing the noise when working with high dimensional time series data, which is useful for predictive analysis in student performance. Some units are interconnected between the lower and the upper layers but they are not connected within the same layer [29].

2.4. Student performance prediction process

In general, three sequential stages are needed to predict students' academic performance, as depicted in Figure 1. The prerequisite before the desired predictive analysis is data preprocessing. Data integration, cleaning, discretization, and filtering operations are some of the most important preprocessing tasks [30]-[33]. Discretization is performed to convert numerical data into nominal values, such as the students' marks to grades and so forth. A number of effective data filtering methods are also feasible, such as filter-based and wrapper-based mechanisms [33], which rank the essential features and attributes and choose the minimum set of attributes required for the learning purpose.

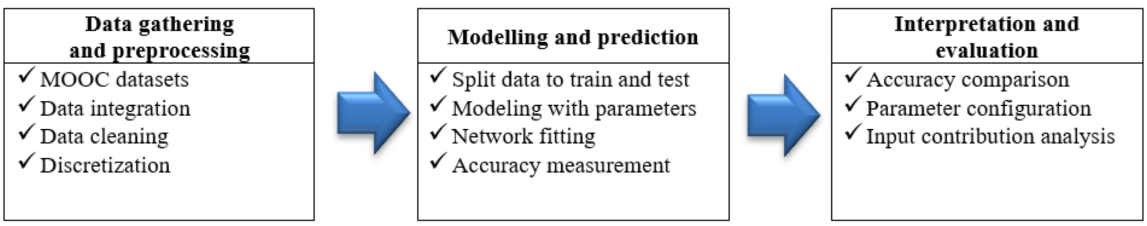

Figure 1. Students' academic performance prediction process

When the dataset is ready, modelling of the prediction algorithm can be performed [34]. Using the proposed DNN architecture, the data will be split into two for both training and testing. The accuracy of the prediction is then measured for the selected dataset in both the training and testing phases. Only the parameters that correspond to the best accuracy values will be chosen. The parameters will also include the connection weight values between the neurons in the proposed architecture [31]. Besides accuracy, other parameters may also be applied, such as recall and F1.

3. THE RESEARCH METHOD

In this paper, a systematic review is carried out via four steps, as depicted in Figure 2. The detailed explanation of each of these four steps is presented in the following subsections. To ensure that the identification operation is of good quality, several inclusion and exclusion (IE) criteria have been set to ensure that only the relevant and right papers are included. There are four criteria for each of the two strategies, as shown in Table 1.

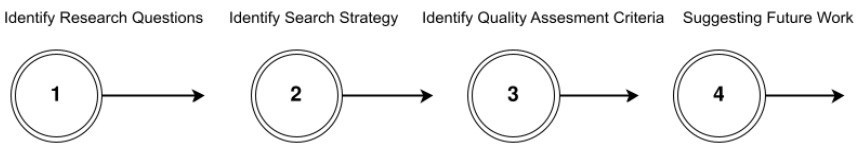

Figure 2. Systematic review methodology

Table 1. Inclusion exclusion (IE) criteria

Inclusion strategy	Exclusion strategy
- Must be for student performance prediction only	- No performance prediction for students
- Must perform prediction using neural network	- No prediction accuracy is reported
- Must specify performance measures and models	- Unpublished results in journal or conference
- Must show the summary of the applied corpus	- No neural network designs

From 202 journal papers found initially across varying reputable online journal databases, 75 of these papers were removed as they are unrelated and duplicates of the other journals. Further filtering process is carried out based on the inclusion and exclusion criteria according to Table 1 to finally select 23 papers. The details of the findings obtained, are briefly elaborated in Table 2 and Table 3. These details include the year of publication, the type of dataset source, algorithm and data dimensions. Although relatively good accuracy values have been achieved, the number of attributes and variables employed are still limited, as in [33], where only 10 variables are considered. Additional attributes such as the roles of lecturers and additional classes are not included in the datasets, hence not considered in the training and testing.

Table 2. Performance prediction for non-online/non-electronic dataset sources in the literature

First author	Year	Deep architecture	Algorithm	Dimensions	Accuracy
S. -C. Tsai [33]	2020	ANNs/DNN	Backpropagation	Spatio-online data	77%
E. T. Lau [32]	2019	ANNs/DNN	LM-backpropagation	Spatio-online data	84.8%
B. Sekeroglu [35]	2019	LSTM	Backpropagation	Temporal	87.78%
M. R. I. Rifat [36]	2019	ANNs/DNN	Backpropagation	Spatio-online data	NA
E. E. Vasileva [37]	2019	ANNs/DNN	Backpropagation	Temporal	NA
J. Sultana [38]	2018	MLP	Backpropagation	Spatio-online data	78.75%
R. Deuja [39]	2018	MLP	Backpropagation	Spatio-online data	97.12%
W. W. T. Fok [40]	2018	CNN	Backpropagation	Temporal	91%
A. Nurhuda [41]	2017	ANNs/DNN	Backpropagation	Temporal	79.87%
M. F. Sikder [42]	2016	ANNs/DNN	LM-backpropagation	Temporal	96.57 %

Using the selected and filtered papers tabulated in Table 2 as well as Table 3, which tabulates the comparisons between different mechanisms run in massive open online courses (MOOC) and LMS dataset sources, a set of research questions (RQ) have been designed to carry out the systematic review of these papers, as given below.

- (RQ1) what is the best DNN architecture for predicting student performance?
- (RQ2) which of the DNN methods are widely applied for predicting student performance?
- (RQ3) which is the best dataset mostly applied in predicting student performance?
- (RQ4) how to improve the prediction accuracy by adding more dimensions of data?

– (RQ5) what are the current trends and directions in the research community related to student performance prediction?

In the next section, these questions will be answered and discussed thoroughly.

Table 3. Performance prediction for MOOC/LMS dataset sources in the literature

First author	Year	Deep architecture	Algorithm	Dimensions	Accuracy
Ş. Aydoğdu [43]	2020	ANNs/DNN	Backpropagation	Spatio-online data	80.47%
H. Waheed [44]	2020	ANNs/DNN	Backpropagation	Spatio-online data	88.62%
L. Qiu [45]	2019	CNN	Backpropagation	Spatio-online data	90.72%
A. S. Imran [46]	2019	ANNs/DNN	Backpropagation	Spatio-online data	98.65%
S. Altaf [47]	2019	Multi-layer feed forward neural network (MLFFNN)	Backpropagation	Spatio-online data	97.4%
R. C. Raga [17]	2019	ANNs/DNN	Backpropagation	Spatiotemporal	91.07%
Y. S. Alsalman [48]	2019	ANNs/DNN	Backpropagation	Spatio-online data	97%
X. Ma [49]	2018	ANNs/DNN	Filter-type feature selection	Spatio-online data	90%
F. Okubo [50]	2017	RNN	Backpropagation	Spatio-online data	90%
T. -Y. Yang [51]	2017	Feedback time series neural network (FTSNN)	Backpropagation	Spatio-online data	60%
S. Chaudhary [23]	2017	FNN	Backpropagation	Spatio-online data	91.5%

4. RESULTS AND ANALYSIS

4.1. Results of the systematic literature review

Based on the research questions set in the previous section, a rigorous study has been carried out to answer these questions which revolve around the study of DNN and its implementation on predicting student performance. The research papers from various established journals have been studied and analysed to solve these questions in order to complete the systematic literature review. The findings of the systematic literature review conducted are described as follows.

4.1.1. RQ1: what is the best DNN architecture for predicting student performance

Based on the systematic review carried out, the best deep neural network architecture applied in predicting student performance is either ANN or DNN, as depicted in Figure 3. The least applied deep neural network method is RNN, which has only been proposed in one journal paper. On the other hand, there are ten papers presented and proposed either RNN or DNN [32], [36], [41], [52], [43], [44], [46], [48], for classifying, modeling, and predicting student performance. This trend is mainly due to the higher prediction accuracy that can be achieved by ANN or DNN as compared to the other methods.

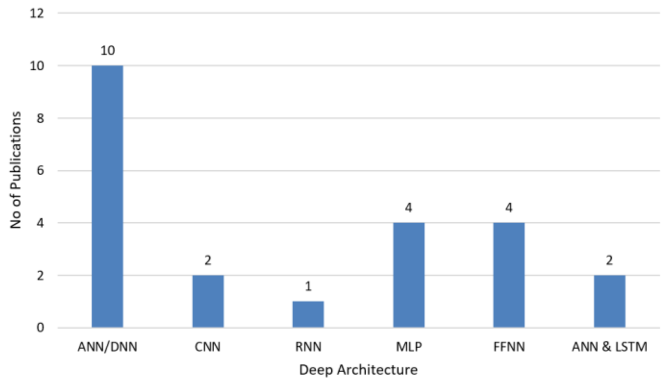

Figure 3. DNN architectures for student performance prediction

Suleiman *et al.* [30], CNN architecture is reported to be able to perform dropout prediction for MOOC based on clickstream data on student learning behavior. In [48], it is reported that DL is effective in predicting student performance. The RNN architecture is presented in [33], for early prediction of the final grades compared with other regression methods. In [34], [37], [53], [50], MLP architecture is compared with logistic regression techniques as well as other ML methods where MLP was observed to be more effective for prediction operations. As for the FFNN methods proposed in [36], [38], [49], [51], excellent prediction accuracy values are also reported. The LSTM architecture, as proposed in [32], demonstrates good performance against other machine learning methods such as the support vector machine (SVM).

4.1.2. RQ2: which of the DNN methods are widely applied for predicting student performance?

The deep-learning method that is widely used in architecture testing is backpropagation (BP), as seen in Figure 4 since there are up to 20 papers presented based on this method. Other than the backpropagation algorithms, Levenberg Marquardt (LM) was found to be applied to the E. T. Lau model [32], and a filter-type feature selection method was applied to the Xiaofeng Ma model [49]. Backpropagation is used to train a multi-layered neural network to learn the proper internal representations that will allow it to learn any arbitrary input-to-output mapping.

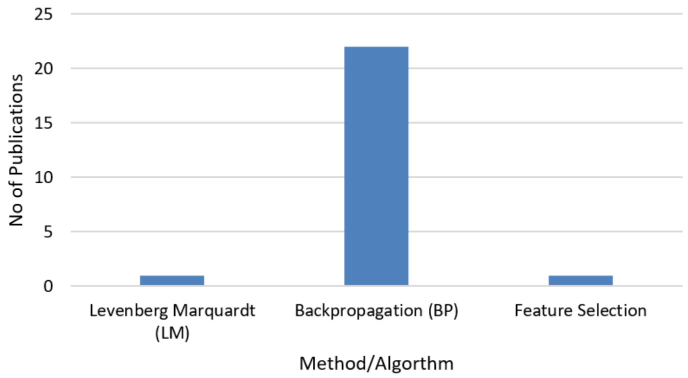

Figure 4. Publications deep-learning method

4.1.3. RQ3: which is the best dataset mostly applied in predicting student performance?

The dataset used to predict student academic performance consists of public and private datasets. Public data sets that can be accessed include the University of California (UC) irvine machine learning repository used in the Yasmeen Shaher Alsalman model [48], Somendra Chaudhary [23], the open university learning analytics (OULA) dataset used in the Hajra Waheed model [44], as seen in Figure 5. However, most of the datasets applied are of the spatio-online type, and only five papers are found to have presented findings based on temporal or spatiotemporal datasets. In the current pandemic situation as well as in the online learning environment, temporal data will be more relevant to be studied as the pandemic situation is found everywhere around the world while online learning is becoming the new norm.

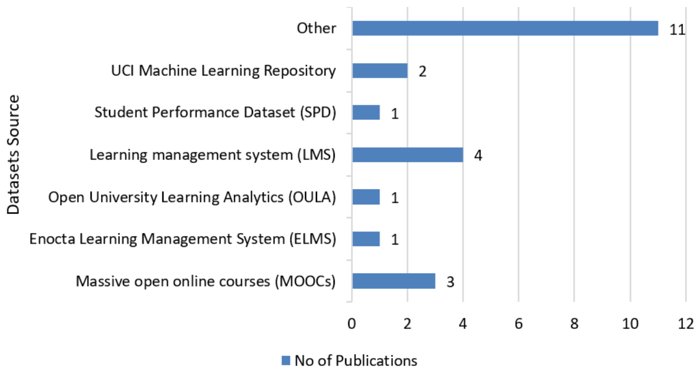

Figure 5. Datasets source with their corresponding numbers of publications

4.1.4. RQ4: how to improve the prediction accuracy by adding more dimensions of data?

Tsai *et al.* [33], there is a clear direction of artificial intelligence (AI) methods, especially deep learning (DL), for research in higher education. However, the number of variables used in the research is only ten, which is relatively low. Ma *et al.* [49], a different model is proposed and run using a dataset integrated from 39 courses. The results reported are better than the baseline method. Okubo *et al.* [50], two different datasets are applied but with a considerably high imbalance ratio. As students' datasets are typically time-series, some models allow dataset updates to reflect the change over the time [46]. To overcome imbalance ratios, further data cleaning might be required such as removing some discrete features or patterns in the datasets [35], [53].

4.1.5. RQ5: what are the current trends and directions in the research community related to student performance prediction?

To enhance the accuracy achievement, more data attributes should be considered, such as ages, family factors, learning styles, course feedback, additional-curricular classes, demographic factors, and so forth [33]. In MOOC situations, dropout prediction may also be improved via additional attributes such as assignment submission, interaction information, and course forums [49]. The number of neurons and corresponding layers should also be adjusted to enhance the accuracy [46]. In future, the impacts of all student activities on their study performance should also be studied, such as those available in the OULA dataset [44]. Combining data from different departments or schools with more students is also possible, as in the case of the cherwell service management (CSM) database [35]. Combining data from different departments or schools with more students is also possible as in the case of CSM database [48]. Furthermore, data from different universities can be combined to provide a more complete picture of performance prediction in general [19]. It can also be observed from our findings that most of the research projects in this area focus on spatio-online data. There are only a few projects that have been carried out based on temporal and spatio-temporal data, which are more challenging and relevant these days as the time factor is essential in creating the dataset. Some of the DL methods proposed to work with temporal and spatio-temporal data are the ordinary ANN and the LSTM methods, which will be further studied as presented next.

4.2. A proposed framework to student' performance prediction

As the temporal and spatiotemporal datasets are found to be more relevant in the current online learning trend across the globe, the preferable predictive model to predict the student's performance in our project will be based on an improved LSTM algorithm. The proposed improvement will involve optimising algorithms that are adam and nadam algorithms to improve the learning process and the achievable performance. Deep LSTM is able to model complex nonlinear relationships over a relatively long period of time. So this is very appropriate for solving time-series types of datasets, such as temporal and spatiotemporal datasets. The proposed framework is given in Figure 6.

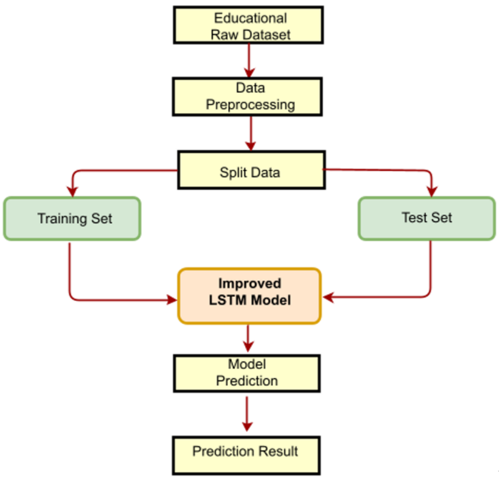

Figure 6. A proposed framework to student performance prediction

LSTM is considered among the most popular deep learning models used today. It is also being applied to time series prediction, which is a particularly hard problem to solve due to the presence of long term trends, seasonal and cyclical fluctuations, and random noise. The performance of LSTM is highly dependent on the choice of several hyper-parameters, which need to be chosen very carefully, in order to get good results. Being a relatively new model, there are no established guidelines for configuring LSTM. The influence of feature selection is significant on the prediction accuracy of LSTM models. Therefore, further significant improvements will be carried out to cater to these problems and design an improved LSTM model for implementing the prediction operation.

5. CONCLUSION

The systematic study carried out in this paper shows that deep learning is gaining interest among researchers for predicting students' academic performance. Apart from the increasingly large volumes of required data, the increase in the data attributes as well as the complex nature of online courses, including MOOC have further contributed to the need for deep learning in addressing this problem. Different variants of neural network architectures have also been developed and applied for predicting performance, and the deep neural network architecture has been reported to produce good prediction accuracy values, especially with more data attributes in larger datasets. From the systematic review carried out in this paper, it can be observed that the most widely used DL method for predicting the performance of students is ANN. For temporal datasets, which are very relevant for the student performance datasets which are generally temporal in nature, the LTSM method has been reported to produce promising accuracy values, which are more than 85%. Exploring and developing deep learning approaches to predict students' academic performance is thus an important and relevant research problem that should be addressed in the future.

REFERENCES

[1] L. W. Santoso and Yulia, "Predicting student performance in higher education using multi-regression models," *Telkomnika (Telecommunication Computing Electronics and Control)*, vol. 18, no. 3, pp. 1354-1360, 2020, doi: 10.12928/TELKOMNIKA.v18i3.14802.

[2] S. T. Ahmed, R. Al-Hamdani, and M. S. Croock, "Enhancement of student performance prediction using modified K-nearest neighbor," *Telkomnika (Telecommunication Computing Electronics and Control)*, vol. 18, no. 4, pp. 1777-1783, 2020, doi: 10.12928/TELKOMNIKA.V18I4.13849.

[3] F. Zulkifli, Z. Mohamed, and N. A. Azmee, "Systematic research on predictive models on students' academic performance in higher education," *International Journal of Recent Technology and Engineering*, vol. 8, no. 2, pp. 357-363, 2019, doi: 10.35940/ijrte.B1061.0782S319.

[4] X. Li, X. Zhu, X. Zhu, Y. Ji, and X. Tang, "Student Academic Performance Prediction Using Deep Multi-source Behavior Sequential Network," *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, pp. 567-579, 2020. [Online]. Available: https://link.springer.com/content/pdf/10.1007/978-3-030-47426-3_44.pdf

[5] J. L. Rastrollo-Guerrero, J. A. Gómez-Pulido, and A. Durán-Domínguez, "Analyzing and predicting students' performance by means of machine learning: A review," *Applied Sciences*, vol. 10, no. 3, 2020, doi: 10.3390/app10031042.

[6] S. A. Salloum, M. Alshurideh, A. Elnagar, and K. Shaalan, "Machine Learning and Deep Learning Techniques for Cybersecurity: A Review," *Proceedings of the International Conference on Artificial Intelligence and Computer Vision (AICV2020)*, 2020, pp. 50-57, doi: 10.1007/978-3-030-44289-7_5.

[7] M. Imran, S. Latif, D. Mehmood, and M. S. Shah, "Student academic performance prediction using supervised learning techniques," *International Journal of Emerging Technologies in Learning*, vol. 14, no. 14, pp. 92-104, 2019, doi: 10.3991/ijet.v14i14.10310.

[8] M. Asiah, K. N. Zulkarnaen, D. Safaai, M. Y. N. N. Hafzan, M. M. Saberi, and S. S. Syuhaida, "A Review on Predictive Modeling Technique for Student Academic Performance Monitoring," *MATEC Web of Conferences, Engineering Application of Artificial Intelligence Conference 2018 (EAAIC 2018)*, 2019, vol. 255, no. 2, doi: 10.1051/matecconf/201925503004.

[9] N. Mduma, K. Kalegele, and D. Machuve, "A survey of machine learning approaches and techniques for student dropout prediction," *Data Science Journal*, vol. 18, no. 1, pp. 1-10, 2019, doi: 10.5334/dsj-2019-014.

[10] N. M. Suhaimi, S. Abdul-Rahman, S. Mutalib, N. H. A. Hamid, and A. M. A. Malik, "Review on Predicting Students' Graduation Time Using Machine Learning Algorithms," *International Journal of Modern Education and Computer Science*, vol. 11, no. 7, pp. 1-13, 2019, doi: 10.5815/ijmecs.2019.07.01.

[11] J. Sultana, M. U. Rani, and M. A. H. Farquad, "Student's performance prediction using deep learning and data mining methods," *International Journal of Recent Technology and Engineering*, vol. 8, no. 1, pp. 1018-1021, 2019. [Online]. Available: https://www.researchgate.net/profile/J-Sultana/publication/335234927_Student's_Performance_Prediction_using_Deep_Learning_and_Data_Mining_methods/links/5d5a6b3e92851c3763694c8f/Students-Performance-Prediction-using-Deep-Learning-and-Data-Mining-methods.pdf

[12] M. Karlik and B. Karlik "Prediction of Student's Performance with Deep Neural Networks," *International Journal of Artificial Intelligence and Expert Systems (IJAE)*, vol. 9, no. 2, pp. 39-47, 2020, [Online]. Available: https://www.cscjournals.org/manuscript/Journals/IJAE/Volume9/Issue2/IJAE-197.pdf

[13] B. K. Francis and S. S. Babu, "Predicting Academic Performance of Students Using a Hybrid Data Mining Approach," *Journal of Medical Systems*, vol. 43, 2019, doi: 10.1007/s10916-019-1295-4.

[14] V. Vijayalakshmi and K. Venkatachalapathy, "Comparison of Predicting Student's Performance using Machine Learning Algorithms," *International Journal of Intelligent Systems and Applications*, vol. 11, no. 12, pp. 34-45, 2019, doi: 10.5815/ijisa.2019.12.04.

[15] B. Mounika and V. Persis, "A Comparative Study of Machine Learning Algorithms for Student Academic Performance," *International Journal of Computer Sciences and Engineering*, vol. 7, no. 4, pp. 721-725, 2019, doi: 10.26438/ijcse/v7i4.721725.

[16] A. O. Oyedeji, A. M. Salami, O. Folorunsho, and O. R. Abolade, "Analysis and Prediction of Student Academic Performance Using Machine Learning," *JITCE (Journal of Information Technology and Computer Engineering)*, vol. 4, no. 01, pp. 10-15, 2020. [Online]. Available: http://jitce.fti.unand.ac.id/index.php/JITCE/article/view/51/37.

[17] R. C. Raga and J. D. Raga, "Early Prediction of Student Performance in Blended Learning Courses Using Deep Neural Networks," *2019 International Symposium on Educational Technology (ISET)*, 2019, pp. 39-43, doi: 10.1109/ISET.2019.00018.

[18] H. Karimi, J. Huang, and T. Derr, "A Deep Model for Predicting Online Course Performance," *Association for the Advancement of Artificial Intelligence*, 2020, [Online]. Available http://www.cse.msu.edu/~wangzh65/AI4EDU/papers/19.pdf

[19] A. Yunita, H. B. Santoso, and Z. A. Hasibuan, "Deep Learning for Predicting Students' Academic Performance," *2019 Fourth International Conference on Informatics and Computing (ICIC)*, 2019, pp. 1-6, doi: 10.1109/ICIC47613.2019.8985721.

[20] S. A. Salloum, R. Khan, and K. Shaalan, "A Survey of Semantic Analysis Approaches," *Proceedings of the International Conference on Artificial Intelligence and Computer Vision (AICV2020)*, 2020, pp. 61-70, doi: 10.1007/978-3-030-44289-7_6.

[21] S. F. S. Alhashmi, M. Alshurideh, B. A. Kurdi, and S. A. Salloum, "A Systematic Review of the Factors Affecting the Artificial Intelligence Implementation in the Health Care Sector," *Proceedings of the International Conference on Artificial Intelligence and Computer Vision (AICV2020)*, 2020, pp. 37-49, doi: 10.1007/978-3-030-44289-7_4.

[22] S. A. Salloum, M. Alshurideh, A. Elnagar, and K. Shaalan, "Mining in Educational Data: Review and Future Directions," *Proceedings of the International Conference on Artificial Intelligence and Computer Vision (AICV2020)*, 2020, pp. 92102, doi: 10.1007/978-3-030-44289-7_9.

[23] S. Chaudhary, A. Imran, and S. V. Kolekar, "Prediction of academic performance using gravitational search based neural network algorithm," *2017 International Conference on Inventive Computing and Informatics (ICICI)*, 2017, pp. 388-393, doi: 10.1109/ICICI.2017.8365379.
[24] Z. Fang, H. Feng, S. Huang, and D. X. Zhou, "Theory of deep convolutional neural networks II: Spherical analysis," *Neural Networks*, vol. 131, pp. 154-162, Nov. 2020, doi: 10.1016/j.neunet.2020.07.029.
[25] A. Mondal and J. Mukherjee, "An Approach to Predict a Student's Academic Performance using Recurrent Neural Network (RNN)," *International Journal of Computer Applications*, vol. 181, no. 6, pp. 1-5, Jul. 2018, doi: 10.5120/ijca2018917352.
[26] J. Yuan, M. Abdel-Aty, Y. Gong, and Q. Cai, "Real-Time Crash Risk Prediction using Long Short-Term Memory Recurrent Neural Network," *Transportation Research Record*, vol. 2673, no. 1, pp. 314–326, Apr. 2019, doi: 10.1177/0361198119840611.
[27] Y. Zhang, R. Xiong, H. He, and M. G. Pecht, "Long Short-Term Memory Recurrent Neural Network for Remaining Useful Life Prediction of Lithium-Ion Batteries," in *IEEE Transactions on Vehicular Technology*, vol. 67, no. 7, pp. 5695-5705, July 2018, doi: 10.1109/TVT.2018.2805189.
[28] Y. Zhang and F. Liu, "An improved deep belief network prediction model based on knowledge transfer," *Future Internet*, vol. 12, no. 11, pp. 1-18, Oct. 2020, doi: 10.3390/fi12110188.
[29] H. Wei, C. Shan, C. Hu, Y. Zhang, and X. Yu, "Software defect prediction via deep belief network," *Chinese Journal of Electronics*, vol. 28, no. 5, pp. 925-932, Sep. 2019, doi: 10.1049/cje.2019.06.012.
[30] S. Suleiman, A. Lawal, U. Usman, S. U. Gulumbe, and A. B. Muhammad, "Student's Academic Performance Prediction Using Factor Analysis Based Neural Network," *International Journal of Data Science and Analysis*, vol. 5, no. 4, Jan. 2019, doi: 10.11648/j.ijdsa.20190504.12.
[31] A. Y. Q. Huang, O. H. T. Lu, J. C. H. Huang, C. J. Yin, and S. J. H. Yang, "Predicting students' academic performance by using educational big data and learning analytics: evaluation of classification methods and learning logs," *Interactive Learning Environments*, vol. 28, no. 2, pp. 206-230, Feb. 2020, doi: 10.1080/10494820.2019.1636086.
[32] E. T. Lau, L. Sun and Q. Yang, "Modelling, prediction and classification of student academic performance using artificial neural networks," *SN Applied Sciences*, vol. 1, no. 9, Aug. 2019, doi: 10.1007/s42452-019-0884-7.
[33] S. -C. Tsai, C. -H. Chen, Y. -T. Shiao, J. -S. Ciou, and T. -N. Wu, "Precision education with statistical learning and deep learning: a case study in Taiwan," *International Journal of Educational Technology in Higher Education*, vol. 17, no. 12, Apr. 2020, doi: 10.1186/s41239-020-00186-2.
[34] D. Diouf, A. Niang, and S. Thiria, "Deep Learning Based Multiple Regression to Predict Total Column Water Vapor (Tcwv) From Physical Parameters in West Africa By Using Keras Library," *International Journal of Data Mining & Knowledge Management Process (IJDKP)*, vol. 9, no. 6, pp. 13-21, 2019, doi: 10.5121/ijdkp.2019.9602.
[35] B. Sekeroglu, K. Dimililer, and K. Tuncal, "Student Performance Prediction and Classification Using Machine Learning Algorithms," in *Proceedings of the 2019 8th International Conference on Educational and Information Technology*, 2019, pp. 7-11. doi: 10.1145/3318396.3318419.
[36] M. R. Islam Rifat, A. Al Imran, and A. S. M. Badrudduza, "EduNet: A Deep Neural Network Approach for Predicting CGPA of Undergraduate Students," 2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT), 2019, pp. 1-6, doi: 10.1109/ICASERT.2019.8934616.
[37] E. E. Vasileva, D. S. Kurushin, and S. S. Vlasov, "Early Prediction of the Grade Point Average of University Students Diploma: Neural Network Approach," *2019 XXII International Conference on Soft Computing and Measurements (SCM))*, 2019, pp. 259-262, doi: 10.1109/SCM.2019.8903629.
[38] J. Sultana, N. Sultana, K. Yadav, and F. AlFayez, "Prediction of Sentiment Analysis on Educational Data based on Deep Learning Approach," *2018 21st Saudi Computer Society National Computer Conference (NCC)*, 2018, pp. 1-5, doi: 10.1109/NCG.2018.8593108.
[39] R. Deuja, R. Karna, and R. Kusatha, "Data-Driven Predictive Analysis of Student Performance In College Using Neural Networks," *2018 IEEE 3rd International Conference on Computing, Communication and Security (ICCCS)*, 2018, pp. 77-81, doi: 10.1109/CCCS.2018.8586809.
[40] W. W. T. Fok *et al*., "Prediction model for students' future development by deep learning and tensorflow artificial intelligence engine," *2018 4th International Conference on Information Management (ICIM)*, 2018, pp. 103-106, doi: 10.1109/INFOMAN.2018.8392818.
[41] A. Nurhuda and D. Rosita, "Prediction Student Graduation on Time Using Artificial Neural Network on Data Mining Students STMIK Widya Cipta Dharma Samarinda," in *Proceedings of the 2017 International Conference on E-Commerce, E-Business and E-Government*, 2017, pp. 86–89. doi: 10.1145/3108421.3108431.
[42] M. F. Sikder, M. J. Uddin, and S. Halder, "Predicting students yearly performance using neural network: A case study of BSMRSTU," *2016 5th International Conference on Informatics, Electronics and Vision (ICIEV)*, 2016, pp. 524-529, doi: 10.1109/ICIEV.2016.7760058.
[43] Ş. Aydoğdu, "Predicting student final performance using artificial neural networks in online learning environments," *Education and Information Technologies*, vol. 25, no. 3, pp. 1913-1927, 2020, doi: 10.1007/s10639-019-10053-x.
[44] H. Waheed, S. U. Hassan, N. R. Aljohani, J. Hardman, S. Alelyani, and R. Nawaz, "Predicting academic performance of students from VLE big data using deep learning models," *Computers in Human Behavior*, vol. 104, 2020, doi: 10.1016/j.chb.2019.106189.
[45] L. Qiu, Y. Liu, Q. Hu, and Y. Liu, "Student dropout prediction in massive open online courses by convolutional neural networks," *Soft Computing*, vol. 23, no. 1, pp. 10287-10301, 2019, doi: 10.1007/s00500-018-3581-3.
[46] A. S. Imran, F. Dalipi, and Z. Kastrati, "Predicting Student Dropout in a MOOC: An Evaluation of a Deep Neural Network Model," in *Proceedings of the 2019 5th International Conference on Computing and Artificial Intelligence*, 2019, pp. 190-195, doi: 10.1145/3330482.3330514.
[47] S. Altaf, W. Soomro, and M. I. M. Rawi, "Student Performance Prediction using Multi-Layers Artificial Neural Networks: A case study on educational data mining," *ACM International Conference Proceeding Series*, 2019, pp. 59-64, doi: 10.1145/3325917.3325919.
[48] Y. S. Alsalman, N. Khamees Abu Halemah, E. S. AlNagi, and W. Salameh, "Using Decision Tree and Artificial Neural Network to Predict Students Academic Performance," *2019 10th International Conference on Information and Communication Systems (ICICS)*, 2019, pp. 104-109, doi: 10.1109/IACS.2019.8809106.
[49] X. Ma, Y. Yang, and Z. Zhou, "Using Machine Learning Algorithm to Predict Student Pass Rates in Online Education," in *Proceedings of the 3rd International Conference on Multimedia Systems and Signal Processing*, 2018, pp. 156-161, doi: 10.1145/3220162.3220188.

[50] F. Okubo, T. Yamashita, A. Shimada, and H. Ogata, "A Neural Network Approach for Students' Performance Prediction," in *Proceedings of the Seventh International Learning Analytics & Knowledge Conference*, 2017, pp. 598-599, doi: 10.1145/3027385.3029479.

[51] T. -Y. Yang, C. G. Brinton, C. Joe-Wong, and M. Chiang, "Behavior-Based Grade Prediction for MOOCs Via Time Series Neural Networks," in *IEEE Journal of Selected Topics in Signal Processing*, vol. 11, no. 5, pp. 716-728, Aug. 2017, doi: 10.1109/JSTSP.2017.2700227.

[52] G. Xu and W. Fang, "Shape retrieval using deep autoencoder learning representation," *2016 13th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP)*, 2016, pp. 227-230, doi: 10.1109/ICCWAMTIP.2016.8079843.

[53] N. Mohd and Y. Yahya, "A Data Mining Approach for Prediction of Students," Depression Using Logistic Regression and Artificial Neural Network," *Proceedings of the 12th International Conference on Ubiquitous Information Management and Communication*, 2018, pp. 1-5, doi: 10.1145/3164541.3164604.

BIOGRAPHIES OF AUTHORS

Edi Ismanto completed education bachelor's degree in the Informatics Engineering Department, State Islamic University of Sultan Syarif Kasim Riau. And master's degree in Master of Computer Science at Putra Indonesia University Padang. Now working as a lecturer in the Department of Informatics, University Muhammadiyah of Riau. With research interests in the field of Machine learning algorithms and AI. He can be contacted at email: edi.ismanto@umri.ac.id.

Hadhrami Ab Ghani received his bachelor degree in electronics engineering from Multimedia University Malaysia (MMU) in 2002. In 2004, he completed his masters degree in Telecommunication Engineering at The University of Melbourne. He then pursued his Ph.D. at Imperial College London in intelligent network systems and completed his Ph.D. in 2011. He can be contacted at email: hadhrami.ag@umk.edu.my.

Nurul Izrin Md Saleh obtained a bachelor's degree in information technology from Multimedia University Malaysia (MMU). He completed his master's degree in computer science at The University of Putra Malaysia. Then complete a Ph.D. at The University of Brunel London in the same field of study. She can be contacted at email: nurulizrin.ms@umk.edu.my.

Januar Al Amien completed education bachelor's degree in the Informatics Engineering Department, STMIK-AMIK Riau. And master's degree in Master of Information Technology at Putra Indonesia University Padang. Now working as a lecturer in the Department of Computer Science, University Muhammadiyah of Riau. With research interests in the field of Machine learning algorithms and AI. He can be contacted at email: januaralamien@umri.ac.id.

Rahmad Gunawan graduated with a bachelor's degree at Gunadarma University with a major in Information Management, a master's degree with Gunadarma University majoring in Electrical Telecommunication. And now works as a lecturer at the Faculty of Computer Science, University of Muhammadiyah Riau. With research interests in the field of Machine learning algorithms and AI. He can be contacted at email: goengoen78@umri.ac.id.