



# The Performance Comparison of Models for Predicting the Risk of Losing Student Loan by Fuzzy Neural Network Method and Multiple Linear Regression Analysis Method

Tawin Tanawong\*, Sanya Khruahong and Adirek Roongrungsi

Department of Computer Science and Information Technology, Faculty of Science, Naresuan University, Phitsanulok, 65000, Thailand

\* Corresponding author. E-mail address: tawint@nu.ac.th

Received: 21 August 2019; Revised: 19 October 2019; Accepted: 5 November 2019; Available online: 27 April 2020

## Abstract

This research presents the comparing the performance of the prediction model of the risk of losing student loan by Fuzzy Neural Network (FNN) and Multiple Linear Regression Analysis (MLR). We used the preparation of 296 samples, assigned the income variables, which are characterised by Fuzzy attribute and had problem-solving with Dummy variables. Also, we proposed the method of converting fuzzy data by making a fuzzy attribute matching of Neural Network (NN) and Multiple Linear Regression. It calculates the fuzzy linguistic term, the fuzzy language and the Crisp value, and after that have filtered variables with Pearson correlation technique and multiple linear regression which get eight independent variables and RiskForPay is a dependent variable. Results from this research, the appropriate model of Fuzzy Neural Network is a division of learning data, and cross-validation fold is 5 with an accuracy of 83.33%  $\pm$  6.02% which models have 8-5-6 structure, momentum 0.2 and learning rate 0.3. The predictive model with multiple linear regression equations has Root Mean Squared Error: 1.513  $\pm$  0.000 and Squared Correlation: 0.081, respectively.

**Keywords:** Fuzzy Neural Network, Multiple Linear Regression Analysis, Student Loan, Fuzzy logic

## Introduction

Thai family consider that education is important; they have the desire to push their children to study at university. However, some family may have some problem with income cause to they can not support their children for learning in the university. Therefore, the Thailand government had organised the Student Loan Fund (SLF) in 1995 for a student in secondary and university students whom low-income families. This fund can help them to complete their education. After graduating, they can get the two years of a debt moratorium, which had not made a scheduled payment of interest or principal. In the past operations, we found that we could not screen students who are capable of paying off debt. Various educational institutions have impacted in the loss of money from non-payment of students, which affect the borrowing of new students in each university. However, the original form of operation has only screened the loans by asking questions, checking history, and only proofing the student's basic information documents. Also, the crucial information is related to the students stored in the Education Services Division of each university, but cannot be used to truly consider the scholarship. Therefore, it should be applied to some techniques for finding the methods for protection of losing money from student loans to be clear by using the prediction model. Data Mining can be applied to create predictive models for monitoring the risk of losing loans to be more excellent reliability.

For analysing the problem of the student loan fund, we proposed a model for predicting the risk of lending to individual students by comparing performance models and forecasting losses student loan between fuzzy

neural networks and multiple linear regression analysis methods. We applied the technology to help manage the loan fund to be more suitable for the payment of students to be more efficient. For this research, there are various parts of the work, which are the preparation of knowledge information related to the fundamentals of former students who have studied at the university and graduated student by using Training data sets and Testing data sets. The classification method is used for classifying the patterns and improving the problem learning model to perform forecasting with a web application which was developed for this research.

### Methods and Materials

Prediction modelling is very crucial and a mainstay of statistical practice. It can be applied in many situations. There are many popular models for implementing the prediction system. In this article, we interested in the Fuzzy Neural Network (FNN) and Multiple Linear Regression Analysis (MLR).

#### Fuzzy Neural Network

Fuzzy Neural Network (FNN) (Mendel, 1995; Osowski & Linh, 2001; Tang, Liu, Zhang, Ke, & Zou, 2018; Wang, Hao, Ma, & Huang, 2010) or Neuro-Fuzzy System is a learning machine that decides the parameters in the fuzzy system (i.e., fuzzy sets, fuzzy rules) by employing approximation methods from neural networks. A problem can be solved with neural networks and fuzzy systems (e.g. regression, pattern recognition, or density estimation) if it has not existed any mathematical model of the given problem. They do have specific weaknesses and benefits, which almost completely disappear by mixing both concepts.

Neural networks (Haykin, 2009) can apply to use if a sufficient amount of observed examples can express the problem. This information is used to train the black box, which, no prior knowledge about the problem needs to be given. It is not straightforward to obtain comprehensible rules from the neural network's structure. Moreover, a fuzzy system demands linguistic rules instead of studying examples as previous knowledge. Besides, the input and output variables have to be specified linguistically. If the information is incomplete, incorrect or contradictory, then the fuzzy system must be tuned. Since it does not have any formal approach, it tuned with the heuristic method, which often takes a long time and error-prone.

Fuzzy logic neurons (Souza, 2018) are functional units able to perform multivariate nonlinear methods. The training algorithm described in for fuzzy neural networks uses clustering for topology description and genetic algorithms for parameter tuning. FNN was studied for the quantitative and qualitative performance comparisons (Wai, Chen, & Liu, 2015). In this article shows that the result of performance comparisons of the FNN control system with various learning rates is more suitable to control the single-stage boost inverter than a double-loop PI control framework, an adaptive control scheme, even though it requires a high computation amount. Furthermore, an online fuzzy neural network (FNN) was designed for velocity sensorless control methodology (Wai, Chen, & Yao, 2016). It could cope with the complex problem of controlling transformation and detailing the parameters in the AOC scheme. It directly ensures the stability of the entire system without the necessity of constraints, detailed system information and assistant compensated controllers despite the permanence of possibilities. Also, Fuzzy Neural Networks was used for the credit rating of the banking system (Astamal & Rahimi, 2019), this article details that a structured model was obtained for determination and categorisation, which is consist of six steps. The results indicate model efficiency compared to bank credit measuring experts.



### Multiple Linear Regression Analysis

Multiple Linear Regression Analysis is a statistical technique that uses several explanatory variables to predict the outcome of a response variable (Astamal & Rahimi, 2019; Hoffmann & Shafer, 2015). The multiple linear regression has the goal of modelling the linear relationship between the explanatory (independent) variables and response (dependent) variable. Multiple Linear Regression Analysis was applied to many pieces of research.

It was applied to monitoring the influence of physical fitness on special judo fitness test performance (Lopes-Silva, Panissa, Julio, & Franchini, 2019), which Multiple Linear Regression Analysis was used to recognize the physical fitness variables that predict the Special Judo Fitness Test (SJFT) performance. However, the result of this research, they had a small ability to predict the SJFT variables.

Besides, Multiple Linear Regression analysis is also used as part of interesting research, Japanese's researchers (Suzuki et al., 2018) applied its to clarify the level of influence and factors that influence bone density according to the survey of health conditions and lifestyle behaviours in dense snow areas. Multiple Linear Regression Analysis is used to calculate their association with bone mineral density and Bayesian network model is used to determine the influence level of each factor as conditional probability. Multiple regression analysis explained that age, sex, fracture, and calcium absorption significantly influenced bone mineral density. The result presented significant factors affecting the presence or absence of fracture experiences and suggesting that enough calcium intake is necessary for preventing bone fractures.

After the literature review, we found that Fuzzy Neural Network (FNN) and Multiple Linear Regression Analysis (MLR) could adapt to predict for the risk of losing student loan.

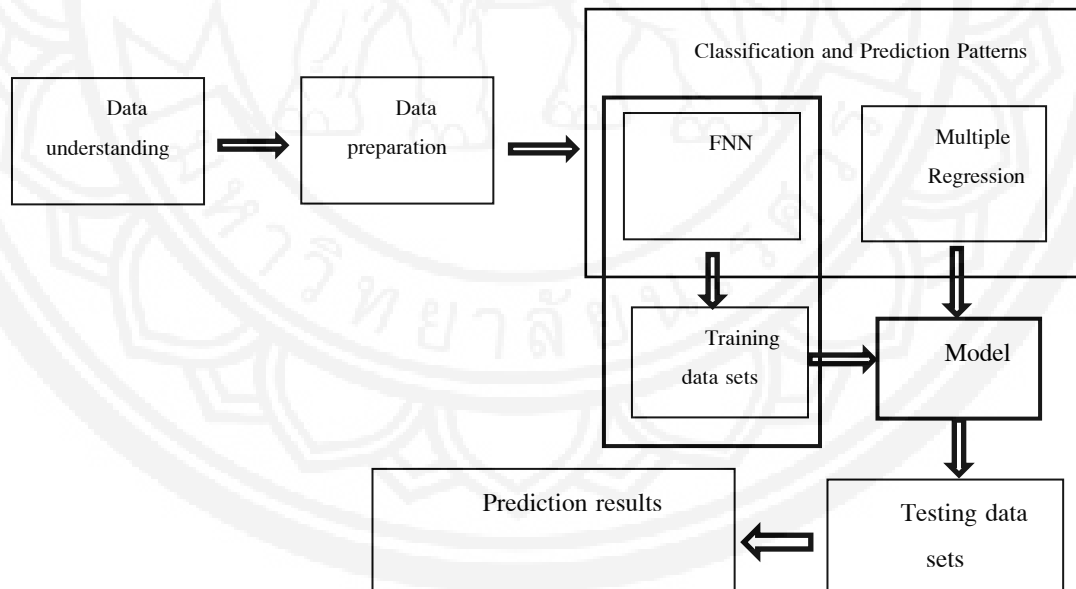


Figure 1 Shows the system architecture

We have defined the concept of this research (Figure 1) to compare between the use of Fuzzy Neural Network techniques and multiple regression techniques for predicting the results of the risk of borrowing the student's rank. We can divide the main work into three parts:

Part 1 is the data collection and selection process by primary and secondary data collection methods, and then select the appropriate imported independent variables by finding the relationship between variables using the correlation coefficient model and multiple linear regression equations.

Part 2 is a comparison of techniques between fuzzy neural networks and linear regression methods to create suitable models for use as rules and embedded code into applications.

Part 3 is the operation of the system to forecast the risk of losing a loan on the application.

#### Data preparation and attribute selection

We collected 296 student samples from four faculties who study in science and technology. For data collection, they were students information who have studied, loaned and graduated. This research uses the IBM SPSS Statistics 22 program to select attributes by finding relationships between variables with Pearson's correlation coefficient technique (as in Equation 1) will use the symbol  $r$  or  $r_{XY}$  to calculate the  $r$  value,  $r$  value or correlation coefficient with the strength of association include 0.01–0.0.9, 0.10–0.29, 0.30–0.49, 0.50–0.69, 0.70–0.89 and 0.90–0.99, with a low level of relationship, low to medium, medium to high, high to very high and almost completely, respectively. Also, the use of independent variables need to test the relationship between independent variables, and independent variables have  $r < 0.75$  or not; if it is less than  $r$ , it can be used. This method is only a test of association in an association, not a causal relationship.

In this research, we found that the relationship between GPAX and Income variables, the correlate is 0.002, in the slightly inverse direction (Trivial). The correlation between Faculty variables and Major variables is 0.956 in almost perfect direction (Near perfect). Moreover, we found that the relationship between the Faculty and Major independent variables and the RiskForPay variable is 0.051 and -0.003, respectively. Details can be shown in Table 1.

$$r = \frac{(1/N-1)(\sum XY - ((\sum X)(\sum Y)/N))}{S_x S_y} \quad (1)$$

$r$  = Correlation coefficient  
 $N$  = Number of samples  
 $S_x$  = Standard deviation x  
 $S_y$  = Standard deviation y

Table 1 shows the relationship between variables with a correlation coefficient

	Faculty	Major	GPAX	Sex	Income	CousineN	StatusHist	RiskForPay
Faculty	1	.956**	.028	-.023	.003	-.023	-.383**	.051
Major	.956**	1	.022	-.042	.027	-.011	-.363**	-.003
GPAX	.028	.022	1	-.109	-.002	.045	.176**	-.114
Sex	-.023	-.042	-.109	1	-.005	-.062	-.157**	-.008
Income	.003	.027	-.002	-.005	1	.096	-.114*	-.006
CousineN	-.023	-.011	.045	-.062	.096	1	.013	-.015
StatusHist	-.383**	-.363**	.176**	-.157**	-.114*	.013	1	.014
RiskForPay	.051	-.003	-.114	-.008	-.006	-.015	.014	1

\*\* Correlation is significant at the 0.01 level (2-tailed).

\*Correlation is significant at the 0.05 level (2-tailed).



### Converting a dummy variable from qualitative data to quantitative data

This research found that the input data are independent variables which are divided into three qualitative variables, Faculty, Income and Sex. Moreover, three quantitative variables include GPAX, cousins and StatusHist, respectively. Therefore, qualitative data is converted to quantitative data. We used the principle of dividing the data, is called *Dummy variables*. For example, the Faculty variable is qualitative data, which can adjust the data to be in the form of quantitative data. The researcher has assigned the value of  $G$  to be the group number of variables. Table 2 shows the details of four group of puppet variables that have the same number of groups as  $G-1$ . Selecting the Archi variable is a reference group and reference. The dependent variable is set to level 0, 1, 2, 3, 4 and 5 with None, very low, Low, Medium, High and very high respectively. Therefore, we have input data to select variables by Multiple regression equations (Multiple regression) as shown in Table 3.

**Table 2** Example of the Dummy variable coding of the Faculty variable

Faculty	Sci	Engi	Aggri
Sci(0)	1	0	0
Engi (1)	0	1	0
Aggri (3)	0	0	1
Archi (2) (Reference group)	0	0	0

**Table 3** Show some data for importing with Multiple Regression Analysis

Sci	engi	aggri	GPAX	Sex	VeryPoor	Poor	Medium	High	cousineN	statusHist	riskForPay
1	0	0	3.5	0	0	0	1	0	0	0	2
1	0	0	2.54	0	0	0	1	0	1	0	3
1	0	0	2.71	0	1	0	0	0	1	0	2
1	0	0	2.64	0	1	0	0	0	0	0	4
1	0	0	2.58	0	0	0	0	0	0	0	1
1	0	0	3.06	0	0	1	0	0	1	0	1
1	0	0	2.01	1	1	0	0	0	1	0	2
1	0	0	2.61	0	0	0	1	0	1	1	0
1	0	0	2.53	0	1	0	0	0	0	1	3
1	0	0	3.01	0	1	0	0	0	0	1	2

**Table 4** Showing some performance values of multiple regression

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error			
1 (Constant)	5.44	8.833		6.540	.000
Sci	-1.24	3.358	-.379	-3.468	.001
Engi	-1.156	.366	-.331	-3.158	.002
Aggri	-1.098	.370	-.318	-2.964	.003
GPAX	-.770	.256	-.185	-3.012	.003
Sex	-.144	.198	-.045	-.729	.016

They analyze the import data using multiple regression methods detail in Table 4. We found that the Standardized Coefficients from this table show only some of the five independent variables. Which, if considering the Beta value, it will be found that the Sci 1 variable has the highest value in the direction of inverse direction with the dependent variable (riskForPay). However, these variables are used to create simulators.

### Fuzzy Attribute Configuration

This research used data mining techniques that use data classification patterns (Data classification) for training and testing data. We propose the use of a neural net technique that has some attributes as a Fuzzy attribute, which the Income variable is a variable that can display the data in a fuzzy manner. For example, high income, low income, income 35K (Bahts). Attributes can access the form of receiving information. Therefore, we propose two types of data formats that are form 1: fuzzy terms such as Very High (VH), High (H), Medium (M), Poor (P) and Very Poor (VP), respectively. Which corresponds to the value defined in trapezoidal type, which defines the scope of data under the variables  $\alpha$ ,  $\beta$ ,  $\gamma$  and  $\delta$  respectively, as shown in Table 5. For the second format, the Crisp value, such as Income, is 25.4K. We have set the independent variable in the form of the fuzzy language. When the user wants to import any independent variable. It can be imported in the form of fuzzy language or Crisp as well, which the system can support both of these values.

**Table 5** Displays the fuzzy language description of the attribute attr<sub>006</sub>: “Income”

fuzzy ID	label name	attrName	type	Distribution Name			
				$\alpha$	$\beta$	$\gamma$	$\delta$
006	Income	Very High	Trapezoidal	0.70	0.75	0.85	1.00
		High	Trapezoidal	0.50	0.55	0.80	0.90
		Medium	Trapezoidal	0.35	0.40	0.60	0.75
		Poor	Trapezoidal	0.20	0.25	0.45	0.55
		Very Poor	Trapezoidal	0.00	0.15	0.30	0.45

### Modeling using multiple linear regression techniques

Modelling with multiple linear regression techniques use independent variables and dependent variables like the modelling of artificial neural networks. The selection of the variables is stepwise selection, which is a method of selecting variables that use the interaction between progressive and backward combinations. We started by selecting variables with the highest correlation coefficient with the highest dependent variable into the equation. Moreover, then choose the variables that are not in the equation. The requirements are variables that can enter the equation using the forward selection method and at the same time, select the variables in the equation. The predictive variables in any equation are likely to be eliminated from the equation using the Backward Selection method.

We can selectively mix both methods at every step until there are no variables that are removed from the equation, and no variables are imported. This equation, we assume the end method and the regression equation with the coefficient of forecasting. Equation 2 is a formula to calculate the linear regression equation. We can calculate the performance using the working model on the RapidMiner Studio 9.2 program, which can show the detailed testing procedures (as shown in Figure 2).

$$Y' = a + b_1X_1 + b_2X_2 + \dots + b_kX_k \quad (2)$$





where  $\hat{Y}$  is predict values of dependent variables (Dependent variable)  
 $a$  is the constant of the predictive equation in the form of raw data  
 $b_1, b_2, \dots, b_k$  is the weight, score or regression coefficient of the forecasting variable from 1 to  $k$  respectively.  
 $X_1, X_2, \dots, X_k$  is the data of forecasting variables (Predictors: Constant) from 1 to  $k$  respectively  
 $k$  is the number of predictive variables.

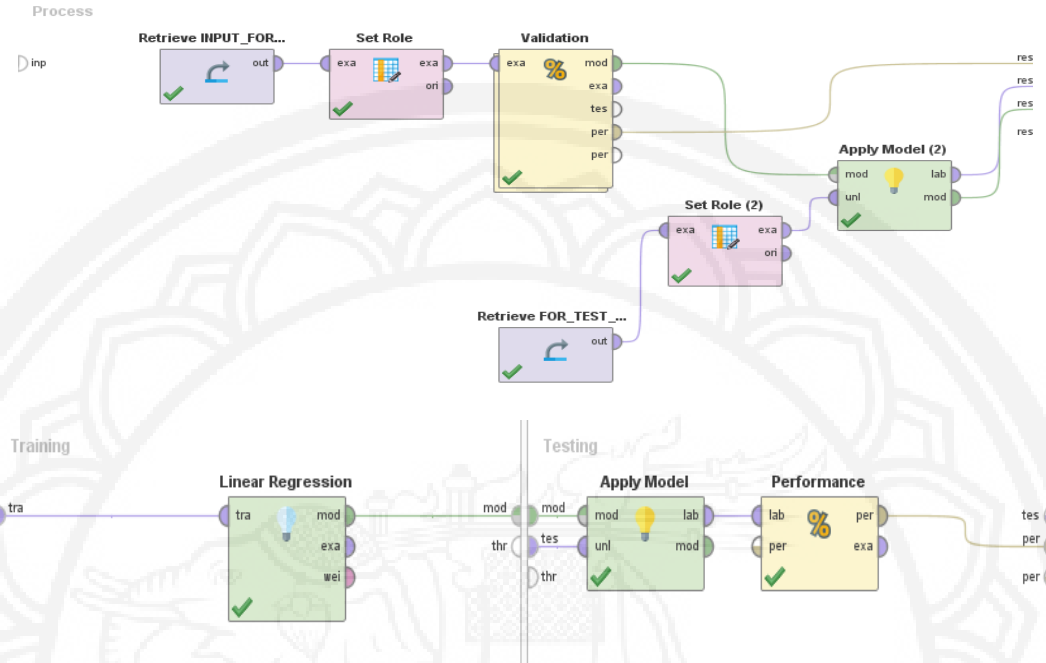


Figure 2 Showing the performance testing process, multiple linear regression equations

### Fuzzy attribute matching

#### Making a Fuzzy attribute matching between Fuzzy linguistic term and Fuzzy linguistic term

Fuzzy linguistic term and Fuzzy linguistic term compare between two fuzzy language terms in the same attribute to find the degree of similarity between any two fuzzy sets, and this research got similar results of the degree between the fuzzy language pairs that were greater than or equal to the thresholds (similarity) of any 2 fuzzy sets. In this research, we have defined the overlap pattern of the member function graph. Set the fuzzy set to be  $A_1$  and  $A_2$  correspond to the centre,  $c_1$  and  $c_2$ , and the boundary point is  $\alpha_i, \beta_i, \gamma_i, \delta_i$  ( $i = 1, 2$ ), where  $c_i = ((\gamma_i + \beta_i))/2$  which assumes that  $c_2 > c_1$  has the following details.

Where  $(\alpha_1 < \alpha_2, \delta_1 > \delta_2, \beta_1 \leq \beta_2, \gamma_2 \leq \gamma_1)$  for this case, Fuzzy attributes  $A_2$  are included in the fuzzy attribute  $A_1$ .  $A_2 \subset A_1$  is calculated (as in Equation 3) by the similarity of the Fuzzy Set, as shown in figure 3.

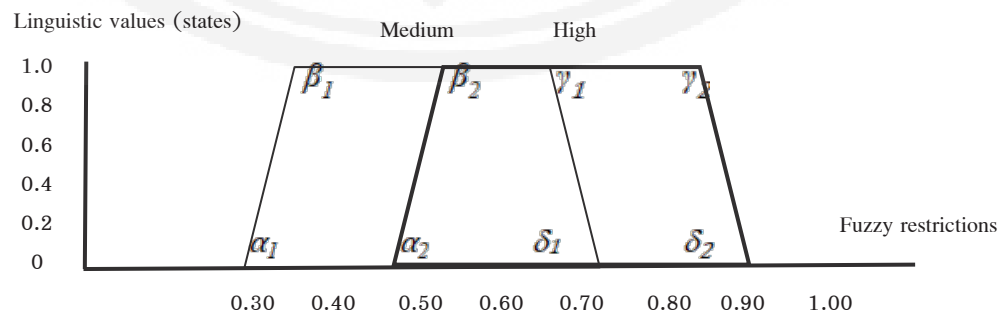


Figure 3 The analogy between fuzzy language terms “Medium” and “High”

$$S = \frac{|A_2|}{|A_1|} = \frac{w_2}{w_1} \quad \text{or} \quad S = \frac{\delta_2 - \alpha_2 + \gamma_2 - \beta_2}{\delta_1 - \alpha_1 + \gamma_1 - \beta_1} \quad (3)$$

where  $w_1 = w_{b_1} + w_{t_1}$ ,  $w_2 = w_{b_2} + w_{t_2}$

#### Making a fuzzy attribute to match between the crisp value and the fuzzy linguistic term

Making Fuzzy attribute will match between Crisp value and Fuzzy linguistic term. In order to know how much the degree of membership, shows the example of finding the value of Crisp to compare with the fuzzy language expression. Which specifying the Crisp value and the fuzzy expression of the Income attribute is 0.54 ( $10^2$  xK (Bahts)) And 'Medium', respectively, can show details in Figure 4.

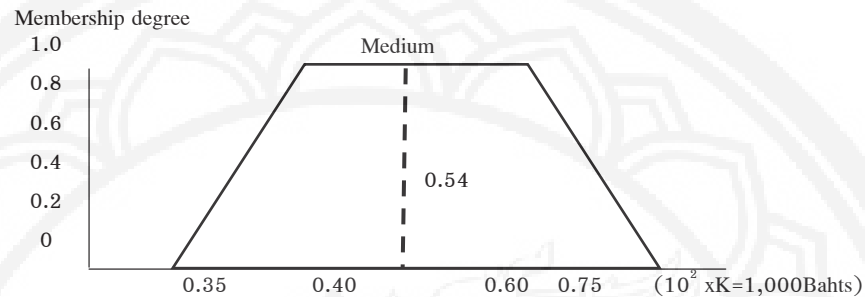


Figure 4 shows the comparison of the Fuzzy attribute Income between the 'Medium and 0.54K

The process of making a match between "Medium" and "0.54". We use data from the member function values are  $\alpha, \beta, \gamma, \delta$  to compare how the "0.54" value has a degree of membership value of the fuzzy language term "Medium". However, the value of the comparison will use conditions as details in formula 4.

$$\mu(x, \alpha, \beta, \gamma, \delta) = \begin{cases} 1 & \text{when } \beta \leq x \leq \gamma \\ 0 & \text{when } x \leq \alpha \text{ or } x \geq \delta \\ \frac{\alpha - x}{\alpha - \beta} & \text{when } \alpha < x < \beta \\ \frac{\delta - x}{\delta - \gamma} & \text{when } \gamma < x < \delta \end{cases} \quad (4)$$

Where x is any Crisp value

We have compared the membership degree between Crisp and the fuzzy expression. If the calculation result is greater than or equal to the threshold value, it will pass the accepted condition. Therefore, the results present that when using the Fuzzy Income attribute, the value of 0.54 compared with other fuzzy terms similar to other, Very High, High, Medium, Poor and Very Poor, equal to 0, 0.8, 1, 0.1 and 0 respectively. Therefore, the Fuzzy Income attribute value of 0.54 can be the closest value to the fuzzy language term "Medium", which we chose to use the weight of the Medium attribute to calculate as well for the artificial neural networks in the model created for forecasting.



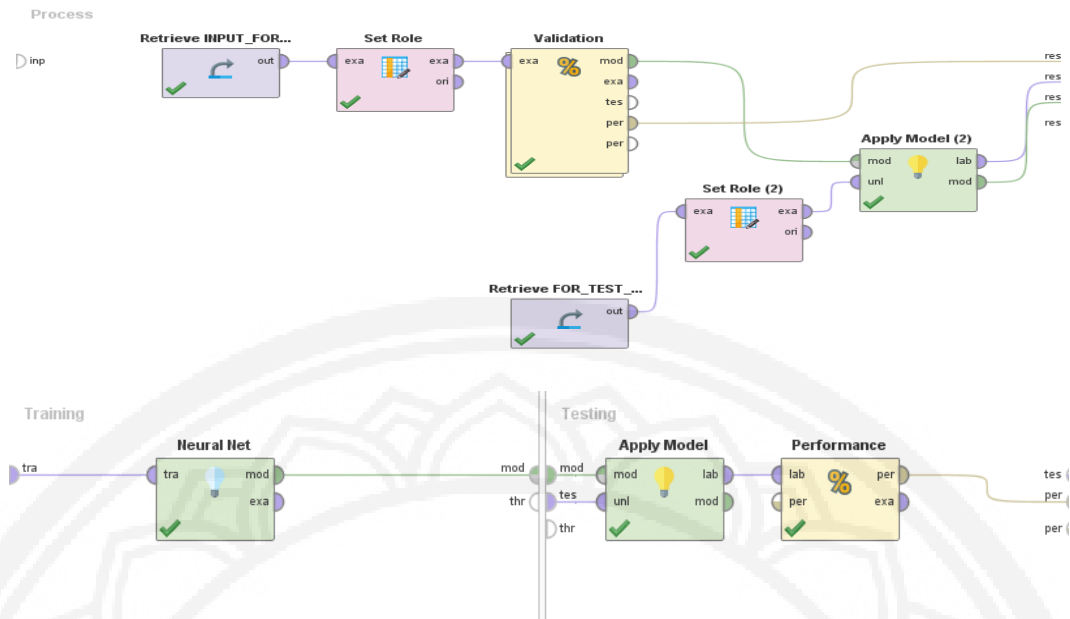


Figure 5 Show the development of cross-model validation fold model 5

### Construction of forecasting models with fuzzy neural network techniques

Selecting a data-sharing model to measure the performance of the model, In this research, we have selected two main data types: Split Test. Which used sample data, 34%, 70% and 80%, using sample data 66%, 30% and 20%, respectively, and Cross-validation test. We set Fold = 5 and Fold = 10, respectively, as shown in Figure 5.

## Results

### Performance testing of fuzzy models, artificial neural networks, and multiple linear regression models

In this research, it was found that the forecasting model of neural networks which the appropriate neural network algorithm has a fuzzy neural network structure, model 8-5-6. We used Income (for example the fuzzy attribute) = Very Poor (VP) as a reference variable in dividing the data in a 5-fold manner, cross-validation, K = 5 momentum, 0.2 learning rate, 0.3, with accuracy, memory, and counterbalance values were 83.33%, 83.10%, 83.30% and 82.40% respectively. There are detailed weight attributes of the model, as shown in Table 6 and Table 7, respectively. For the multiple linear regression model, there is an equation model with the values of sci, engi, aggri, GPAX - 0.851, equal to -0.851, - 0.816, - 0.819 and - 0.689 respectively.

### Calculation of the results of forecasting models

Calculations to find the results of that model will be divided into two steps.

### Fuzzy attribute value conversion

In this step, it is necessary to convert the value of the fuzzy attribute to use the values obtained to calculate the forecast of the model. Table 8 shows only some of the input data that you want to test. We will find that the Fuzzy Income attribute must be calculated using the formula 4. Therefore, calculating Input (2) and Input



(3) values were  $\mu(x=0.54:\alpha=0.35, \beta=0.40, \gamma=0.60, \delta=0.75)$  and  $\mu(x=0.33:\alpha=0.20, \beta=0.25, \gamma=0.45, \delta=0.55)$  which to find the membership levels which are close to "Medium" equal to 1 and "Poor" is equal to 1, respectively.

**Table 6** The weight between a hidden node and an output node

Class 'Low' (Sigmoid)	Class 'Medium' (Sigmoid)	Class 'High' (Sigmoid)
Node 1: -2.690	Node 1: -0.688	Node 1: -1.209
Node 2: 5.552	Node 2: -0.614	Node 2: -12.160
Node 3: -3.705	Node 3: -0.236	Node 3: -4.761
Node 4: -3.638	Node 3: -0.236	Node 4: 3.916
Node 5: 3.082	Node 5: -0.868	Node 5: -4.417
Threshold: -0.451	Threshold: -1.232	Threshold: -2.966
Class 'veryLow' (Sigmoid)	Class 'None' (Sigmoid)	Class 'veryHigh' (Sigmoid)
Node 1: 1.436	Node 1: 1.397	Node 1: 1.161
Node 2: -2.927	Node 2: -0.673	Node 2: -0.192
Node 3: 2.875	Node 3: 1.583	Node 3: -0.093
Node 4: 1.326	Node 4: 0.535	Node 4: 0.732
Node 5: -0.870	Node 5: 0.532	Node 5: -0.045
Threshold: -3.151	Threshold: -2.777	Threshold: -2.414

**Table 7** The weight between the Input node and Hidden node

Attributes	Node1(Sigmoid)	Node2(Sigmoid)	Node3(Sigmoid)	Node 4(Sigmoid)	Node 5(Sigmoid)
Sci	1.108	0.381	6.511	-8.500	0.254
Engi	-4.994	-0.847	2.169	1.670	6.113
Aggri	3.806	-0.314	-1.842	-0.077	7.691
GPAX	-2.372	6.9761	2.906	8.434	2.353
Sex	-6.841	-0.664	4.639	4.078	4.184
Income	6.783	8.669	9.542	-2.920	-11.275
CousineN	3.972	-5.261	3.651	-9.078	5.077
StatusHist	-0.762	0.561	-1.6691	0.417	6.144
Threshold	-6.532	-6.050	-3.864	-4.413	5.753

**Table 8** Samples of losses student loan forecasts with TFNN technique and Multiple Linear Regression

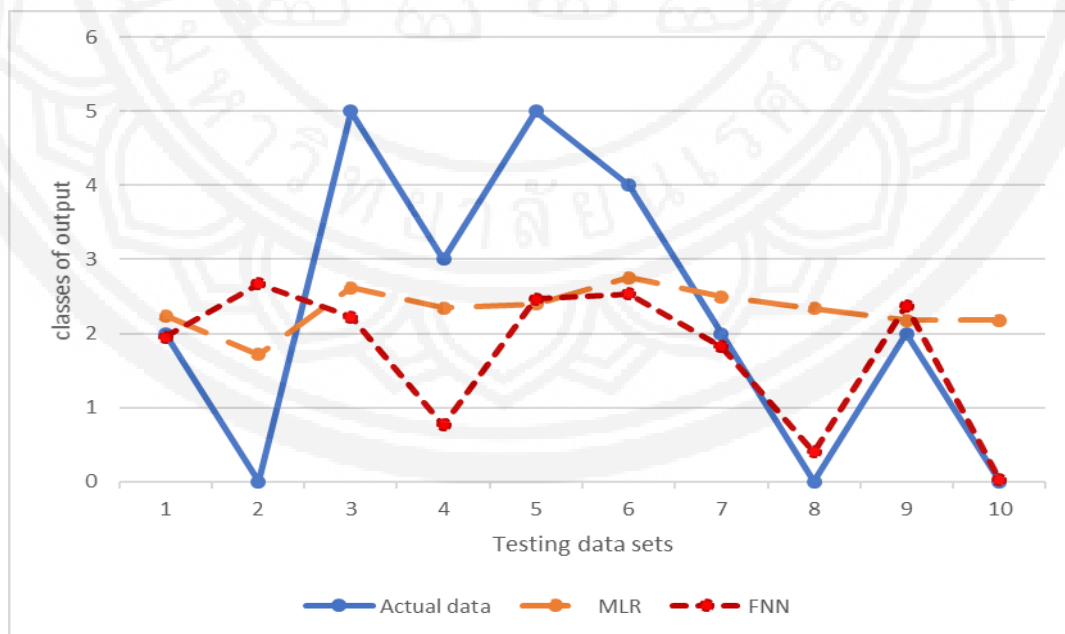
	Input (1)	Input(2)	Input(3)	Input(4)	Input(5)
Sci	1	0	0	0	1
Engi	0	0	1	1	0
Aggri	0	1	0	0	0
GPAX	2.55	2.8	3.25	2.78	2.09
Sex	1	1	0	1	0
Income (Fuzzy attribute)	High	0.54 ( $\cong$ Medium)	0.33( $\cong$ Poor)	Low	Medium
CousineN	2	1	1	1	0
StatusHist	1	1	0	0	0
...	...	...	...	...	...



Therefore, the calculated value can be concluded that the sample data imported into the forecasting system using fuzzy neural network techniques. The value of the RiskForPay attribute is approaching, Sigmoid node  $Low = x_{Low}$  which is set to "Low" level. Therefore, the study results found that, when comparing the work between the FNN and MLR, it is found that both methods use the same principle to convert the fuzzy attributes to be used as input for testing. However, testing results were found that using the FNN model, will give more accuracy by comparing the results as shown in Table 9 and Figure 6, respectively. However, the work obtained from the FNN model is necessary to improve the use of fuzzy data to be appropriate and have a variety of imported data to increase the efficiency of forecasting accuracy.

**Table 9** Prediction results of FNN and MLR

The FNN model		The MLR		The actual data	
Generated data		Generated data			
Class	Value	Class	Value	Class	Value
Low	1.950	Low	2.236	Low	2
Medium	2.672	Low	1.720	None	0
Low	2.221	Medium	2.620	veryHigh	5
veryLow	0.764	Low	2.344	Medium	3
Medium	2.463	Low	2.396	veryHigh	5
Medium	2.535	Medium	2.755	High	4
Low	1.826	Medium	2.496	Low	2
None	0.399	Low	2.338	None	0
Low	2.378	Low	2.177	Low	2
None	0.022	Low	2.176	None	0



**Figure 6** Showing the results of the comparison between fuzzy artificial networks and multiple regression linear equations



### Conclusion

This research presents a comparison of modelling in the form of FNN and MLR for predicting the risk of losing student loan. They are presented starting from the management of fuzzy attributes, including screening data for importing to create models from the guidelines. We found that the FNN model with the structure 8-5-6 with the correct value Accuracy, memory, and counterbalance values were 83.33%, 83.10%, 83.30% and 82.40% respectively. It has a higher performance than the MLR model with the Performance vector value, that is the root mean squared error value: 1.513 +/- 0.000 and the squared correlation value: 0.081 respectively. However, the model of the forecasting model works, and we found that the configuration of the Fuzzy Attributes which they can work under Income; is the only one defined in a fuzzy way. It is difficult to determine the accuracy. In this research, relying only based on the average configuration and then dividing into phases may have an impact on the calculation error. Therefore, it is vital to find ways to determine the fuzzy level to be more reliable. In addition, the input factors for data training and gathering more information for data storage are complicated to conduct research that must be relevant to many people. The department in the government sector is the university unit which agency responsible for the funding management, including the private sector related to various banks. If these can control in the research; it may lead to being the efficient prediction that can be accepted and applied according to the objectives of various businesses.

### References

- Astamal, M. A., & Rahimi, R. (2019). Designing an Expert System for Credit Rating of Real Customers of Banks Using Fuzzy Neural Networks. *Advances in Mathematical Finance and Applications*, 4(1), 89-102. doi:10.22034/amfa.2019.577561.1128
- Haykin, S., (2009). *Neural networks and learning machines* (3rd ed.). Retrieved from [https://cours.etsmtl.ca/sys843/REFS/Books/ebook\\_Haykin09.pdf](https://cours.etsmtl.ca/sys843/REFS/Books/ebook_Haykin09.pdf)
- Hoffmann, J. P., & Shafer, K. (2015). *Linear regression analysis*. Washington, DC: NASW Press.
- Lopes-Silva, J. P., Panissa, V., Julio, U. F., & Franchini, E. (2019). *Influence of Physical Fitness on Special Judo Fitness Test Performance*. Retrieved from [https://www.researchgate.net/publication/328876361\\_Influence\\_of\\_Physical\\_Fitness\\_on\\_Special\\_Judo\\_Fitness\\_Test\\_Performance\\_-\\_A\\_Multiple\\_Linear\\_Regression\\_Analysis](https://www.researchgate.net/publication/328876361_Influence_of_Physical_Fitness_on_Special_Judo_Fitness_Test_Performance_-_A_Multiple_Linear_Regression_Analysis)
- Mendel, J. M. (1995). Fuzzy logic systems for engineering: a tutorial. *Institute of Electrical and Electronics Engineers*, 83(3), 345-377. doi: 10.1519/JSC.00000000000002948
- Osowski, S., & Linh, T. H. (2001). ECG beat recognition using fuzzy hybrid neural network. *Institute of Electrical and Electronics Engineers*, 48(11), 1265-1271. Retrieved from <https://ieeexplore.ieee.org/document/959322>
- Souza, P. V. C. (2018). Regularized fuzzy neural networks for pattern classification problems. *International Journal of Applied Engineering Research*, 13(5), 2985-2991. Retrieved from [https://www.ripublication.com/ijaer18/ijaerv13n5\\_122.pdf](https://www.ripublication.com/ijaer18/ijaerv13n5_122.pdf)



- Suzuki, T., Shimoda, T., Takahashi, N., Tsutsumi, K., Samukawa, M., Yoshimura, S., & Ogasawara, K. (2018). Factors Affecting Bone Mineral Density Among Snowy Region Residents in Japan: Analysis Using Multiple Linear Regression and Bayesian Network Model. *Interactive Journal of Medical Research*, 7(1), 10. doi: 10.2196/preprints.8555
- Tang, J., Liu, F., Zhang, W., Ke, R., & Zou, Y. (2018). Lane-changes prediction based on adaptive fuzzy neural network. *Expert Systems with Applications*, 91, 452–463. doi.org/10.1016/j.eswa.2017.09.025
- Wai, R.-J., Chen, M.-W., & Liu, Y.-K. (2015). Design of adaptive control and fuzzy neural network control for single-stage boost inverter. *Institute of Electrical and Electronics Engineers*, 62(9), 5434–5445. Retrieved from <https://ieeexplore.ieee.org/document/7054543>
- Wai, R.-J., Chen, M.-W., & Yao, J.-X. (2016). Observer-based adaptive fuzzy-neural-network control for hybrid maglev transportation system. *Neurocomputing*, 175, 10–24. doi:10.1016/j.neucom.2015.10.006
- Wang, G., Hao, J., Ma, J., & Huang, L. (2010). A new approach to intrusion detection using Artificial Neural Networks and fuzzy clustering. *Expert Systems with Applications*, 37(9), 6225–6232. doi: 10.1016/j.eswa.2010.02.102