23rd Annual Conference of the
Transportation Science Society of the Philippines
Quezon City, Philippines, 8 August 2016

Comparison of Logit and Neural Network Models in Inter-Island Discrete Choice Analysis

Krister Ian Daniel Z. ROQUEL
Graduate Student
Civil Engineering Department
De La Salle University
2401 Taft Avenue, Manila
E-mail: krister.roquel@dlsu.edu.ph

Alexis M. FILLONE
Full Professor
Civil Engineering Department
De La Salle University
2401 Taft Avenue, Manila
E-mail: alexis.fillone@dlsu.edu.ph

Abstract: Logit-based models have often been used for discrete choice analysis. However, conventional logit models preserve a linear relationship that requires variables that are independent of each other, which is generally not the proper assumption. In this paper, the researcher addresses the non-linear behavior and inter-dependence of variables using neural networks in modeling inter-island travel choice. The researcher employed neural network analysis to a previous work to test the applicability of neural network in discrete choice models for inter-island travel. It was found that the neural network model is a good fit in describing travel choice behavior, while the logit model is more inclined to model the decision making process. Also, it was found that the neural network model is also capable of accurately predicting the minority, which has long been a problem when using logit models, as these are usually treated as errors.

Key words: discrete choice, multinomial logit, neural network, inter-island travel

1. INTRODUCTION

Logit-based models have often been used for discrete choice analysis. These are based on the random utility theory, which employs an abstract measurement of the degree of satisfaction for any choice an individual makes, with the assumption that rational people act to maximize their utility. However, conventional logit models preserve a linear relationship that requires variables that are independent of each other, which is generally not the proper assumption. In this paper, the researcher would like to address the non-linear behavior and inter-dependence of variables using neural networks in modeling inter-island travel choice.

Previous works on the application of neural networks on discrete choice behavior have shown potentials and advantages of employing neural networks over the traditional logit models. As early as the late 1990s, Nijkamp, et al. (1996) conducted a study on the comparison of neural network and logit analysis in modeling inter-urban transport flows. Bentz and Merunka (2000), Hensher and Ton (2000), Cantarella and Luca (2002), Vythoulkas and Koutsopoulos (2003), Norets (2008), Nakayama, et al. (2008), and Dia (2010) all have contributions on the field with their respective researches on using artificial neural networks on discrete choice applications. Even until recently, Pulugurta, et al. (2013) still conducts studies on the comparison of the models developed using various approaches.

As choice decisions usually involve approximations that are not precisely captured by logit models, neural network models would always have a place in discrete choice analysis due to their capability of function inference based on observations. The latter does not need any prior knowledge of the characteristics of the variables and can account for non-linearity, which makes for an easier and more convenient model development process. In this paper, the researcher employs neural network analysis to a previous work (Roquel, 2013), to test the applicability of neural network in discrete choice models for inter-island travel

2. RELATED LITERATURE

The basic concept of characterizing the traveller's mode choice is taken as that travellers would choose the option which maximizes their utility for every unit cost they pay, as described by the economic consumer theory. This theory provides the means for the transformation of assumptions about desires into a demand function expressing the action of a consumer under given circumstance. Furthermore, the random utility theory assumes that when facing a choice situation, individuals assign random utilities, based on their personal preferences, to each alternative considered and then choose the alternative having the highest derived utility. In actual modeling analysis, the aim is to express the utility of an alternative as a function of the attributes of the alternative and the tastes and socio-demographic attributes of the decision-maker (Hess, 2005). The individual's derived utility can be decomposed into three components given in the following equation.

$$U_n(j) = \theta_{jn} + \beta_n X_{jn} + \varepsilon_{jn} \tag{1}$$

where θ_{jn}: intrinsic utility of alternative j for individual n,
β_n: vector of parameters estimated for an individual n,
X_{jn}: vector of the attributes of alternative j for individual n, and
ε_{jn}: random error

The function $f(\beta_n, X_{jn})$ is free from any prior assumptions allowing linear formulation in the area of discrete choice modelling, such that the observed utility shall be simply β_{nXjn} (Rajaonarison, Bolduc & Jayet, 2005). The inclusion of the random error, or unobserved utility, means that the deterministic choice now becomes probabilistic, leading to a random utility model. With this, the alternative with the highest observed utility shall have the highest probability of being chosen. In multinomial logit modeling, the probability equation can be written as equation (2),.

$$P(j) = \frac{\exp(U_j)}{\sum_{j'=1}^{J} \exp(U_{j'})} \tag{2}$$

where P(j): probability of choosing mode j

Given the requisite data, a logit model can be estimated that assigns a probability to an individual n travelling from origin A to destination B, choosing mode j. The model shall be able to capture the relevant variables that affect the utility, or benefit, of choosing a particular transport mode (Ewing, Schroeer & Greene, 2004).

On the other hand, in neural network modeling, computational models are constructed to simulate the processing and learning functions of a human brain (Walczak & Cerpa, 2003). The general neural network structure (See Figure 1) is formed by a group of parallel, processing elements named neurons. Neurons in a certain layer of the neural network are connected to those from the previous layer by a number of weighted connections. In addition, there is an extra weight named bias, which is summed to the rest of input weights (Al-Zoubi, et al., 2007).

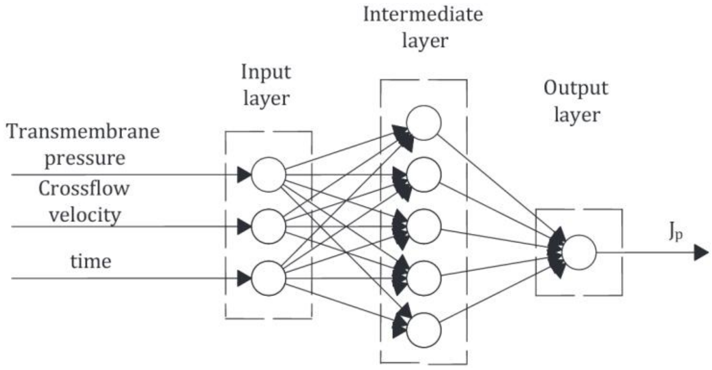


Figure 1. Artificial Neural Network Structure

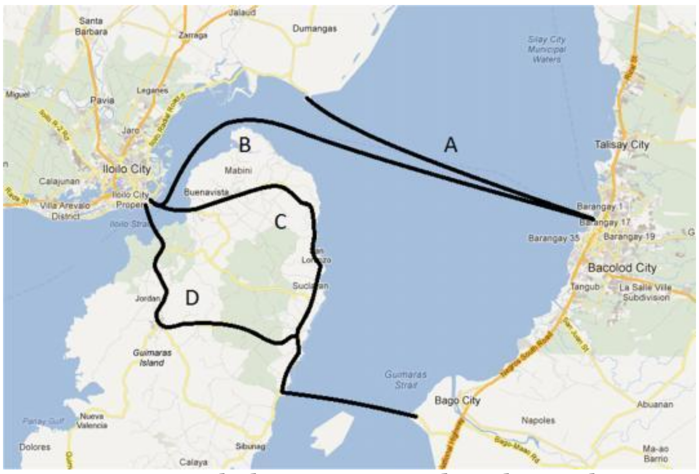
Neurons are usually distributed in different layers, according to input, intermediate (hidden), and output layers (Chen & Kim, 2006; Razavi, Mortazavi, & Mousavi, 2004). Thus, the output of a certain layer acts as an input signal for the neurons in the following layer. In order to calculate an output of a neuron, a transfer function is required for its net input to be transformed. As a consequence of all these connections, the learning process can be fitted by selecting the optimal combination of neurons and weights for each studied system.

In feed-forward neural networks, signals are sent forward, and then the errors are propagated backwards. The back propagation algorithm uses supervised learning, which means that algorithms are provided with examples of the inputs and outputs for the network to compute, and then the errors (the difference between the actual and expected results) are calculated. The idea of the backpropagation algorithm is to reduce this error, until the neural network learns the training data. The training begins with random weights, and the goal is to adjust them so that the error will be minimal (Gershenson, 2003).

3. STUDY AREA

The data used in the study were gathered from terminals serving the inter-island network in the heart of the Visayan region in the Philippines. Major contributors to inter-island traffic in the region are the provinces of Iloilo and Negros Occidental, which are two highly urbanized provinces with populations of over 2.2 M and 2.9 M, respectively (NSO, 2009). Figure 2 shows the location of the study region, while Figure 3 shows the inter-island travel options currently available to the public. As shown, inter-island travel can be done in four ways (A, B, C, and D) in this travel network.

Figure 2. Location of Iloilo and Negros Occidental Provinces



Figrue 3. Major Iloilo-Negros Occidental Travel Routes

With an average of 140 trips per week, the Fastcraft ferry (Route B) caters to most of the demand. RORO (roll-on, roll-off) ferry travel, on the other hand, which offers around 100 trips per week on the average, serves as an effective alternative (Route A). This travel can also be made through inter-modal travel through the island of Guimaras. Iloilo-Guimaras passenger travel can be done using pumpboats, embarking from Iloilo City and alighting at either Buenavista (Route C) or Jordan (Route D). Port-to-port transportation across Guimaras island can be made through jeepneys, multicabs, and vans. Guimaras-Negros Occidental travel can then be performed using pumpboats from San Lorenzo to Pulupandan, completing the Iloilo-Negros travel.

The basic travel options for the Iloilo City to Negros Occidental travel can be summarized into various categories as shown in Table 1. Based on the data shown, it can be seen that a great deal of the inter-island travelling population, 70.56%, uses the fastcraft ferry option (Route B).. This option has the shortest total travel time and does not involve intermodal transfers. However, this option is the most expensive among all options, costing around more than twice the total travel costs incurred using the nearest alternative. This can be interpreted to mean that the travelling population prioritizes travel time and comfort, in terms of the number of transfers, greatly over travel cost.

Table 1. Iloilo-Negros Occidental Inter-Island Travel Options (Daily Basis)

<table>
<tr><th>Route</th><th colspan="2">Transport Mode</th><th>Ave No. of Pass. Per Trip</th><th>Ave No. of Trips</th><th colspan="2">Ave Travel Time [Hour]</th><th colspan="2">Travel Fare (per pax) [Php]</th><th>Transfer</th><th>Users</th></tr>
<tr><td rowspan="3">A</td><td colspan="2">PUJ, Van, Multicab [A-1]</td><td>-</td><td>-</td><td>1</td><td rowspan="3">3.65</td><td>25</td><td rowspan="3">130</td><td rowspan="3">2</td><td rowspan="3">1805 (28.40%)</td></tr>
<tr><td colspan="2">Tricycle [A-2]</td><td>-</td><td>-</td><td>0.5</td><td>25</td></tr>
<tr><td colspan="2">RORO (Roll-on Roll-off) [A-3]</td><td>95</td><td>19</td><td>2.15</td><td>80</td></tr>
<tr><td>B</td><td colspan="2">Fastcraft Ferry</td><td>195</td><td>23</td><td colspan="2">1.5</td><td colspan="2">335</td><td>0</td><td>4485 (70.56%)</td></tr>
<tr><td rowspan="5">C/D</td><td rowspan="3">Pump boat</td><td>[C-1]</td><td>41</td><td>140</td><td>0.33</td><td rowspan="5">3.88 / 3.5</td><td rowspan="2">14</td><td rowspan="5">154 / 134</td><td rowspan="5">3</td><td rowspan="5">66 (1.04%) {See note}</td></tr>
<tr><td>[D-1]</td><td>45</td><td>150</td><td>0.25</td></tr>
<tr><td>[C/D-3]</td><td>33</td><td>2</td><td>0.75</td><td>60</td></tr>
<tr><td rowspan="2">PUJ, Van, Multicab</td><td>[C-2]</td><td>-</td><td>-</td><td>2.75</td><td>80</td></tr>
<tr><td>[D-2]</td><td>-</td><td>-</td><td>2.5</td><td>60</td></tr>
<tr><td rowspan="2">Legend:</td><td colspan="3" rowspan="2">- : Value does not affect the numbers being studied</td><td>Ave</td><td colspan="2">3.1325</td><td colspan="2">188.25</td><td colspan="2" rowspan="2"></td></tr>
<tr><td>St. Dev.</td><td colspan="2">0.842</td><td colspan="2">98.395</td></tr>
</table>

Note: Value of "66 (1.04%)" is based on the assumption that all users of travel option C/D-3 originally came from Iloilo City. Otherwise, use value of "7 (0.11%)", based on the statistic that only 1 out of 10 of those using option C/D-3 originally came from Iloilo, in accordance to the statement made by the officiating body at the wharf hosting the said travel option.

4. MODEL DATA

The variables were categorized into a total of 11 categories to simplify the descriptions of the variables, as shown in the Appendix. Also shown, the travel choices were reduced to A, B, and, C, where options C and D were merged into one as almost no data was gathered for the latter.

5. LOGIT MODELING

In the development of models, all modeling variables were used in different combinations to come up with the best models possible. In evaluating which models are suitable in describing the travel mode choice of the travelling population, many criteria were considered. First, the coefficients of the variables were checked if the sign (positive or negative) agrees with prior knowledge, considering what quantity the variable is representing (utility or disutility). Furthermore, the coefficients' statistical significance are checked through its respective P-values, log likelihood functions, and Rho-squared measures. Lastly, accuracy of models in predicting the travel choice was considered.

The following multinomial logit (ML) models, with a logit structure shown in Figure 4, were developed using only three options, A, B and C, using Alternative C as the base alternative. Using the logit models, the probability of an individual to choose a particular alternative can be computed using equation (3).

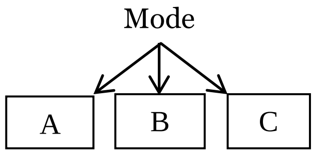


Figure 4. Multinomial Logit Structure

$$P(j) = \frac{e^{U_j}}{e^{U_A} + e^{U_B} + e^{U_C}} \tag{3}$$

Where:
- U_j : utility of alternative j
- U_A : utility of alternative A or RORO Ferry
- U_B : utility of alternative B or Fastcraft Ferry
- U_C : utility of alternative C or Guimaras Intermodal Trip

Table 2. Multinomial Models Developed with its Variables

Variables	**Base Model**	**ML1**	**ML2**	**ML3**
	Coefficient	**Coefficient**	**Coefficient**	**Coefficient**
A_A	1.61425**	.47441	-1.05824**	3.73793**
A_B	2.20655**	-.87900*	1.76049**	5.23193**
TOTCOST		-.00559**		
TOTTIME		-.48424**		
COMFORT		3.9924**		1.20335**
LNDTIME			-.01342**	
C_TVEH			-.01238**	-.01043**
T_ORPR				-.03467**
WAITTME				.00956**
T_PRDE				-.01168**
AxINC1		.00011**	.00011**	.00011**
AxAGE1		-.05496**	-.05293**	-.05304**
BxINC2		.00013**	.00013**	.00013**
BxAGE2		-.05735**	-.05377**	-.05584**
Goodness of Fit Measures				
$L(\bar{\beta})$	-1068.047	-788.7015	-782.8576	-634.8718
L(0)	-	-1377.6598	-1377.6598	-1377.6598
$-2[L(0)-L(\bar{\beta})]$	-	1177.9166	1189.6044	1479.576
$-2[L(C)-L(\bar{\beta})]$	-	558.691	570.3788	860.3504
ρ^2	-	0.42751	0.43175	0.53699
$\bar{\rho}^2$	-	0.26155	0.26702	0.40277

* -passed the 0.1 level of significance   ** - passed the 0.05 level of significance

As seen in Table 2, for the ML1 model, TOTTIME, TOTTIME and COMFORT were used as alternative-specific deterministic variables, while LNDTIME and C_TVEH were used in model ML2, and T_ORPR, WAITTME, T_PRDE, C_TVEH and COMFORT for model ML3. For all three models ML1, ML2 and ML3, AGE and INCOME were used as generic deterministic variables. Going over the coefficients, it can be seen that TOTCOST, TOTTIME, LNDTIME, C_TVEH, T_ORPR, and T_PRDE have negative signs, meaning the items are considered disutilities, which follows priori knowledge since these consider values spent by the individual. For the variables COMFORT and WAITTME, the coefficients are positive. As for INCOME and AGE, the coefficients have consistent positive and negative signs, respectively.

Quantities involving cost and time being significant were expected with the common understanding of travel mode choice scenario. These variables involve quantities that are most directly connected to the choice situation as these are directly spent by the individual as a choice decision is made. Comfort being significant with a positive coefficient was also expected. Income was also found to be statistically significant with a positive coefficient. This can be explained simply as the enabling effect of income. People with higher income are less sensitive to higher costs and are capable to pay more, in exchange for other benefits like shorter travel time and/or higher comfort, among others.

Age, in general, was found to be significant, with negative coefficients. This indicates that older people are more likely to use the intermodal option passing through Guimaras province, even though it has significantly higher travel time as compared with the other two alternatives. This can be interpreted to mean that sensitivity to travel time decreases as an individual gets older. This may also be connected to older people being less in a hurry, being less constrained by their schedules. Another possible explanation is the automation of choice decision through practice, where people would prefer using the alternative they had been using long before, for example, in a time where the other two relatively newer options were still unavailable.

6. NEURAL NETWORK MODELING

In the development of neural network models, the variables were included in sets, to simplify the possible combinations of variables. Table 3 shows the variable categories included in each input data set used, where “1” corresponds to the set being included, and “0”, otherwise. As shown, set A has the least number of input variables at a total of only 6 (comprised of 3 travel experience variables and 3 passenger personal information variables), while set P has the most at 74 variables including all the variables available.

Table 3. Variables for Input Data Set

	Variable Category	Trip Purpose	Passenger Travel Information	Travel Experience	Travel Choice Information	Access Information	Egress Information	Others	Passenger Personal Information	Other Passenger Personal Information	Other Passenger Financial Information	General Information	Total Number of Variables
No. of Variables		5	5	3	9	9	9	4	3	3	6	18	
Input Data Set	A	0	0	1	0	0	0	0	1	0	0	0	6
	B	0	1	0	0	0	0	0	1	0	0	0	8
	C	0	0	0	1	0	0	0	0	0	0	0	9
	D	0	1	1	0	0	0	0	1	0	0	0	11
	E	0	0	0	1	0	0	0	1	0	0	0	12
	F	1	1	1	0	0	0	0	1	0	0	0	16
	G	0	1	0	1	0	0	0	1	0	0	0	17
	H	1	1	0	1	0	0	0	1	0	0	0	22
	I	1	1	1	1	0	0	0	1	0	0	0	25
	J	1	1	1	0	0	0	0	1	0	0	1	34
	K	1	1	1	1	1	0	0	1	0	0	0	34
	L	1	1	1	1	0	1	0	1	0	0	0	34
	M	1	1	1	1	0	0	0	1	0	0	1	43
	N	1	1	1	1	1	1	0	1	0	0	0	43
	O	1	1	1	1	1	1	1	1	1	0	1	68
	P	1	1	1	1	1	1	1	1	1	1	1	74

As for the target data, three (3) output nodes were set, corresponding to the choice among the three travel choices available (A, B, and C), where a value of “1” corresponds to the passenger’s choice, and “0” for the other options that were not chosen. The Levenberg-Marquardt algorithm was used as the training function, as it enables the network to find the solution even if it starts very far off the minimum. The gradient descent method was used for the adaption learning function. The performance function used was the mean square error (MSE). The number of hidden layers was set to only one to come up with just a simple neural network.

The activation function used in the input-hidden connection was hyperbolic tangent, while sigmoidal logistic was employed in the hidden-output connection. This was done to account for the negative contributions of some variables within the network, but end with a strictly positive output, as the target output is only “1” or “0”. In testing to find the best combination of input data and number of hidden neurons, neural network models were developed while varying the number of hidden neurons for each input data set. Figure 5 shows its graphical representation of the R^2 values of the neural networks developed, while Table 4 shows the details.

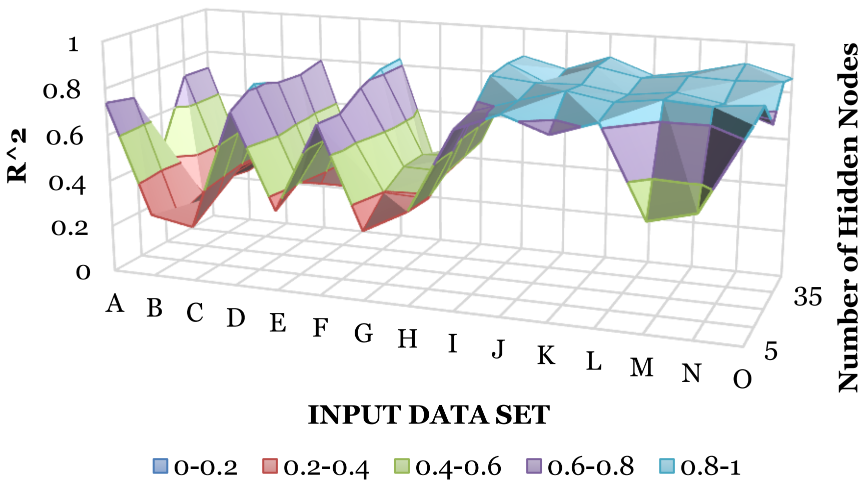


Figure 5. R^2 Performance of Neural Networks

Table 4. R^2 Performance of Neural Networks

Input Data Set	Number of Hidden Neurons				
	5	15	25	35	45
A	0.738551	0.714211	0.438655	0.725819	0.722636
B	0.267713	0.203022	0.248054	0.261878	0.241258
C	0.234643	0.31351	0.388565	0.383347	0.377549
D	0.744217	0.817523	0.779248	0.770621	0.795272
E	0.342775	0.409037	0.346026	0.279312	0.277824
F	0.724133	0.716291	0.782765	0.822667	0.8281
G	0.296546	0.397316	0.453899	0.454155	0.291665
H	0.395918	0.390238	0.449892	0.464592	0.480915
I	0.747533	0.77171	0.874973	0.886309	0.877257
J	0.823992	0.831197	0.858384	0.803013	0.831598
K	0.758833	0.883487	0.864175	0.89842	0.897453
L	0.822032	0.851634	0.836201	0.861945	0.829247
M	0.453414	0.876358	0.860238	0.883356	0.861407
N	0.503972	0.863896	0.827627	0.877145	0.907847
O	0.856476	0.88298	0.761989	0.887402	0.85855
P	0.846639	0.835762	0.761658	0.8904	0.90117

As shown, sets A, D, F, and I onwards have considerably reliable R^2 values. Going back to Table 3, it can be seen that the similarity of these input data sets is the inclusion of travel experience variables. In set A, where only travel experience and passenger personal information were used, the R^2 values only reached a little over 0.7. If variables on passenger travel information were added, as shown in set D, the R^2 values reached 0.8 when the number of hidden neurons was set at 15. When variables dealing with trip purpose were added, the R^2 also attained values over 0.8, but needed more hidden neurons and iterations. Furthermore, when variables are added, the R^2 values tend to show a slight increase, but require significantly longer time for network development. Thus, to have a simple, yet still statistically reliable model, the choices were cut down to sets A, D, and F. Table 5 shows a summary of the variables included in these sets, as well as the R^2 values for the training, validation, and testing of the best neural networks using sets A, D, and F, respectively.

Table 5. Best Neural Networks Developed

		6-5-3 NN		**11-15-3**		**16-45-3**	
Variables		Travel Experience	Used_A Used_B Used_C	Travel Experience	Used_A Used_B Used_C	Travel Experience	Used_A Used_B Used_C
		Passenger Personal Information	Age Gender Income	Passenger Personal Information	Age Gender Income	Passenger Personal Information	Age Gender Income
				Passenger Travel Information	Num_grp Chl_grp Frequency Beflunch Wkday	Passenger Travel Information	Num_grp Chl_grp Frequency Beflunch Wkday
						Trip Purpose	Purwork Purvaca Purschl Purbusi Purhome
Input Nodes		6		11		16	
Hidden Neurons		5		15		45	
R^2	Training	0.73630I1		0.810594		0.848867	
	Validation	0.751793		0.828082		0.819496	
	Testing	0.735975		0.85705		0.742958	
	All	0.738551		0.820129		0.828100	

Following the guidelines in the appropriate number of hidden neurons, the three networks were evaluated. The first condition sets the maximum number of hidden neurons to be twice the number of input nodes plus one. The 16-45-3 model does not satisfy this condition ($2(16) + 1 = 33 < 45$), and is thus, removed. The second guideline states that the number of hidden neurons should be between the average number of input and output nodes and their sum. Both the 6-5-3 and 11-15-3 models satisfy the first part of this condition, but only the 11-15-3 model fails the next ($11 + 3 = 14 < 15$). However, as the R^2 value of the 6-5-3 model is relatively low, and since the 11-15-3 model only slightly failed to satisfy the guidelines, the latter was chosen as the better model.

To determine the optimum number of hidden neurons, neural network models were developed while varying the number of hidden neurons from 5 to 25. Figures 6 and 7 show the R^2 and mean square error performances of the models, respectively. As shown, the highest R^2 values for training, validation, and testing were attained when the number of hidden neurons was at 15. Also shown, the lowest mean square error was reached with 15 hidden neurons. Thus, this paper recognizes the 11-15-3 neural network (i.e., 11 input variables; 15 hidden neurons; 3 output nodes) as the best model to describe the discrete choice behavior being studied. Figure 8 shows the structure of the best model, while Figure 9 shows its R plots.

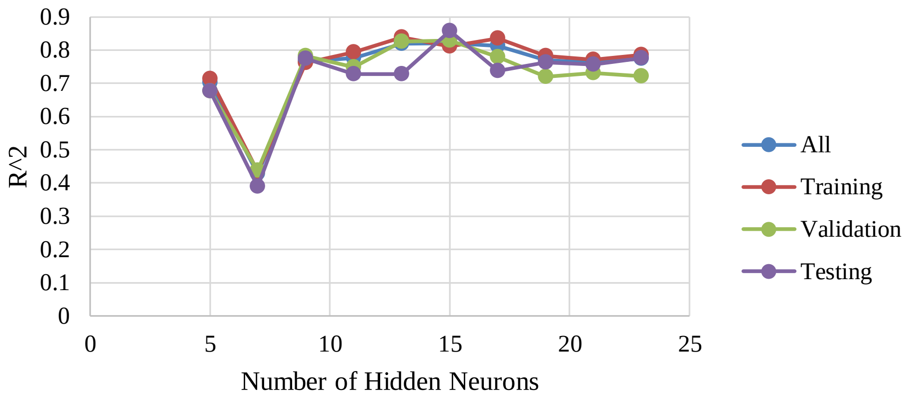


Figure 6. R^2 Performance of Set D Neural Networks

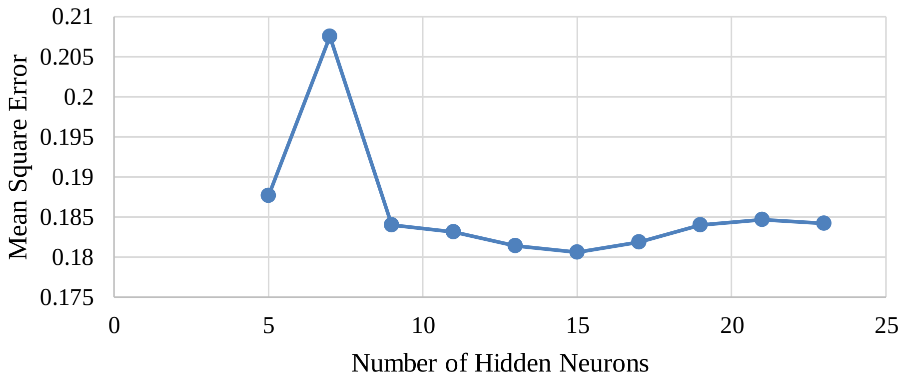


Figure 7. Mean Square Error of Set D Neural Networks

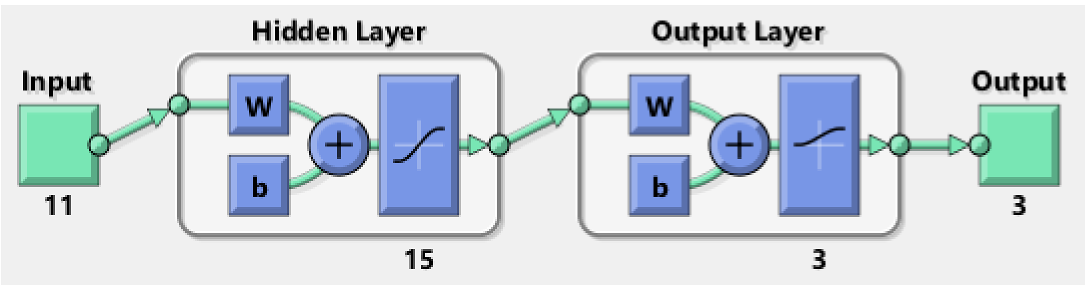


Figure 8. 11-15-3 Neural Network Structure

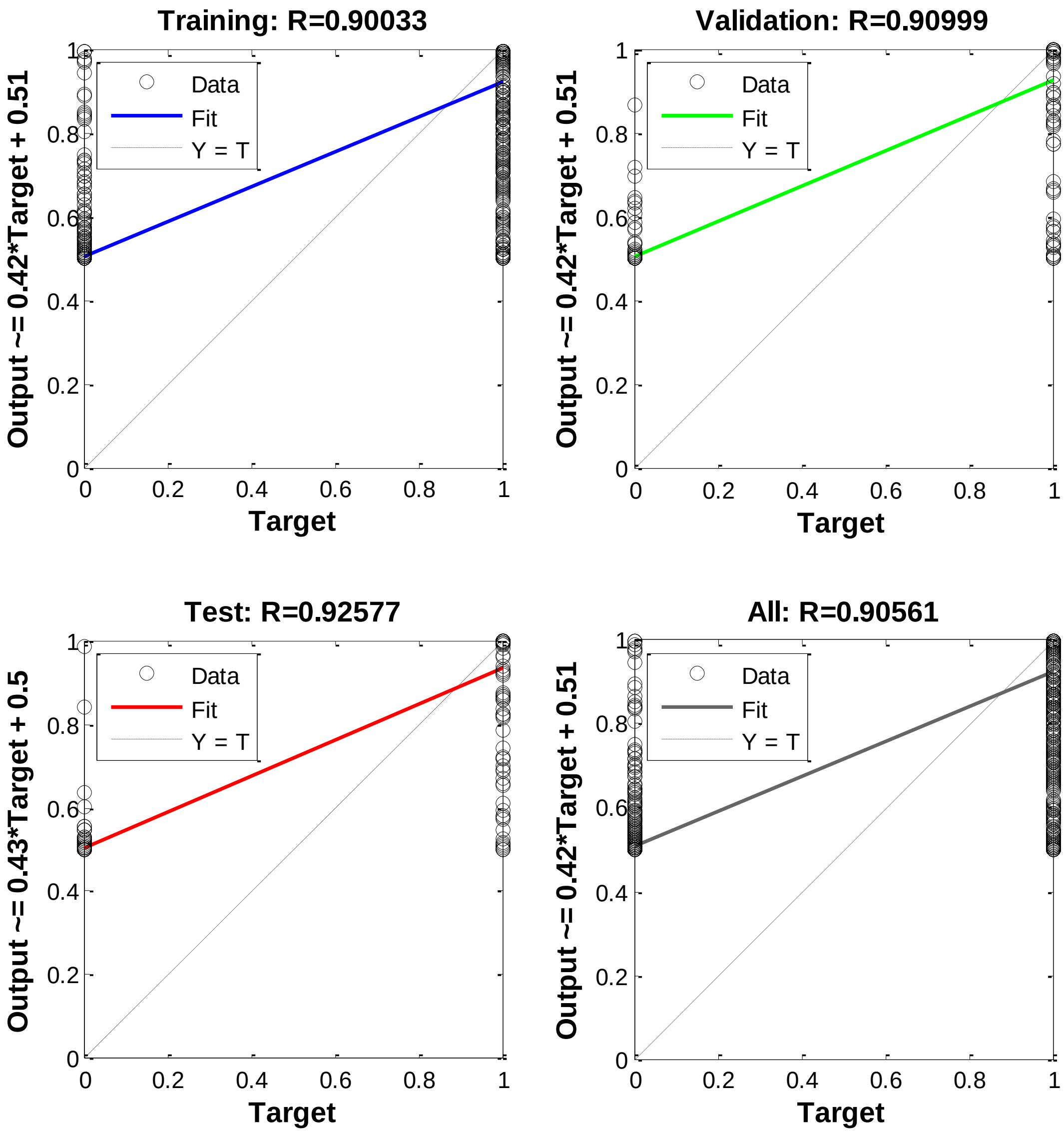


Figure 9. R Plots of the Best 11-15-3 Neural Network Developed

The variables used in the final neural network model include passenger travel information (number of people in travel group, number of children in travel group, frequency of travel, time of day, day of week), travel experience information (experience of using options A, B, or C in the past), and passenger personal information (age, gender, income). This does not follow the common idea that travel time and travel cost are the most significant factors contributing to a travel mode choice. As previously mentioned, the statistically good models are those which primarily included travel experience information. This can be interpreted as the neural network's effort to model the behavior and not necessarily the choice decision process.

7. SUMMARY

Out of all of the variables found to be significant, only AGE and INCOME were found to be significant in both logit and neural network models. All other variables in the neural network were not found to be significant in the logit models, just as those other variables found to be significant in the logit models were insignificant in the neural network models. This shows that the models developed captured different facets of the same discrete choice situation. The logit models can be understood to be more focused on modeling the decision making process of the passenger, while the neural network concentrated on modeling the overall historical behavior.

Table 6 shows a comparison of the R^2 and prediction accuracy of the best models developed. R^2 values for the logit models were computed using interpolation, as shown in Figure 10. As shown in the table, the best neural network has a higher R^2 value compared with the best ML and NL models. Also, the 11-15-3 NN has the highest prediction accuracy at almost 93%. This shows that the neural network model is a better fit in describing the travel choice behavior of the transport network studied as compared with either multinomial or nested logit model.

Table 6. Performance of Logit and Neural Network Models

Measure	ML3	NL3	11-15-3 NN
Pseudo R2	0.40277 (From Table 4)	0.40589 (From Table 5	-
R	-	-	0.90561 (From Figure 10)
R2	0.80332	0.80707	0.82013
Prediction Accuracy [%]	70.33493	70.33493	92.98246

Table 7 shows the disaggregated prediction accuracy of the top three models, where a 65.95% prediction accuracy means that 65.95% of those who chose option A were predicted to choose option A. As shown, the neural network model is also capable of accurately predicting the minority (Choice C), having a prediction accuracy of 100%, as compared with the 26.51% of both ML3 and NL3. Logit models have long been unable to accurately model the minority, as these are usually treated as errors by the model. In the neural network, on the other hand, the minority is the one having the perfect prediction rate. This shows that the neural network takes every observation as a true and perfectly valid observation, and thus, tries to model it along with all other observations. The prediction accuracy, computed to be at 102.12% for Choice B, can be explained as the model predicting more individuals choosing option B than the actual number, corresponding to some prediction errors.

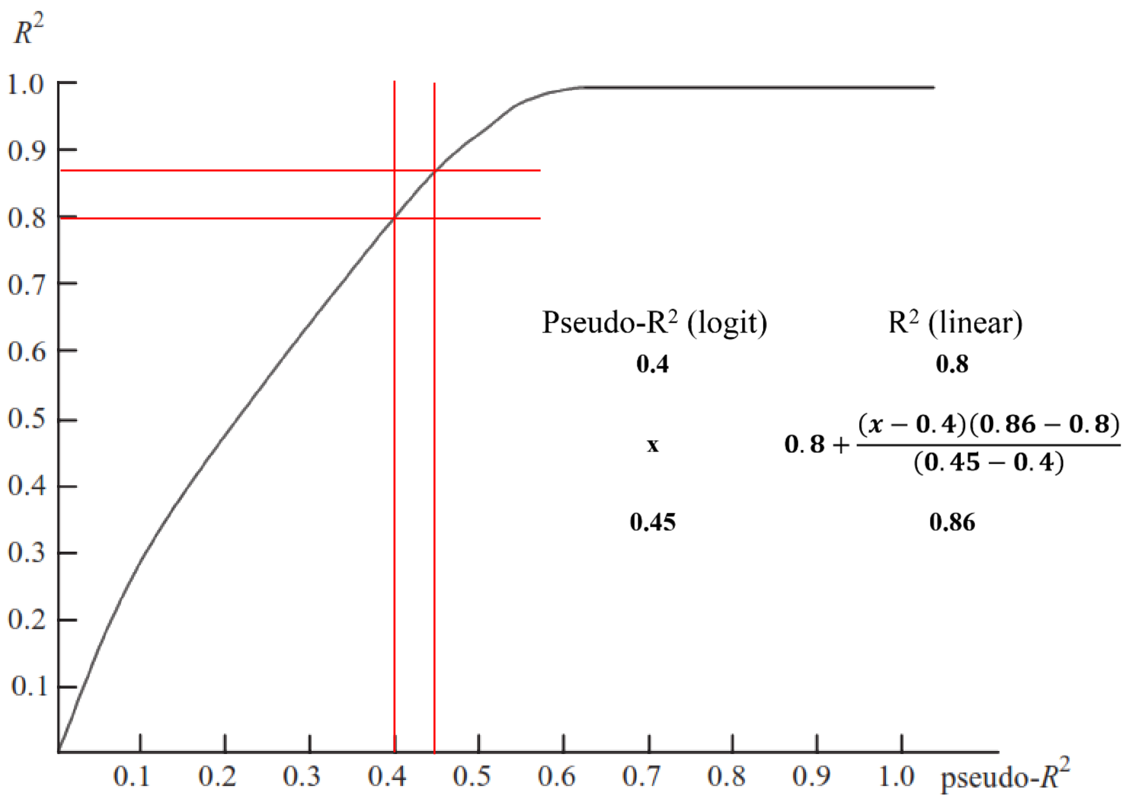


Figure 10. Relationship of Logit Pseudo-R2 and Linear R2

Table 7. Disaggregate Prediction Accuracy [%]

		Model		
		ML3	NL3	11-15-3 NN
Choice	A	65.94724	65.94724	96.16307
	B	77.58621	77.58621	102.122
	C	26.50602	26.50602	100.00

8. CONCLUSIONS & RECOMMENDATIONS

These findings do not, however, mean that neural networks are always better than logit models. If anything, this paper only shows that neural networks can also be used in modeling intra-regional travel, aside from urban trips that have been the focus of most other researches. Also, while the neural network can statistically better model the travel choice being studied, logit models explicitly show the numerical contributions of the variables that ultimately add up to a decision. This allows for the computation of external quantities like the value of time of the population, which can be used in many other applications, unlike the black-box characteristic of neural networks that does not provide any insight on the structure of the function being approximated.

This paper also recognizes the applicability of using data sets in determining the best combinations of input data. As the total number of input variables amount to 74, there would be much difficulty in accounting for all possible combinations. Thus, looking at the small improvements of R^2 values as more input variables and hidden neurons are added, the researcher found it best to keep the neural network as simple and uncrowded as possible. Furthermore, as the research was performed with the aim of finding a more efficient approach in developing discrete choice models, grinding through strenuous modeling using all possible combinations of variables, while finding the optimum number of hidden neurons at the same time, would not have been the way to go.

As for the computation of relative importance of variables, in testing its significance in the discrete choice model, the researcher recommends conducting connection weight analysis on the neural network. As the previous work already has discussions on marginal effects and elasticities for the logit models developed, determining the relative importance of the variables found to be significant in the neural network can be used to further evaluate the applicability of neural networks in predicting travel choices. Being able to get the same findings would only strengthen the idea of the applicability of neural networks in discrete choice analysis.

ACKNOWLEDGMENT

The researchers would like to thank Dr. Andres Winston C. Oreta and Dr. Lessandro Estelito O. Garciano for sharing their knowledge and guidance in the ANN component of this study.

REFERENCES

Al-Zoubi, H., Hilal, N., Darwish, N., & Mohammad, A. (2007). Rejection and modeling of sulphate and potassium salts by nanofiltration membranes: neural network and Spiegler-Kedem model, Desalination 206, 42–60

Bentz, Y., & Merunka, D. (2000). Neural networks and the multinomial logit for brand choice modeling: A hybrid approach. Retrieved from: cs.uni-muenster.de/Professoren/Lippe/diplomarbeiten/html/eisenbach/UntersuchteArtikel/BeMe00.pdf

Cantarella, G., & Luca, S. (2002). Multilayer feedforward networks for transportation mode choice analysis: An analysis and a comparison with random utility models. Retrieved from: www.sciencedirect.com/science/article/ii/S0968090X05000148

Chen, H. & Kim, A. (2006). Prediction of permeate flux decline in crossflow membrane filtration of colloidal suspension: a radial basis function neural network approach, Desalination 192, 415–428

Dia, H. (2010). Evaluation of discrete choice and neural network approaches for modeling driver compliance with traffic information. Retrieved from: www.researcgate.net/publlication/51028304_Evaluation_of_discrete_choice_and_neural_network_approaches_for_modeling_driver_compliance_with_traffic_information

Gershenson, C. (2003). Artificial neural networks for beginners. Retrieved from: https://arxiv.org/ftp/cs/papers/0308/0308031.pdf

Hensher, D., & Ton, T. (2000). A comparison of the predictive potential of artificial neural networks and nested logit models for commuter mode choice. Retrieved from: www.sciencedirect.com/science/article/pii/S1366554599000307

Hess, S. (2005). Advanced discrete choice models with applications to transport demand. Retrieved from: http://www.cts.cv.ic.ac.uk/documents/theses/HessPhD.pdf

Nakayama, S., Takayama, J., & Yamashita, Y. (2008). A non-linear analysis of discrete choice behavior by the logit model. Retrieved from: www.trip.t.u-tokyo-ac.jp/eastsjwp/2008/0804/0804.pdf

Nijkamp, P., Reggiani, A., & Tritapepe, T. (1996). Modeling inter-urban transport flows in Italy: A comparison between neural network analysis and logit analysis. Retrieved from: www.sciencedirect.com/science/article/pii/S0968090X96000174

Norets, A. (2008). Estimation of dynamic discrete choice models using artificial neural network approximations. Retrieved from: www.econ.brown.edu/fac/Andriy_Norets/papers/anorets07_ANNMCMC.pdf

Pulugurta, S., Arun, A., & Errampalli, M. (2013). Use of artificial intelligence for mode choice analysis and comparison with traditional multinomial logit model. Retrieved from: www.sciencedirect.com/science/article/pii/S1877042813045436

Rajaonarison, D., Bolduc, D. & Jayet, H. (2005). The k-deformed multinomial logit model. Economics Letters, Vol. 86 No.1, 13-20.

Razavi, M., Mortazavi, A., & Mousavi, M. (2004). Application of neural networks for crossflow milk ultrafiltration simulation, Int. Dairy J. 14, 69–80

Roquel, K., & Fillone, A. (2013). Mode choice analysis of inter-island passenger travel from Iloilo to Negros Occidental, Philippines. Retrieved from: easts.info/on-line/proceedings/vol9/PDF/P125.pdf

Vythoulkas, P., & Koutsopoulos, H. (2003). Modeling discrete choice behavior using concepts from fuzzy set theory, approximate reasoning and neural networks. Retrieved from: www.sciencedirect.com/science/article/pii/S0968090X02000219

Walczak, S. & Cerpa, N. (2003). Artificial neural networks, in: R.A. Meyers (Ed.), Encyclopedia of Physical Science and Technology, third ed., Academic Press, San Diego (CA), pp. 631–645

APPENDIX

Category	Variable	Description
Trip Purpose	purwork	1 – trip purpose is work; 0 – if not
	purvaca	1 – trip purpose is vacation; 0 – if not
	purschl	1 – trip purpose is school; 0 – if not
	purbusi	1 – trip purpose is business; 0 – if not
	purhome	1 – trip purpose is home; 0 – if not
Passenger Travel Information	num_grp	Number of people in travel group
	chl_grp	Number of children in travel group
	freqncy	Frequency of travel
	beflunch	1 – time of travel is before 12:00 P.M.; 0 – if not
	wkday	1 – day of travel is a weekday; 0 – if not
Travel Experience	usedrta	1 – have experience using route A; 0 – if none
	usedrtb	1 – have experience using route B; 0 – if none
	usedrtc	1 – have experience using route C; 0 – if none
Travel Choice Information	a_time	Travel time when using option A
	b_time	Travel time when using option B
	c_time	Travel cost when using option C
	a_tcost	Travel cost when using option A
	b_tcost	Travel cost when using option B
	c_tcost	Travel time when using option C
	a_wttme	Waiting time when using option A
	b_wttme	Waiting time when using option B
	c_wttme	Waiting time when using option C
Access Information	a_comorpr	Comfort of accessing option A
	b_comorpr	Comfort of accessing option B
	c_comorpr	Comfort of accessing option C
	a_torpr	Time of accessing option A
	b_torpr	Time of accessing option B
	c_torpr	Time of accessing option C
	a_corpr	Cost of accessing option A
	b_corpr	Cost of accessing option B
	c_corpr	Cost of accessing option C
Egress Information	a_comprde	Comfort of egressing option A
	b_comprde	Comfort of egressing option B
	c_comprde	Comfort of egressing option C
	a_tprde	Time of egressing option A
	b_tprde	Time of egressing option B
	c_tprde	Time of egressing option C
	a_cprde	Cost of egressing option A
	b_cprde	Cost of egressing option B
	c_cprde	Cost of egressing option C
Others	a_cbag	Additional cost for baggage when using option A
	b_cbag	Additional cost for baggage when using option B
	c_cbag	Additional cost for baggage when using option C
	b_rdtrp	1 – option B user bought roundtrip tickets; 0 – if not
Passenger Personal Information	age	Age of passenger
	gender	1 – passenger is male; 0 – if female
	income	Personal monthly income of passenger
Other Passenger Personal Information	single	1 – passenger is single; 0 – if not
	married	1 – passenger is married; 0 – if not
	num_chl	Number of children of passenger
Other Passenger Financial Information	num_mot	Number of motorcycles owned by passenger
	num_car	Number of cars owned by passenger
	num_van	Number of vans owned by passenger

	num_suv num_jpn vacatn	Number of SUVs owned by passenger Number of jeepneys owned by passenger Number of vacations passenger takes yearly
General Travel Choice Information	a_totcom b_totcom c_totcom a_lndtime b_lndtime c_lndtime a_seatime b_seatime c_seatime a_freqncy b_freqncy c_freqncy a_tottime b_tottime c_tottime a_aircon b_aircon c_aircon	Total comfort when using option A Total comfort when using option B Total comfort when using option C Total time travelling on land when using option A Total time travelling on land when using option B Total time travelling on land when using option C Total time travelling at sea when using option A Total time travelling at sea when using option B Total time travelling at sea when using option C Operation frequency of option A Operation frequency of option B Operation frequency of option C Total time when using option A Total time when using option B Total time when using option C Time spent in air-conditioned facility when using option A Time spent in air-conditioned facility when using option B Time spent in air-conditioned facility when using option C