PREDICTING STUDENTS' ACADEMIC PROGRESSION: A BAYESIAN NEURAL NETWORK APPROACH

LIEN NGUYEN[1], JACK JIA[2], ANH NGUYEN[3], STEVE LING[4], NIGEL FINCH[1]

[1]Centre for Teaching, Research and Scholarship, Sydney Institute of Higher Education, Sydney, Australia
[2] School of Business, Charles Sturt University-Sydney Campus, Sydney, Australia
[3] School of Electrical Engineering, Hanoi University of Science and Technology, Hanoi, Vietnam
[4] School of Electrical and Data Engineering, University of Technology Sydney, Sydney, Australia
E-MAIL: lien.nguyen@sydneyinstitute.edu.au, zjia@csu.edu.au, anh.nguyenvan1@hust.edu.vn, Steve.Ling@uts.edu.au, Nigel.Finch@sydneyinstitute.edu.au

Abstract:

The prediction of students' academic progression and early identification of at-risk students have been considered as essential factors contributing to the improvement of student retention rates for higher education institutions in Australia. This study is concerned with the development of Bayesian neural networks (BNNs) for predicting students' academic progression in the first semester of their studies. By using data extracted from 353 international undergraduate students who enrolled in full-time studying load at Sydney Institute of Higher Education (SIHE), an artificial neural network was developed and trained by the Bayesian inference framework to explore the ability of BNNs in the application of predicting students' academic progression. With the classification results of 85% Sensitivity and 70% Specificity on the training set, and 81% Sensitivity and 74% Specificity on the testing test, it is indicated that the developed prediction model based on BNNs can successfully predict the student's progression. Additionally, it is demonstrated in this study that the better generalizability and the ability to quantify uncertainty of predictions tend to be the key advantages of BNNs over standard neural networks in the application of predicting students' academic progression.

Keywords:

Bayesian neural network, Machine Learning, higher education, student's academic progression, At-Risk students

1. Introduction

The rates of students' academic progression, retention and completion of their educational programs have been regarded as critical factors which reflect not only the education quality of institutions but also significantly affect their reputation and financial sustainability. Student attrition (drop-out rate) has been recognized as a challenging risk for universities and higher education institutions, especially for private higher education providers. In Australia, data from 2005 to 2014 show that the attrition rate for public universities is around 15% [1] [2]. This rate has shown a gradual increase over the years, with reported median attrition rates in 2021 of 20.17% for Australian universities and 27.42% for non-university higher-education providers [3]. This risk has been becoming more prolonging and causing more profound impact since the massive offering of the online learning mode after COVID-19 pandemic.

International students who commence their tertiary education tend to experience a higher risk of making unsatisfactory progression due to various transition challenges. These include transitions from secondary to tertiary studies, transitions to a new cultural and academic environment as well as language barriers. Resultantly, monitoring students' academic progression and early identifying students who are at risk of making unsatisfactory progression in their first semester of enrolment has been acknowledged as one of critical student retention strategies for all higher education institutions. An early prediction of students' academic performance plays an important role in helping institutions to provide students with early intervention strategies including academic support and counselling to enhance students' academic success rate. Although various frameworks for identifying at-risk students at early stages have been generally developed and implemented, traditional approaches appear less effective due to ongoing challenges. These encompass the complexity of identifying indicators, the administrative burden of data processing and analysis, as well as inconsistencies and inaccuracies occurred in data interpretation. Recently, the emerging use of Machine Learning in the higher education sector has been explored with the aim of developing more precise, consistent and logical algorithms that can effectively predict students' academic progression [4] [5].

Artificial neural networks have been widely recognized as a powerful tool of classification and recognition which can successfully model non-linear correlations between inputs and outputs and adapt itself to new patterns to provide effective solutions to predicting or forecasting problems in various disciplines. The use of artificial neural networks in predicting students' academic progression has recently attracted a lot of research [6] [7]. Different neural network models with different training techniques (both deterministic and probabilistic) have been introduced to enhance the performance of prediction. The Bayesian inference framework has been widely applied in training neural networks because of its capabilities to enhance model regularization, particularly in applications with limited data like medical diagnosis [8] [9]. In addition, the ability to quantify uncertainties associated with predictions features another important factor that makes Bayesian neural networks become popular in discipline areas such as healthcare, finance or education.

This paper aims to introduce an innovative method of predicting students' academic performance in their first semester by using artificial neural networks. In this study, the Bayesian inference framework will be applied to train the developed neural networks. This paper consists of four sections. Section 1 provides an introduction and the context of the study. Section 2 discusses the methodology used in our study. Results of the study will be analyzed and discussed in Section 3. Section 4 draws conclusions for the current study and provides some recommendations for future work.

2. Methods

2.1. Data Preparation and Data Analysis

In this study, the dataset is extracted from international undergraduate students at Sydney Institute of Higher Education (SIHE). To ensure the generalizability of the study, students from different cohorts who commenced their studies from March 2023 to March 2024 at SIHE are included. All students selected in this study were enrolled on the full-time studying load, which requires four subjects of study in their first semester. Students' personal details are de-identified, and all data are collected, processed and analyzed anonymously to ensure research ethics and data privacy.

First, a step of data pre-processing is conducted to remove duplicates, outliers and defective data points. As a result, a cleaned dataset of 353 students (corresponding with 353 data points) is extracted and used as the final dataset for the current study. Each data point consists of eight demographic and academic parameters as defined in Table 1.

TABLE 1. Definitions of Data Parameters

	Parameter Name	Data Type
Parameter 1	Gender	Categorical
Parameter 2	Nationality	Categorical
Parameter 3	Age at Commencement	Numerical
Parameter 4	Grade 12 Results	Numerical
Parameter 5	English Entrance Test Results	Numerical
Parameter 6	Enrolled Program	Categorical
Parameter 7	Attendance Percentage in the first 4 weeks	Numerical
Parameter 8	Academic Performance at the end of 1st semester	Categorical

For the classification purpose, the dataset will be processed so that data types are suitable to be fed into the developed prediction model. To do this, all categorical parameters are quantified into numerical data. Details of this data pre-processing steps are presented as follows:

- Parameter 1: Students' genders which can take two categorical values of Female and Male are converted into 0 (Female) and 1 (Male).
- Parameter 2: Students' nationalities can take various values due to the diversity of students' countries of origin. However, based on the fact that the majority of international students at SIHE are from subcontinental countries, the nationalities of students are converted into: 3 (Indian), 2 (Pakistani), 1 (Nepalese) and 0 (all other nationalities).
- Parameter 3: The students' age at commencement is calculated as the difference between the date they commenced their enrolled program and their date of birth. This features a new demographic attribute which has been added into the pool of parameters compared to our previous study [10].
- Parameter 4: To eliminate the differences in grading systems between countries, students' academic results in their Grade 12 which is used for admission are converted into equivalent percentages.
- Parameter 5: As entry requirements, several types of English language tests are recognized for admission at SIHE. Thus, for consistency, all students' English language results are converted into the equivalent IELTS scores.
- Parameter 6: At SIHE, there are two undergraduate programs that have been offered since 2022, namely Bachelor of Business (BBUS) and Bachelor of IT (BIT). Accordingly, the parameter of Enrolled Program can take two values which are converted into 0 (BBUS) and 1 (BIT).
- Parameter 7: Attendance is a common indicator for identifying at-risk students. In convention, students with lower attendance rate flag a high chance of failing to

meet satisfactory academic progression. At SIHE, the average attendance rate across all enrolled subjects in the first four weeks of studying is evaluated at the end of week 4 to identify students who are at risk of making unsatisfactory progression in their first semester.

- Parameter 8: The academic performance of each student is assessed by using the final results of their first semester's completion. As aforementioned, the data used in this study are collected from students with 4 subjects of enrolment. Thus, the performance of each student is categorized as Satisfactory (students who pass 3 or 4 subjects out of 4 enrolled subjects in their first semester) and Unsatisfactory (students who fail 2 or more out of 4 enrolled subjects in their first semester). As a result, the dataset is divided into two groups: Satisfactory which consists of 118 data points and Unsatisfactory which consists of 235 data points.

Following the step of cleaning and pre-processing data, different statistical techniques including descriptive analysis and inferential analysis will be applied to establish the parameters which significantly differentiate the two groups of Satisfactory and Unsatisfactory.

2.2. Classification using Bayesian Neural Networks

Artificial neural networks have been widely used in the area of predicting students' academic performance. Various training methods have been applied to enhance the performance of the networks. In our previous study, a combination of Genetic Algorithm and Levenberg-Marquardt algorithm was employed to train the developed neural network which led to classification results of 85% Sensitivity and 69% Specificity on a training set of 200 data points and 82% Sensitivity and 66% Specificity on a separated testing set of 100 data points [7]. Based on these results, it was established that the developed neural network can effectively identify at-risk students at an early stage. It is noted that with the proposed method, to avoid the inherent overfitting problem in training neural networks, a cross-validation technique was applied, in which the training set is divided into two subsets of training and validation. Despite the results indicated that the developed algorithm had a good generalizability when applying to unseen testing data, we aim to explore new training techniques with better robustness to overfitting without having to spend data on a separate validation dataset.

Bayesian neural network (BNN) has been known as an efficient machine learning technique which integrates the Bayesian inference framework in training artificial neural networks [9]. This probabilistic approach not only enhances the generalizability of the networks to unseen data but also enables the quantification of uncertainty in predictions. The ability to quantify predictions' uncertainties plays an important role in human-related classification and prediction tasks as it reduces the inconvenience and frustration of being wrongly predicted or diagnosed, avoiding misdirected actions as well as improving decision making. To do this, in BNNs, the network's parameters (weights and biases) and outputs are treated as probability distributions rather than fixed values like in deterministic training approaches. This means that instead of making a single prediction, BNNs provide a range of possible outcomes, alongside a probability distribution that reflects the model's confidence in those predictions. By providing uncertainty estimation, BNNs allow for more informed decision-making, in which a high uncertainty signals the need for caution or further investigation before acting on the prediction, while a low uncertainty increases confidence in the prediction's accuracy.

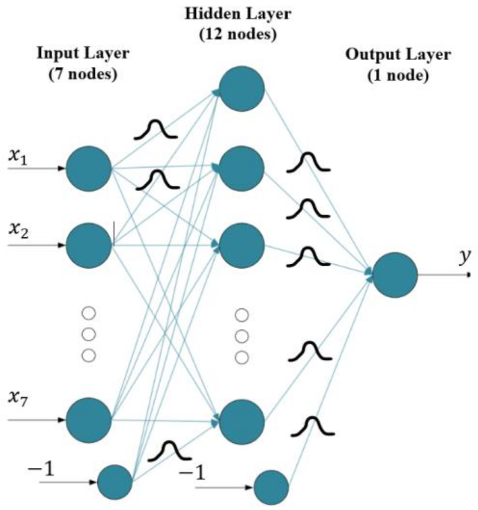

FIGURE 1. Structure of Bayesian Neural Network

In this study, a feed-forward neural network is developed as a classification unit. The structure of the network encompasses three layers: an input layer, a hidden layer and an output layer. There are seven input nodes which are seven quantified parameters as described in Table 1 (Parameters 1-7). The network has one output node which produces the prediction of students' academic progression at the end of their first semester as At-Risk or Normal. The ground truth of each prediction corresponds to the value of Parameter 8 (Table 1), which is set at 1 in the case the student actually makes unsatisfactory academic progression at the end of the semester, and 0 in the case the student can make satisfactory progression at the end of the semester. As shown in previous research, the number of hidden nodes of BNNs can be considered as a parameter to be learned in training the networks [10]. However, the main aim of the current study is to explore the ability of predicting students' academic progression by using BNNs as well as quantifying confidence level of the predictions. Accordingly, the number of hidden nodes in the hidden layer is fixed at 10 in this study. This number is chosen based on our previous experiences which indicated a range of 8-12 hidden nodes would work best for a neural network structure of 7 inputs and 1 output. The structure of the network is shown in Figure 1.

In the developed BNN, the network consists of a set of trainable network parameters *w*. Given the training dataset *D*, the posterior distribution over the weights *P(w\D)* is obtained by the following:

$$P(w|D) = \frac{P(D|w)P(w)}{P(D)}$$

where *P(D\w)* is the likelihood of the training data D given the parameter setting *w*, *P(w)* is the prior beliefs about the network's weights before observing any data, and *P(D)* is the evidence of observing the data under all possible configurations of the network parameters:

$$P(D) = \int P(D|w)P(w)dw$$

The challenge of training BNNs is the intractability of the evidence *P(D)* due to the requirement of integrating over the high-dimensional parameter space. This leads to the intractability of the posterior probability *P(w\D)*, requiring the use of an approximation method. In this study, Variational Inference will be applied, in which *P(w\D)* will be approximated with a tractable simpler distribution *Q(w)* by minimizing the divergence between the approximated distribution *Q(w)* and the true posterior distribution *P(w\D)* [11]. Once the posterior distribution *P(w\D)* is obtained, predictions on unseen data will be obtained by taking expectation on the predictive posterior distribution:

$$P(y|x) = E[P(y|x,w)] = \int P(y|x,w)P(w|D)dw$$

The uncertainty of predictions is quantified by the variance of the predictive posterior distribution which provides the confidence of the network in making predictions. To facilitate the results comparison and analysis, the uncertainty value of each prediction will be normalized to range between 0 and 1. The predictions' normalized uncertainty values less than 0.2 are considered as predictions with low confidence, while values higher or equal to 0.2 are considered as predictions with high confidence.

2.3. Result Interpretation and Evaluation Method

In this study, performance of the classification model will be evaluated by using the following metrics:

$$Sensitivity = \frac{TP}{TP + FN}$$

$$Specificity = \frac{TN}{TN + FP}$$

where *TP*, *TN*, *FP*, *FN* are defined as in Figure 2. Sensitivity and Specificity are common measures of classification, prediction or diagnosis accuracy. It is noted that for any diagnosis or prediction context, Sensitivity and Specificity are inversely related, wherein one increases as the other decreases. Depending on the nature of the diagnosis or prediction problems, a specific level of sensitivity or specificity would be required. For instance, in medical diagnosis, sensitivity which represents true positive rate should be prioritized since any correct diagnosis of medical problems is considered more important. Subsequently, by convention, sensitivity will be set at a higher rate, which will lead to a lower specificity. It is evident that a lower Specificity, which indicates a higher rate of incorrectly identifying individuals without a condition as having the condition, can result in inconveniences, as well as increased time and financial costs. However, it does not typically lead to fatal consequences, especially in the medical diagnosis context.

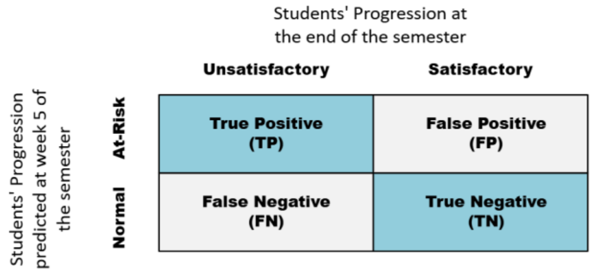

FIGURE 2. Confusion Matrix Definition

In this current study, the Sensitivity measures how well the algorithm can identify at-risk students and the Specificity measures how well it can identify normal students at an early stage of their first semester. It is noted that for the context of predicting students' academic progression, the algorithm capability to correctly predict at-risk students is more prioritized, provided with a reasonable value of Specificity. In this research, a Receiver Operation Characteristic (ROC) curve will be plotted after the training process, in which the output threshold that is corresponding to 85% Sensitivity will be selected and used to test the developed network on the testing set.

3. Results

3.1. Data Analysis Results

The dataset used in this study is extracted from 353 undergraduate students at Sydney Institute of Higher Education (SIHE). The entire dataset is separated into two groups: the Satisfactory group which consists of 118 data points from students who successfully progressed in their first semester of studying at SIHE, and the Unsatisfactory group which consists of 235 data points from students who made unsatisfactory academic progression in their first semester at SIHE. Each data point includes eight parameters as listed in Table 1.

Even though the main objective of this study is to explore the capability of a classification model based on Bayesian neural networks in predicting students' academic progression, statistical analyses are conducted to identify key demographic and academic attributes that can be added to the pool of parameters to enhance prediction accuracy. In all statistical tests, *p*-values less than 0.05 are considered as significant tests and presented in bold in Table 2.

TABLE 2. Data Analysis Results

Parameter Name	Group comparison *p*-values
Gender	**< 0.05**
Nationality	0.072
Age at Commencement	**< 0.05**
Grade 12 Result	0.097
English Entrance Test Results	0.153
Enrolled Program	**< 0.05**
Attendance Percentage in the first 4 weeks	**< 0.001**

It is shown that statistical results of the current study consistently match the results achieved in our previous study which indicated that the three parameters of students' Gender, Enrolled Program and Attendance Percentage in the first four weeks are the most significant features [7]. Potential explanations under these significances were discussed in our previous paper [7]. In this study, the age at commencement of students is newly added to the pool of parameters and noted as another significant feature ($p<0.05$). This result is predictable as students' age has been widely used as a strong demographic attribute which contributes to the effectiveness of predicting students' academic performance. It has been acknowledged that students who start their undergraduate studies earlier after their secondary education have higher likelihood of making satisfactory academic progression in their first semester in higher education. This can be explained by several factors including the continuity of studying momentum from their secondary education, the better ability of having peers who are at the same range of age, as well as the less burden in terms of family and financial responsibilities that older or mature students normally face with. Certainly, it is not a common rule that younger students will always perform better. Due to the transitions into the new education system and academic environment, it can be very challenging for young students with a lack of social skills to handle well in their first semester. This affirms the importance for higher education providers to have strategies of early identification and providing early intervention and academic support to help students overcome difficulties in their studies, especially in their first semester of studying.

Even though students' Academic Results in Grade 12 and Entrance English Results are important features contributing to students' success in higher education, this study's results indicated that there are insignificant differences between the two groups. The results achieved in this study consistently align with our previous study's findings [7]. This can be supported by the fact that all students must pass specific criteria of English Language Proficiency as well as Secondary Education Results to get admission into tertiary education. Moreover, the conversion between different grading systems is likely to have an impact on the results of statistical analyses.

Even though not all extracted parameters are statistically significant, all seven parameters (Parameters 1-7 as listed in Table 1) will be used as inputs into the BNN classification unit due to the well-known capabilities of recognizing underlying patterns within data and modelling non-linear relationships between inputs and outputs of neural networks.

3.2. Classification Results

To train the developed network by the Bayesian inference framework, the overall dataset is randomly divided into a training set and a testing set with an approximate ratio of 2:1 as follows:

- The training set consists of 235 data points including 80 Satisfactory points and 155 Unsatisfactory points.
- The testing set consists of 118 data points including 38 Satisfactory points and 80 Unsatisfactory points.

For comparison purposes, two other deterministic training methods including Levenberg- Marquardt algorithm (LM-NN) and a combination of Genetic Algorithm and Levenberg-Marquardt algorithm (GA+LM-NN) will be implemented [11] [12]. Details of the GA+LM training method can be found in our previous work [7]. To avoid overfitting, an inherent problem of the LM algorithm, the cross-validation technique is applied in each LM training step, in which the training set will be randomly subdivided into two subsets, a LM-training subset and a LM-validation subset with an approximate ratio of 3:1 as follows:

- The LM-training subset consists of 175 data points including 60 Satisfactory and 125 Unsatisfactory points.
- The LM-validation subset consists of 60 data points including 20 Satisfactory and 40 Unsatisfactory points.

Classification results of the three training methods are presented in Table 3 and Table 4. The reported results include the mean performance and the best performance across 20 runs for each training method. In all training runs, the Sensitivity of the training set is fixed at 85% as presented in Section 2. Given that the dataset in this study is more biased towards the Unsatisfactory label, making the network more sensitive to it, a Sensitivity value of 85% is considered reasonable for predicting at-risk students. This ensures an acceptable level of true negative rate. The confusion matrix corresponding to the training run which produces the best performance is presented in Figure 3 and Figure 4.

TABLE 3. Mean Classification Results

Training method	Training set		Testing set	
	Sen[a]	**Spe[b]**	**Sen[a]**	**Spe[b]**
BNN	85%	68%	80%	67%
LM-NN	85%	61%	69%	62%
GA+LM-NN	85%	66%	78%	61%

TABLE 4. Best Classification Results

Training method	Training set		Testing set	
	Sen[a]	**Spe[b]**	**Sen[a]**	**Spe[b]**
BNN	85%	70%	81%	74%
LM-NN	85%	61%	76%	63%
GA+LM-NN	85%	66%	80%	66%

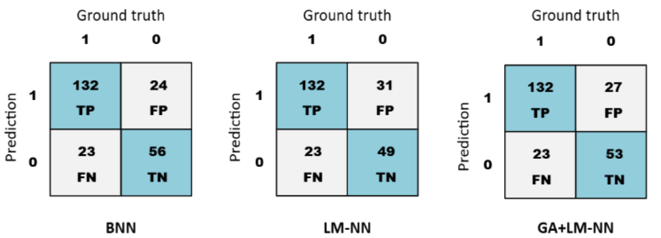

FIGURE 3. Comparison of confusion matrices on the training set in the training run with best classification results

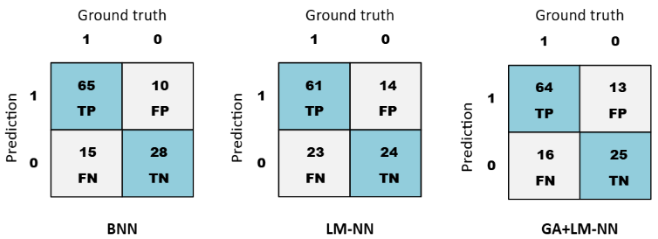

FIGURE 4. Comparison of confusion matrices on the testing set in the training run with best classification results

With the mean performance of 85% Sensitivity and 68% Specificity on the training set, and 80% Sensitivity and 67% Specificity on the testing test, it is indicated that the developed BNN has successfully overcome the overfitting problem of deterministic training methods without sacrificing data on a separate validation subset. The poor mean classification results on testing set of LM-NN method is predictable due to the inherent limitations of being trapped into local optimal of the LM method. The best results of 20 training runs for each training method show comparable performances between the three training methods, with a slightly better performance of the BNN compared to the other two methods. The results on the testing set demonstrate that the developed neural network can recognize the unseen demographic and academic patterns of unseen students and effectively predict students' academic performance at an early stage of their first semester of studies.

TABLE 5. Confidence Levels of Predictions on Testing Set

Predictions on Testing Set		Confidence level	
		High	**Low**
Unsatisfactory Prediction	Correctly predicted (TP)	53	12
	Wrongly predicted (FP)	6	4
Satisfactory Prediction	Correctly predicted (TN)	19	9
	Wrongly predicted (FN)	9	6
	Total	87	31

The results of predictions' confidence on the testing set are presented in Table 5. As described in Section 2, the normalized uncertainty of predictions provides a measure of confidence in the model's predictions. In this study, a threshold of 0.2 is used to categorize predictions into low and high confidence levels. The results indicate that the developed BNN tends to be more confident in predicting Unsatisfactory outcomes, with 79% of predictions classified as high confidence, compared to Satisfactory outcomes, with 65% of predictions are classified as high confidence. This is consistent with the fact that the dataset used in this study is biased with more data in the Unsatisfactory group. These results demonstrate that the developed BNN can effectively quantify uncertainties in its predictions. It is obvious that the ability to inform how confident the network is in its prediction via uncertainty quantification is a significant advantage of using the Bayesian inference framework to train the developed neural network. In the real context, predictions with low confidence should be referred to lecturers and the institute's academic learning support team for further evaluation and assessment before intervention decisions can be made.

4. Conclusions

This paper presented an innovative method of predicting students' academic progression for the objective of early identifying students who are at risk of making unsatisfactory academic progression in their first semester of higher education studies by using a classification unit based on Bayesian neural networks (BNNs). Using anonymous data from 353 students at Sydney Institute of Higher Education (SIHE) who enrolled in undergraduate programs, seven demographic and academic parameters were extracted and utilized as inputs into a neural network classification unit. The developed network was trained by using the Bayesian inference framework, in which networks' parameters and output are treated as probability distributions instead of producing fixed values like in deterministic training approaches. The expectation on the predictive distribution is considered as the final prediction outcome of the network, and the variance of the predictive distribution is considered as the quantified uncertainty of the prediction. The achieved classification results demonstrated that the developed method can effectively predict students' academic progression at an early stage. This early identification of students who are potentially making unsatisfactory academic progression is regarded as an essential intervention strategy to enhance students' academic success rate, especially in their first semester of studies in higher education.

For future research, the network structure optimization will be further explored, specifically to optimize the number of hidden nodes in hidden layer by Bayesian framework. More advanced algorithms with more complex network structures will also be explored to enhance the prediction performance. Furthermore, at this phase of the research, data are limited to international undergraduate students enrolled at SIHE over one year's period. This is likely to result in unavoidable biases in data analysis and predictive results interpretation. Accordingly, diversifying data sources with the participation of more higher education institutions in future research are likely to enhance the generalizability and robustness of the developed predictive model.

References

[1] "Final Report – Improving retention, completion and success in higher education", Higher Education Standards, 2017.

[2] TEQSA, "Characteristics of Australian higher education providers and their relation to first-year student attrition", June 2017.

[3] TEQSA, "Key findings from the 2021 risk assessment cycle", November 2022.

[4] Samina Sarwat, Naeem Ullah, Saima Sadiq, Robina Saleem, Muhammad Umer, Ala' Abdulmajid Eshmawi, Abdullah Mohamed, and Imran Ashraf, "Predicting Students' Academic Performance with Conditional Generative Adversarial Network and Deep SVM", Sensors, 22, 4834, June 2022.

[5] Mustafa Yağcı, "Educational data mining: prediction of students' academic performance using machine learning algorithms", Smart Learn. Environ. 9, 11, 2022.

[6] Carlos F. Rodríguez-Hernández, Mariel Musso, Eva Kyndt, Eduardo Cascallar, "Artificial neural networks in academic performance prediction: Systematic implementation and predictor evaluation", Computers and Education: Artificial Intelligence, Volume 2, 2021.

[7] Lien Nguyen, Anh Nguyen, Jack Jia, Steve Ling and Nigel Finch, "Prediction of Students' Academic Progression using Machine Learning", 2024 IEEE Conference on Artificial Intelligence (CAI), Singapore, pp. 1105-1110, 2024.

[8] Ponkrshnan Thiagarajan, Pushkar Khairnar and Susanta Ghosh, "Explanation and Use of Uncertainty Quantified by Bayesian Neural Network Classifiers for Breast Histopathology Images", IEEE Transactions on Medical Imaging, vol. 41, no. 4, pp. 815-825, April 2022.

[9] Laurent V. Jospin, Hamid Laga, Farid Boussaid, Wray Buntine and Mohammed Bennamoun, "Hands-On Bayesian Neural Networks—A Tutorial for Deep Learning Users", IEEE Computational Intelligence Magazine, vol. 17, no. 2, pp. 29-48, May 2022

[10] Anh V. Nguyen, Lien B. Nguyen, Steven Su and Hung T. Nguyen, "The advancement of an obstacle avoidance bayesian neural network for an intelligent wheelchair", 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Osaka, Japan, pp. 3642-3645, 2013.

[11] Martin T. Hagan and Mohammad B. Menhaj, "Training feedforward networks with the Marquardt algorithm", IEEE Transactions on Neural Networks, vol. 5, pp. 989-993, 1994.

[12] Yanrui Ding, Yujie Cai, P.D. Sun, B. Chen, "The Use of Combined Neural Networks and Genetic Algorithms for Prediction of River Water Quality", Journal of Applied Research and Technology, Volume 12, Issue 3, Pages 493-499, 2014.