Journal of International Technology and Information Management

Volume 30 | Issue 2 Article 4

2021

A comparison of artificial neural networks and the statistical methods in predicting MBA student's academic performance

Ojoung Kwon
California State University, Fresno, okwon@mail.fresnostate.edu

Harry Hui Xia
California State University, Fresno, hxiacsu@gmail.com

Serin Zhang
California State University, Fresno

A comparison of artificial neural networks and the statistical methods in predicting MBA student's academic performance

Ojoung Kwon
(*California State University, Fresno*)
Harry Hui Xia
(*California State University, Fresno*)
Serin Zhang
(*California State University, Fresno*)

ABSTRACT

MBA has become one of the most popular and vital professional degrees internationally. The MBA program admission process's essential task is to choose the best analysis tools to accurately predict applicants' academic performance potential based on the evaluation criteria in making admission decisions. Prior research finds that the Graduate Management Admission Test (GMAT) and undergraduate grade point average (UGPA) are common predictors of MBA academic performance indicated by graduate grade point average (GGPA). Using a sample of 250 MBA students enrolled in a state university with AACSB accreditation from Fall 2010 to Fall 2017, we test and compare the effectiveness of artificial neural networks (ANNs) against traditional statistical methods of ordinary least squares (OLS) and logistic regression in MBA academic performance prediction. We find that ANNs generate similar predictive power as OLS regression in predicting the numerical value of GGPA. By dichotomizing GGPA into categorical variables of "successful" and "marginal," we identify that ANNs offer the most reliable prediction based on total GMAT score and UGPA while logistic regression delivers superior performance based on other combinations of the predictors. Our findings shed light on adopting ANNs to predict academic performance potential with a strong implication in MBA admissions to select qualified applicants in a competitive environment.

Keywords: artificial neural networks, ANNs, MBA academic performance

INTRODUCTION

The Master of Business Administration (MBA) is one of the most prevalent and essential professional degrees worldwide (Baruch & Leeming, 2001) since it is usually viewed as a ticket to the executive suite (Kelan & Jones, 2010).

MBA admissions have always focused on "selection," which is heavily influenced by demand continually exceeding supply. A recent Graduate Management Admission Council (GMAC) analysis of 246 full-time MBA programs that disclose enrollment data on U.S. News & World Report's website showed that 119,338 applications resulted in 18,829 enrollments, a ratio of 16 percent with a somewhat higher offer-to-application ratio at 31 percent (Chowfla, 2021). It is critical and challenging for the MBA admission process to ensure the selection of the appropriate applicants with the necessary qualifications for successfully completing the program (Dakduk et al., 2016), which calls for an effective measure of the academic performance potential of the applicants in making admission decisions. Therefore, how to adequately and accurately predict MBA candidates' academic performance becomes a vital issue for business schools, accreditation agencies, and scholars interested in education development. (Dakduk et al., 2016; Kuncel et al., 2007).

MBA programs consider a variety of factors when making admission decisions. Typical evaluation criteria include overall undergraduate grade point average (UGPA), junior/senior GPA, undergraduate major and institution, Graduate Management Admissions Test (GMAT) score, references, goals statement, a personal interview, and others. (Naik & Ragothaman, 2004). Among them, GMAT, a pragmatically derived test measuring cognitive and academic skills (Koys, 2010), and UGPA, a precise and accessible indicator certifying the previous achievements (Garbanzo Vargas, 2012), are the most common and best predictors of MBA academic performance proxied by graduate grade point average (GGPA) (Ahmadi et al., 1997; Dakduk et al., 2016; Gupta & Turek, 2015).

Traditionally, practitioners and researchers use statistical methods, including multiple regression, stepwise regression, discriminant analysis, to predict applicants' success in the MBA program (Wright & Palmer, 1997). The drawback of these statistical methods is that they usually assume normality and homoscedastic variances. If these assumptions are violated in real-world data structures, the predictability of regression is diminished. Moreover, there are other typical challenges found in earlier research, such as a rather skewed distribution of GGPA (Abedi, 1991) and the low value of R-squared (R^2) with multiple regression and stepwise regression (Pharr & Bailey, 1993).

Because of the aforementioned limitations of traditional statistical methods, machine-learning programs and artificial neural networks (ANNs) have become increasingly popular for classification and decision-making (Naik & Ragothaman,

.

2004; Ragothaman & Naik, 1994). For the past several decades, ANNs, originated in mathematical neurobiology attempting to model the human brain's capabilities, have been used as an essential tool for quantitative modeling. ANNs are successfully applied to solve various problems, such as pattern classification, time series analysis, and prediction, in almost all business (Strader et al., 2020), industry (Chan, 2007; Wray et al., 2003), and science (Zhang, 2005).

The purpose of this research is to shed light on the ongoing debate on the performance of ANNs in academic performance prediction. We examine and compare ANNs to the popular statistical methods, namely, ordinary least squares regression (OLS) and logistic regression, to predict MBA student performance indicated by GGPA. We find that ANNs generate similar predictive power as OLS regression in predicting the numerical value of GGPA. Meanwhile, by dichotomizing GGPA into "successful" and "marginal" with a threshold of 3.6, we observe that ANNs deliver superior performance based on UGPA and GMAT total while logistic regression outperforms ANNs and OLS regression based on other predictors to project the categorical value of GGPA.

LITERATURE REVIEW AND RESEARCH PROPOSITION

The MBA admission process is crucial to selecting qualified applicants and controlling the programs' quality and reputation, in which a common and critical task is to determine suitable methods to predict future academic success (Romero & Ventura, 2010). The objective of such predictions is to estimate the academic performance potential of the applicants. In practice, the predictive values can be numerical/continuous values of GGPA through regression analysis to determine the relationship between GGPA and one or more independent variables (Draper & Smith, 1998) such as GMAT and UGPA. Alternatively, they can be categorical/discrete values through classification, a procedure in which individual items (GGPA) are placed into groups based on quantitative information regarding one or more characteristics (such as "successful" and "marginal") inherent in the items (Espejo et al., 2010).

Besides traditional statistical methods, due to the importance of MBA academic performance prediction, which is often fraught with variety, ambiguity, and complexity, ANNs are appealing as a predictive tool precisely because of their expected effectiveness in such a situation (Lippmann, 1987).

In practice, along with traditional statistical regression methods, ANNs have been used to predict student grades (Gedeon & Turner, 1993) and applicants' likely performance (Oladokun et al., 2008).

However, the extant literature on academic performance predictive tools provides mixed evidence. Gorr et al. (1994) compare linear regression, stepwise polynomial

.

regression, and fully-connected, single middle layer ANNs with an index for predicting student GPA in professional school admissions and discover that none of the empirically estimated methods show any statistically significant improvement. To address skewed distribution, by introducing skewness in the dependent variable, a comparative analysis of prediction of MBA academic performance using ANNs and regression show that both bias and absolute percentage error are higher among the results generated by the ANNs method (SubbaNarasimha et al., 2000).

Another stream of study focuses on categorical approaches. Hardgrave et al. (1994) evaluate the ability of five different methods: OLS regression, stepwise regression, discriminant analysis, logistic regression, and ANNs to predict the success of graduate MBA students, which generate poor results with the best method accurately predicting 60% of the cases. They also find that three categorical methods: discriminant analysis, logistic regression, and ANNs, seem to outperform the numerical regression methods. Asogwa and Oladugba (2015) argue that ANNs outperform Multinomial Logistic Regression in terms of the Average Classification Correct Rate for classifying students based on their academic performance. On the contrary, Walczak and Sincich (1999) conclude that the accuracy of ANNs is not significantly greater than the logistic regression analysis.

It can be observed from the literature that, despite a wealth of research, neither ANNs nor statistical methods deliver conclusive superiority in academic performance prediction, the key task in the MBA admission process. Therefore, our research proposition is to examine and compare the effectiveness of ANNs to the popular statistical methods in predicting MBA students' academic performance. By considering the academic performance proxied by GGPA as a dependent variable, the predictive analysis is conducted using ANNs, OSL regression, and logistic regression, respectively. To test the predictive accuracy, we will compare the outcomes of the numerical and categorical value of GGPA through statistical tests and F1-score to evaluate the performance of various predictive methods (Paliwal & Kumar, 2009).

.

DATA AND RESEARCH METHODOLOGY

Data

Data Collection

Our total sample contains 279 students enrolled in the traditional MBA program at a state university with AACSB accreditation from Fall 2010 to Fall 2017 to develop a tailored ANNs prediction method. No personal identification information is collected for this research, and we obtained the approval of using such data according to the related university policy and procedure. Our final data set contains 250 records after eliminating 29 records with missing data. Most of these eliminated records do not have information on GMAT analytical writing assessment (AWA) scores.

The descriptive statistics for the related variables show that mean (median) of UGPA is 3.22 (3.23) with a standard deviation of 0.42, GMAT total is 550 (540) with a standard deviation of 62, and GGPA is 3.59 (3.60) with a standard deviation of 0.27. Based on the median GGPA of 3.60, we dichotomize GGPA into two categories: “successful” and “marginal” with a threshold of 3.6.

Random Sampling

We randomly partition the dataset of 250 records into five groups with a constant sample size (each contains 50 records or 20% of the total sample). Each group is considered a fold, and hence, there are five folds. We perform a 5-fold cross-validation analysis. For each iteration, we use four folds of 200 records, which is 80% of the total sample, as the training set, and fit the models to the remaining one fold (50 records) as the testing set for measuring the accuracy of ANNs against OLS and logistics regression analysis.

For ANNs, the training set is further divided into two subgroups: training data (160 records or 80% of four folds training set) and validation data (40 records or 20% of four folds training set).

.

Variables Selection and Predictive Methods

Following the extant literature (Kass et al., 2012; Kuncel et al., 2007; Oh et al., 2008), we group MBA admission criteria of UGPA, GMAT total and subtest scores (verbal, quantitative, and analytical writing assessment) as independent variables or input into four models as shown below to predict MBA student academic performance indicated by GGPA or as output for ANNs, OSL, and logistic regression, respectively. We exclude other non-contributing information (i.e., gender, age, race, and undergraduate institution) from our final data set after running the initial validity checks. We find that such an approach is well aligned with the GMAC report (Talento-Miller & Rudner, 2008).

Model 1: *UGPA*
Model 2: *GMAT*
Model 3: *UGPA + GMAT*
Model 4: *UGPA + Verbal + Quant + AWA*

Where
UGPA is the Undergraduate Grade Point Average
GMAT is the GMAT total score
Verbal is the verbal part of the GMAT score
Quant is the quantitative part of the GMAT score
AWA is the analytical writing assessment part of the GMAT score

We choose OLS regression and ANNs to predict the numerical value of GGPA, indicating MBA students' academic performance. Besides, the predicted GGPA from OLS regression and ANNs is further converted to categorical value by dichotomized into two categories with the threshold of 3.6: "successful" (GGPA $\geq$ 3.6) and "marginal" (GGPA $<$ 3.6), to compare to the outcome of logistic regression method.

Artificial Neural Networks (ANNs) Method

Many different ANNs architectures are available and are tested in previous studies. In this research, we use the Multi-Layer Perceptron (MLP) with the backpropagation algorithm. MLP is one of the most popular ANNs architectures and is widely accepted (Alyuda Research, 2006). An MLP is a feed-forward ANNs architecture with the ability to keep improving its performance (i.e., reducing generated output errors) by iteratively changing the interconnecting weight of the architecture among all input layer connections, hidden layer, and output layer (Gardner & Dorling, 1998).

When used with MLP, the logistic sigmoid function reduces the outliers' effect (Hill et al., 1994; Maier & Dandy, 2000). The backpropagation algorithm is commonly used for MLP network training (Dawson & Wilby, 2001).
This algorithm might reduce the overall network error between network outputs and target values by adjusting the networks' interconnecting weights iteratively (Gardner & Dorling, 1998). Thus, MLP with the logistic sigmoid function becomes our choice in this study.

Software Used

We use the Alyuda NeuroIntelligence (ANI) to create ANNs models. Genetic Algorithm (GA) has the capabilities in pattern recognition, categorization, and association, and therefore, it has been widely applied in ANNs. Trippi and Turban (1992) show that a genetic algorithm enables ANNs to learn and adapt to changes through machine learning for automatically solving complex problems based on a set of repeated instructions. GA enables ANNs to produce improved solutions by selecting input variables with higher fitness ratings. ANI uses GA (built-in) and Fuzzy Logic to retain the best network.

Six Steps to Build ANNs

We follow the ANI's six-step neural network design process to build up the network: data analysis, data preprocessing, network design, training, testing, and query. The logistic function is applied to design the network. The logistic function has a sigmoid curve of $F(x) = 1/ (1+ e^{-x})$ with output range of [-1, 0.1]. A batch backpropagation model with a stopping training condition by error value (≤ 0.1) is used to find the best network during the network training.

(1) Data Analysis

The first step of data analysis is to flag missing data, wrong data types, and outliers. ANI generates these data with color codes (yellow and red) to quickly detect and resolve the issues. There are two sets of data used in the ANNs model: training set (including training data and validation data) and testing data. The training data is used to train the neural network and adjust network weights. The validation data is used to tune network parameters other than weights, calculate generalization loss, and retain the best network. The testing set is used to test how well the neural network performs on new data after the network is trained. Finally, we used the same testing set (50 records or 20% of the total sample) to examine the estimated errors between the actual and predicted values.

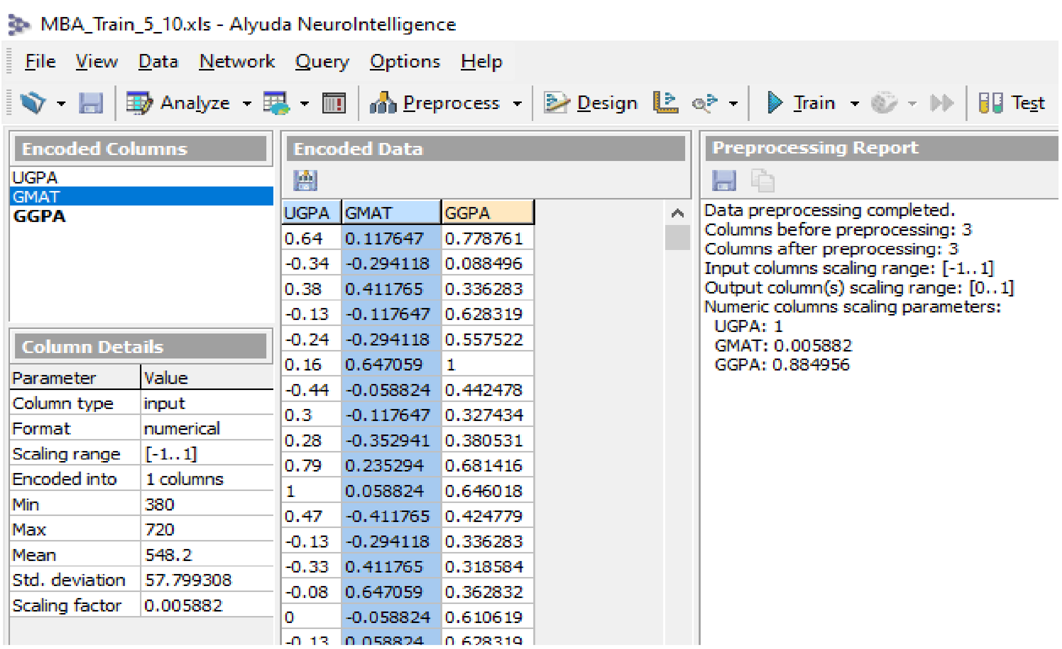
ISSN: 1941-6679-On-line Copy

.

2 Data Preprocessing

Because ANNs work only with numeric data, data normally need to be transformed (or preprocessed) to be suitable for the neural network before being fed to ANNs. Dates, time, and categories need to be transformed into numerical values. Numerical values are scaled to (0 to 1) or (-1 to 1), and textual (or categorical) values are converted into numeric ones (i.e., female = 1 and male = 0). In our study, ANI automatically transforms UGPA, GMAT, and GGPA into numeric values, as indicated in Figure 1.

Figure 1. Data Preprocessing Screen of NeuroIntelligence



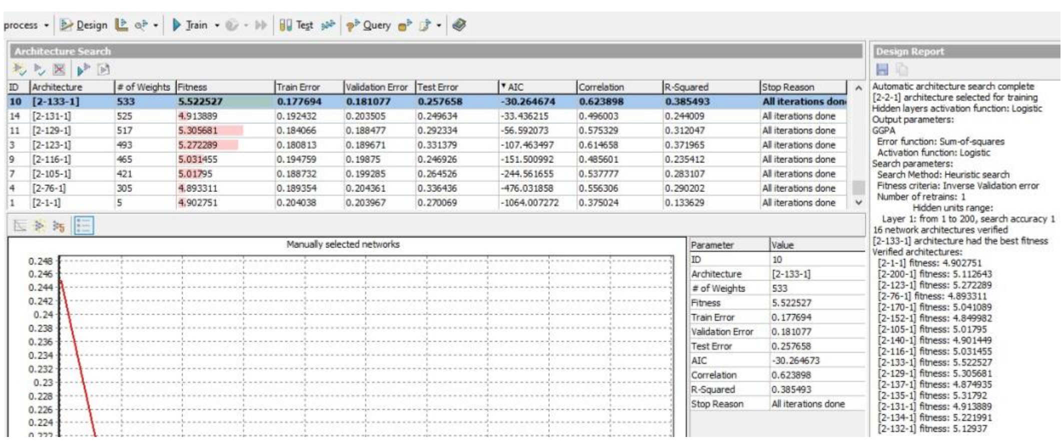
3 ANNs Design

To determine the numbers of Hidden Nodes (HN) for the ANNs, we combine the rules from previous research, which discusses the relationship among the HN, the numbers of Inputs (I) for the input layer, and the numbers of Outputs (O) for the output layer. According to Fletcher and Goss (1993), the numbers of HN should range from (I/2+O) to (2I+1). Palani et al. (2008) suggest that the HN should range from (I/3+O) to (2I+1). Moreover, Alyuda Research (2006) suggests that the HN should range from I/2 to 4I.

.

Furthermore, Gazzaz et al. (2012) combine the former three rules for ANNs application with ANI and state that the HN should range between I/3 and 4I. As a result, for this study, we set the HN to lay down between I/3 and 4I and larger than O. (I/3 < HN < 4I ∩ HN > O).
Based on the existing literature, different criteria can be applied to ANI for the best ANNs architecture searching. Gazzaz et al. (2012) propose R-squared (R^2) as a model selection criterion in forecasting the water quality index. Huang (2013) uses minimum testing error as criteria to select ANNs architecture for Exchange Rate Prediction Model. Gaurang et al. (2010) discuss and indicate the significant efficiency of the Akaike information criterion (AIC) in ANNs architecture searching. We argue that AIC is the best ANNs architecture searching criteria in our study, in which the architecture with the highest AIC is selected for the network training.
Besides, machine learning takes a great deal of time. It necessitates running many combinations to find the best ANNs architecture before the actual training starts. On top of that, there are so many more combinations to process until acceptable learning algorithms for each case can be found. Figure 2 shows a screenshot of the "Finding the Best Architecture" process of NeuroIntelligence.

Figure 2. Finding the Best Architecture Screen of NeuroIntelligence



.

4 ANNs Training

According to Alyuda NeuroIntelligence (2010), "the backpropagation algorithm is the most popular algorithm for training multi-layer perceptrons and is often used by researchers and practitioners. The main drawbacks of backpropagation are: slow convergence, need to tune up the learning rate and momentum parameters, and high probability of getting caught in local minima." Gaussian distribution of network inputs is used to retrain and restore the best network and randomize weights. By doing so, over-training, such as memorizing data instead of generalizing and encoding data relationships, can be prevented and thus reduce network errors. In this study, 10% jitter (random noise) is added to avoid over-training and local minima. Weights randomization can avoid sigmoid saturation that causes slow training.
There are several training stop criteria from previous papers on ANI. Anwar and Watanabe (2010) set the termination of training after 20,000 iterations or Mean Squared Error (MSE) < 0.000001, and the learning & momentum rate at 0.1 for backpropagation. Gazzaz et al. (2012) has applied 0.000001 as the network MSE improvement, 0.01 of training set MSE, and maximum for 10,000 iterations. Also, Gazzaz et al. (2012) retrain ten times, according to the ANI manual. Meng (2008) applies 50,000 iterations and network error (MSE) as 0.01 in predicting the return on IPO in China stock market. Since the training process is uncertain, more training times for the ANNs will have a better chance of achieving more accurate results. For this research, the training is set to stop when 100,000 iterations are completed or stop training when it reaches 0.10 (or 10%) average training errors.

5 Testing

ANI automatically performs network testing after training completion. The "Actual vs. Output" table displays error values for each record from the input dataset. It allows us to filter the table to show only records from training data or validation data using the corresponding toolbar buttons. We can browse a table using the scrollbar or navigation keys to inspect which records produce more significant errors. We also used the "Actual vs. Output Graph" (see Figure 3) or "Scatter Plot" (see Figure 4) to visualize the gap between the actual vs. neural network output.

6 Query

Finally, a batch of the testing set (50 records or 20% of the total sample) is fed into the ANI via the Batch Query mode. The predicted output from ANI is downloaded into Excel to calculate the difference (the forecasting error) between the actual and predicted GGPA.

.

Figure 3. Actual vs. Output Graph of NeuroIntelligence

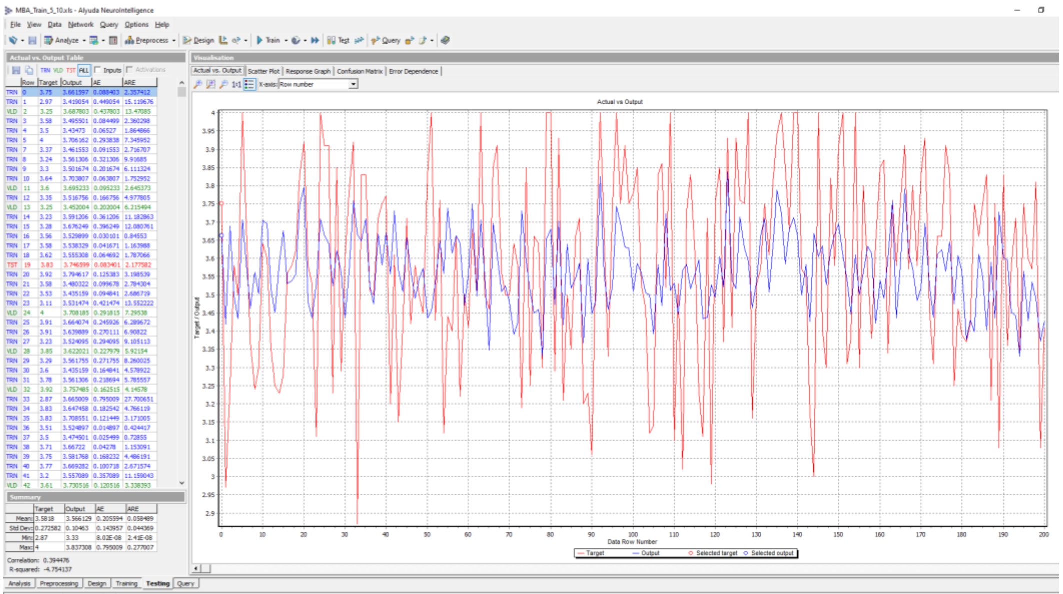
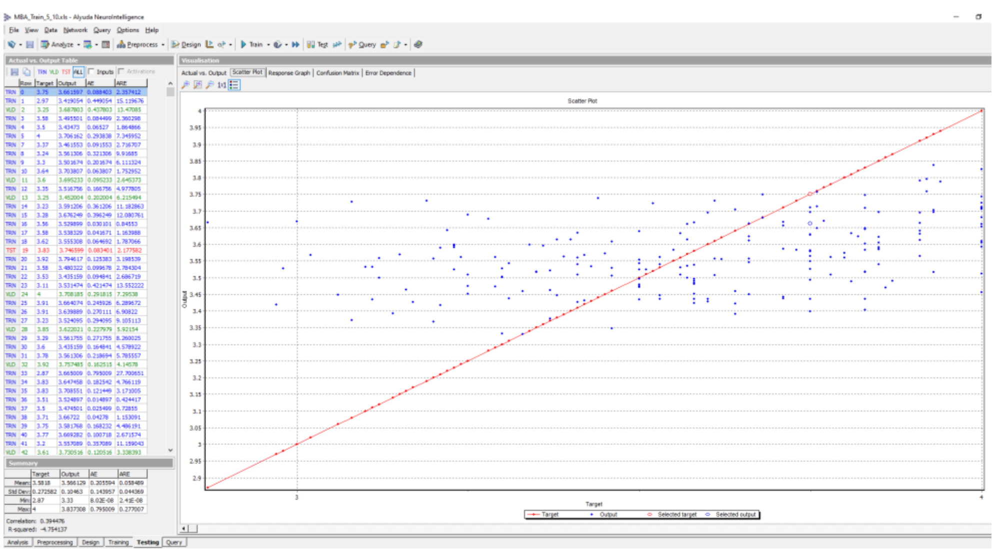


Figure 4. Scatter Plot of NeuroIntelligence



.

OSL Regression

As a traditional statistical tool, the OLS regression method has many advantages. It is easy to use, validate, and typically generate the best combination of predictors using stepwise regression. However, the regression method is a linear model with relatively high forecasting errors when forecasting a nonlinear environment. Besides, the regression method can only predict one dependent variable at a time. It works well with our research design because we would like to predict MBA student academic performance indicated by GGPA in this study. Below are the four regression models designed according to variables selection stated in 3.1.3 :

Model 1: $GGPA = \beta_0 + \beta_1 \times UGPA$
Model 2: $GGPA = \beta_0 + \beta_1 \times GMAT$
Model 3: $GGPA = \beta_0 + \beta_1 \times UGPA + \beta_2 \times GMAT$
Model 4: $GGPA = \beta_0 + \beta_1 \times UGPA + \beta_2 \times Verbal + \beta_3 \times Quant + \beta_4 \times AWA$

Logistics Regression

Logistic regression analyzes categorical/discrete values in modeling binary outcomes such as pass/failure or win/lose, which is often considered a binary classifier. The response (or outcome) variable is modeled according to the probability of success, $P(Y = 1)$, and the probability of failure, $P(Y = 0)$, which is called a Bernoulli process. The ratio of $P(Y=1)$ and $P(Y=0)$ is called the odds ratio, and the log of this ratio is the *logit*. The *logit* is then modeled linearly with the predictor variables used in Model 1 to 4 (same as the OLS regression).
We fit four different logistic regression models with the same sets of predictors used in OLS regression analysis. The response variable, GGPA, is dichotomized before fitting the models. If the GGPA is greater than or equals to 3.6, it is coded as 1, indicating “successful.” If it is below 3.6, coded as 0, indicating “marginal.” We apply these fitted models to the test data using 5-fold cross-validation and obtain the estimated probability of GGPA of the students in the test data.
If the estimated probability is greater than or equals to 0.5, the predicted GGPA through logistic regression is assigned as 1 (successful) and 0 (marginal) otherwise.

.

Predictive Ability Analysis

Numerical Value Prediction Methods

We evaluate the results of numerical value prediction of GGPA using ANNs and OLS regression for 5 test data folds with the Mean of Errors (MoE) and the standard deviation of all five folds. The MoE is calculated as:

FE = ABS (GGPA – PGGPA) / GGPA
MoE = Average (FE) for all 5 test data folds

where
FE is the forecasting error that is the absolute value of the difference between the actual GGPA and the predicted GGPA (PGGPA) then divided the actual GGPA (SubbaNarasimha et al., 2000)
ABS stands for the absolute value
GGPA is a Graduate GPA and is the actual GPA that the MBA students earned
PGGPA is a Predicted GGPA and is the output of ANNs and the OLS regression method

The predictive method with lower MoE and standard deviation indicates greater accuracy or higher predictive ability.

Categorical Value Prediction Methods

The prediction accuracy of ANNs, OLS, and logistic regression methods is compared using the F1 scores calculated from 5-fold cross-validation results. The comparison is based on a categorical prediction, in which we analyze the capability of the models to predict the categorized response variable, successful, or marginal GGPA of the MBA students. The response variable GGPA indicating MBA students' academic performance is dichotomized into two categories: "successful" and "marginal" with the threshold of 3.6. We apply the F1 score as the measure of the predictive ability of the three methods.

$$F1 = 2 \times \frac{Precision \times Recall}{Precision \ + Recall}$$

F1 score is widely used as a measure of accuracy in statistical analysis for binary classification. It is the harmonic average of precision and recall. The precision is the ratio of correctly predicted positive results and the total positive prediction.

.

The recall, also called sensitivity, is the ratio of correctly identified positive prediction and the total actual positive outcome. F1 score takes any values between 0 and 1, with 1 indicating the perfect precision.

ANALYSIS AND FINDINGS

Results of the Numerical Value Predictive Methods

Table 1 shows the comparison of MoE and the standard deviation of the results of the four models. We find that Model 3, using both UGPA and GMAT total as input variables, generates the predicted GGPA with the lowest MoE of 5.64% with a standard deviation of 4.47% through OLS regression. ANNs outcomes are consistent with OLS regression demonstrating Model 3 with the lowest MoE of 5.70% with a standard deviation of 4.68%. Such results indicate that Model 3 is the best one to predict GGPA among the four models. In practice, Model 3 is well-aligned with our current admission policy before the Covid-19 pandemic.
Our traditional MBA program requires applicants to submit UGPA and GMAT scores as key considerations to measure academic performance potential in the admission process.

Table 1. The Averages of 5-fold OLS and ANNs results

Model #	**Variables**	**OLS Regression**		**ANNs**	
		Mean of Errors (MoE)	Standard deviation (S.D.)	Mean of Errors (MoE)	Standard deviation (SD)
1	*UGPA*	6.03%	4.43%	6.30%	4.33%
2	*GMAT*	5.81%	4.63%	5.95%	4.56%
3	***UGPA+GMAT***	**5.64%**	4.47%	**5.70%**	4.68%
4	*UGPA+Verbal+Quant+AWA*	5.66%	4.53%	5.82%	4.81%

.

We conducted a paired t-test to see whether there is a difference between the two means of MoE. The hypothesis is set as:

H_o: $\mu_d = 0$
H_a: $\mu_d <> 0$

where μ_d is the difference between the means of OLS MoE and ANNs MoE

Table 2. t-Test: Paired Two Sample for Means

	All Four Models		**Model 3**	
	OLS MoE	*ANNs MoE*	*OLS MoE*	*ANNs MoE*
Mean	5.79%	5.85%	5.64%	5.70%
Variance	4.2285E-05	3.8417E-05	5.6058E-05	5.4099E-05
Observations	20	20	5	5
Pearson Correlation	0.970		0.997	
df	19		4	
t Stat	-1.841		-2.189	
Sig. (T<=t) two-tail	**8.13%**		**9.38%**	
t Critical two-tail	2.093		2.776	

The results of the paired t-test of OLS and ANNs means are shown in Table 2. The mean for the OLS regression method's MoE of all four models is 5.79% vs. 5.85% for ANNs. If we compare Model 3 alone, the OLS regression method still delivers a better result than ANNs (mean of MoE of 5.64% vs. 5.70%). However, although we may argue that the MoE of OLS is lower than ANNs, they are statistically insignificant at the level of 5% (8.13% for all four models and 9.38% for Model 3), which demonstrates that there is no real difference in predictive power between the OLS regression method and ANNs in forecasting the numerical value of GGPA.

.

Results of the Categorical Value Predictive Methods

For the categorical value of GGPA, we define “Successful” as GGPA $\geq$ 3.6, whereas “Marginal” as GGPA < 3.6. Table 3 shows the averages of 5-fold cross-validation results from ANNs, OLS, and logistic regression (LR) methods. From the table, we note that the percentages of correct predictions among “Successful” are greater than those of “Marginal,” which might be caused by the fact that there could be more variability among MBA students who have GGPA < 3.6 and calls for further investigation to locate the reasons behind this.

Table 3. The Average Percentages of Correct Predictions and F1 Scores

Model #	**Percentage of Correct Predictions among “Successful”**			**Percentage of Correct Predictions among “Marginal”**			**F1 Scores**		
	ANNs	OLS	LR	ANNs	OLS	LR	ANNs	OLS	LR
1	66.40%	61.70%	60.10%	59.90%	58.70%	59.80%	57.70%	56.50%	**59.50%**
2	65.10%	66.30%	63.60%	60.00%	61.10%	63.10%	53.90%	58.90%	**62.50%**
3	67.60%	71.80%	67.30%	66.00%	64.90%	64.40%	**66.90%**	65.40%	63.80%
4	63.50%	71.70%	67.30%	63.90%	65.50%	66.40%	65.10%	65.80%	**66.10%**

Table 3 also summarizes the average F1 scores of the three predictive methods using different combinations of predictors (four models). The logistic regression outperforms ANNs and OLS regression in three models except for Model 3. Such an outcome aligns with the fact that the F1 score is a measure of accuracy for binary classification, in which the logistic regression method specializes. For both OLS and logistic regression, it is clear that as the number of variables increases, the F1 score increases. However, for ANNs, the F1 score is highest when the input variables are UPGA and GMAT total for Model 3, which indicates that the ANNs are sensitive to the different combinations of input variables. Such a pattern is also true among the percentages of correct predictions for “Successful” and “Marginal,” respectively, in Model 3. Therefore, in terms of the categorical value prediction, ANNs are the best when both UGPA and GMAT total are predictors based on the F1 score, while logistic regression outperforms in all other predictive models using different combinations of UGPA, GMAT total, and GMAT subtest scores.

.

CONCLUSIONS

Making the right admission decisions to select suitable applicants with the potential to succeed in the MBA programs is critical to any business school in this competitive market. Our study investigates three popular MBA student academic performance (proxied by GGPA) predictive methods: ANNs, OLS, and logistic regression. By employing the numerical value of GGPA, we prove that ANNs deliver similar accuracy as OLS regression based on UGPA, GMAT total, and subtest scores (verbal, quantitative, and analytical writing assessment). Using categorical variables of “successful” and “marginal” with a threshold of 3.6 GGPA, we find that ANNs generate the most significant predictive power based on UGPA and GMAT total while logistic regression delivers more accurate results in the models using other predictor combinations. Our findings contribute to the extant literature by shedding light on the ongoing debate of the performance of ANNs in academic performance prediction with direct evidence of ANNs as an effective predictive tool to measure MBA academic performance potential. This study has a strong implication to the decision-makers in the MBA admission process by offering another tested tool for the critical task of selecting qualified applicants.

Although ANNs are particularly accurate in categorical MBA academic performance prediction based on the popular admission criteria of UGPA and GMAT total, we find no overwhelming proof that ANNs can clearly outperform traditional statistical methods in other models and deliver exceptional value considering the extra training time and more sophisticated resources ANNs demand. The possible explanation may relate to the limitations of this study. First, it is subject to a universal challenge of range restriction in educational research with a relatively high mean for GGPA and a small standard deviation resulting in limited variability. Second, this research only involves limited data from one university’s MBA program, which also constrains the generalizability of the findings. Finally, the restricted sample size and number of variables are disadvantageous to ANNs, which are more suitable for predictions based on large sample size and nonlinear relation between predictors and dependent variables.

Therefore, an opportunity for future researchers to further examine the effectiveness of ANNs is to create an adequate sample size by collecting more data from other academic programs and institutions. In addition, student academic performance depends on various factors, more than those considered in this research. It would be beneficial to extend this study by incorporating other variables that influence MBA academic performance outcomes, such as work experience, achievement motivation, soft skills, or self-efficacy, in future research of testing the effectiveness of ANNs in academic performance prediction.

.

REFERENCES

Abedi, J. (1991). Predicting Graduate Academic Success from Undergraduate Academic Performance: A Canonical Correlation Study. *Educational and Psychological Measurement*, *51*(1), 151–160. https://doi.org/10.1177/0013164491511014

Ahmadi, M., Raiszadeh, F., & Helms, M. (1997). An examination of the admission criteria for the MBA programs: A case study. *Education*, *117*(4), 540–547.

Alyuda NeuroIntelligence. (2010). *Alyuda NeuroIntelligence Manual*. http://www.alyuda.com/neural-networks-software.htm.

Alyuda Research. (2006). *Alyuda Forecaster XL2.3.* www.alyuda.com

Anwar, S., & Watanabe, K. (2010). Predicting Future Depositor`s Rate of Return Applying Neural Network: A Case-study of Indonesian Islamic Bank. *International Journal of Economics and Finance*, *2*(3), p170. https://doi.org/10.5539/ijef.v2n3p170

Asogwa, O., & Oladugba, A. (2015). *On The Comparison of Artificial Neural Network (ANN) and Multinomial Logistic Regression (MLR)*. Undefined. /paper/On-The-Comparison-of-Artificial-Neural-Network-and-Asogwa-Oladugba/e28603e6669c3a091af13bff055db7ee6909671a

Baruch, Y., & Leeming, A. (2001). The added value of MBA studies – graduates' perceptions. *Personnel Review*, *30*(5), 589–602. https://doi.org/10.1108/EUM0000000005941

Chan, J. (2007). A Predictive Analytic Model for Value Chain Management. *Journal of International Technology and Information Management*, *16*(1). https://scholarworks.lib.csusb.edu/jitim/vol16/iss1/3

Chowfla, S. (2021). *Selection in Management Education Admissions | AACSB*. https://www.aacsb.edu/blog/2019/july/changing-nature-selection-graduate-management-education

.

Dakduk, S., Malavé, J., Torres, C. C., Montesinos, H., & Michelena, L. (2016). Admission Criteria for MBA Programs: A Review. *SAGE Open*, *6*(4), 2158244016669395. https://doi.org/10.1177/2158244016669395

Dawson, C. W., & Wilby, R. L. (2001). Hydrological modeling using artificial neural networks. *Progress in Physical Geography: Earth and Environment*, *25*(1), 80–108. https://doi.org/10.1177/030913330102500104

Draper, N. R., & Smith, H. (1998). More on Checking Fitted Models. In *Applied Regression Analysis* (pp. 205–215). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118625590.ch8

Espejo, P. G., Ventura, S., & Herrera, F. (2010). A Survey on the Application of Genetic Programming to Classification. *IEEE Transactions on Systems, Man and Cybernetics. Part C, Applications and Reviews*, *40*(2), 121–144. https://doi.org/10.1109/TSMCC.2009.2033566

Fletcher, D., & Goss, E. (1993). Forecasting with neural networks: An application using bankruptcy data. *Information & Management*, *24*(3), 159–167. https://doi.org/10.1016/0378-7206(93)90064-Z

Garbanzo Vargas, G. M. (2012). Factores asociados al rendimiento académico en estudiantes universitarios, una reflexión desde la calidad de la educación superior pública. *Revista Educación*, *31*(1), 43. https://doi.org/10.15517/revedu.v31i1.1252

Gardner, M. W., & Dorling, S. R. (1998). Artificial neural networks (the multilayer perceptron)—A review of applications in the atmospheric sciences. *Atmospheric Environment*, *32*(14), 2627–2636. https://doi.org/10.1016/S1352-2310(97)00447-0

Gaurang, P., Ganatra, A., Kosta, Y., & Devyani, P. (2010). Searching Most Efficient Neural Network Architecture Using Akaike's Information Criterion (AIC). *International Journal of Computer Applications*, *5*. https://doi.org/10.5120/126-242

Gazzaz, N. M., Yusoff, M. K., Aris, A. Z., Juahir, H., & Ramli, M. F. (2012). Artificial neural network modeling of the water quality index for Kinta River

.

(Malaysia) using water quality variables as predictors. *Marine Pollution Bulletin*, *64*(11), 2409–2420. https://doi.org/10.1016/j.marpolbul.2012.08.005

Gedeon, T. D., & Turner, H. S. (1993). Explaining student grades predicted by a neural network. *Proceedings of 1993 International Conference on Neural Networks (IJCNN-93-Nagoya, Japan)*, *1*, 609–612 vol.1. https://doi.org/10.1109/IJCNN.1993.713989

Gorr, W. L., Nagin, D., & Szczypula, J. (1994). Comparative study of artificial neural network and statistical models for predicting student grade point averages. *International Journal of Forecasting*, *10*(1), 17–34. https://doi.org/10.1016/0169-2070(94)90046-9

Gupta, A., & Turek, J. (2015). Empirical investigation of predictors of success in an MBA programme. *Education & Training*, *57*(3), 279–289. http://dx.doi.org.hmlproxy.lib.csufresno.edu/10.1108/ET-10-2012-0100

Hardgrave, B. C., Wilson, R. L., & Walstrom, K. A. (1994). Predicting graduate student success: A comparison of neural networks and traditional techniques. *Computers & Operations Research*, *21*(3), 249–263. https://doi.org/10.1016/0305-0548(94)90088-4

Hill, T., Marquez, L., O'Connor, M., & Remus, W. (1994). Artificial neural network models for forecasting and decision making. *International Journal of Forecasting*, *10*(1), 5–15. https://doi.org/10.1016/0169-2070(94)90045-0

Huang, H. C. (2013). Using a Hybrid Neural Network to Predict the NTD/USD Exchange Rate. *Proceedings of the 2012 International Conference of Modern Computer Science and Applications*, 433–439. https://doi.org/10.1007/978-3-642-33030-8_70

Kass, D., Grandzol, C., & Bommer, W. (2012). The GMAT as a Predictor of MBA Performance: Less Success Than Meets the Eye. *Journal of Education for Business*, *87*(5), 290–295. https://doi.org/10.1080/08832323.2011.623196

Kelan, E. K., & Jones, R. D. (2010). Gender and the MBA. *Academy of Management Learning & Education*, *9*(1), 26–43. https://doi.org/10.5465/amle.9.1.zqr26

.

Koys, D. (2010). GMAT versus Alternatives: Predictive Validity Evidence from Central Europe and the Middle East. *Journal of Education for Business*, *85*(3), 180–185. https://doi.org/10.1080/08832320903258618

Kuncel, N. R., Credé, M., & Thomas, L. L. (2007). A Meta-Analysis of the Predictive Validity of the Graduate Management Admission Test (GMAT) and Undergraduate Grade Point Average (UGPA) for Graduate Student Academic Performance. *Academy of Management Learning & Education*, *6*(1), 51–68. https://doi.org/10.5465/amle.2007.24401702

Lippmann, R. (1987). An introduction to computing with neural nets. *IEEE ASSP Magazine*, *4*(2), 4–22. https://doi.org/10.1109/MASSP.1987.1165576

Maier, H. R., & Dandy, G. C. (2000). Neural networks for the prediction and forecasting of water resources variables: A review of modeling issues and applications. *Environmental Modeling & Software*, *15*(1), 101–124. https://doi.org/10.1016/S1364-8152(99)00007-9

Meng, D. (2008). A Neural Network Model to Predict Initial Return of Chinese SMEs Stock Market Initial Public Offerings. *2008 IEEE International Conference on Networking, Sensing and Control*, 394–398. https://doi.org/10.1109/ICNSC.2008.4525247

Naik, B., & Ragothaman, S. (2004). Using neural networks to predict MBA student success. *College Student Journal*, *38*(1), 143–150.

Oh, I.-S., Schmidt, F. L., Shaffer, J. A., & Le, H. (2008). The Graduate Management Admission Test (GMAT) is Even More Valid Than We Thought: A New Development in Meta-Analysis and Its Implications for the Validity of the GMAT. *Academy of Management Learning & Education*, *7*(4), 563–570. https://doi.org/10.5465/amle.2008.35882196

Oladokun, V. O., Adebanjo, A. T., & Charles-Owaba, O. E. (2008). *Predicting students academic performance using artificial neural network: A case study of an engineering course*. http://ir.library.ui.edu.ng/handle/123456789/1796

.

Palani, S., Liong, S.-Y., & Tkalich, P. (2008). An ANN application for water quality forecasting. *Marine Pollution Bulletin*, *56*(9), 1586–1597. https://doi.org/10.1016/j.marpolbul.2008.05.021

Paliwal, M., & Kumar, U. A. (2009). A study of academic performance of business school graduates using neural network and statistical techniques. *Expert Systems with Applications*, *36*(4), 7865–7872. https://doi.org/10.1016/j.eswa.2008.11.003

Pharr, S., & Bailey, J. (1993). Admission/continuance standards as predictors of academic performance of business students. *Journal of Education for Business*, *69*(2), 69. https://doi.org/10.1080/08832323.1993.10117659

Ragothaman, S., & Naik, B. (1994). Using Rule Induction for Expert System Development: The Case of Asset Writedowns. *International Journal of Intelligent Systems in Accounting and Finance Management*, *3*(3), 187–203. https://doi.org/10.1002/j.1099-1174.1994.tb00065.x

Romero, C., & Ventura, S. (2010). Educational Data Mining: A Review of the State of the Art. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, *40*(6), 601–618. https://doi.org/10.1109/TSMCC.2010.2053532

Strader, T., Rozycki, J., ROOT, T., & Huang, Y.-H. (John). (2020). Machine Learning Stock Market Prediction Studies: Review and Research Directions. *Journal of International Technology and Information Management*, *28*(4), 63–83.

SubbaNarasimha, P. N., Arinze, B., & Anandarajan, M. (2000). The predictive accuracy of artificial neural networks and multiple regression in the case of skewed data: Exploration of some issues. *Expert Systems with Applications*, *19*(2), 117–123. https://doi.org/10.1016/S0957-4174(00)00026-9

Talento-Miller, E., & Rudner, L. M. (2008). The Validity of Graduate Management Admission Test Scores: A Summary of Studies Conducted From 1997 to 2004. *Educational and Psychological Measurement*, *68*(1), 129–138. https://doi.org/10.1177/0013164407305581

.

Trippi, R., & Turban, E. (1992). *Neural Networks in Finance and Investing: Using Artificial Intelligence to Improve Real-World Performance*. https://doi.org/10.1016/s0169-2070(96)00728-5

Walczak, S., & Sincich, T. (1999). A Comparative Analysis of Regression and Neural Networks for University Admissions. *Information Sciences*, *119*(1–2), 1–20. https://doi.org/10.1016/S0020-0255(99)00057-2

Wray, B. A., Markham, I. S., & Mathieu, R. G. (2003). An Artificial Neural Network Approach to Learning from Factory Performance in a Kanban-Based System. . . *Number*, *12*(2), 15.

Wright, R. E., & Palmer, J. C. (1997). Examining performance predictors for differentially successful MBA students. *College Student Journal*, *31*(2). https://www.elibrary.ru/item.asp?id=4737163

Zhang, P. G. (2005). Neural Networks. In O. Maimon & L. Rokach (Eds.), *Data Mining and Knowledge Discovery Handbook* (pp. 487–516). Springer US. https://doi.org/10.1007/0-387-25465-X_22

.