Exercise 5 (10 Points) Solution

Exercise Description:

Load the Exercise5.xml file and the organics.sas7bdat dataset to SAS Enterprise Miner for questions 1-12. Question 14 needs to be completed in SAs Viya. The organics dataset is already available for you in the Viya system.
Make sure the model roles and measurement levels of the variables are as shown below.

Name	Model Role	Measurement Level	Description
ID	ID	Nominal	Customer loyalty identification number
DemAffl	Input	Interval	Affluence grade on a scale from 1 to 30
DemAge	Input	Interval	Age, in years
DemCluster	Rejected	Nominal	Type of residential neighborhood
DemClusterGroup	Input	Nominal	Neighborhood group
DemGender	Input	Nominal	M = male, F = female, U = unknown
DemRegion	Input	Nominal	Geographic region
DemTVReg	Input	Nominal	Television region
PromClass	Input	Nominal	Loyalty status: tin, silver, gold, or platinum
PromSpend	Input	Interval	Total amount spent
PromTime	Input	Interval	Time as loyalty card member
TargetAmt	Rejected	Interval	Number of organic products purchased
TargetBuy	Target	Binary	Organics purchased? 1 = Yes, 0 = No

A. Neural Networks Optimized on Average Square Error

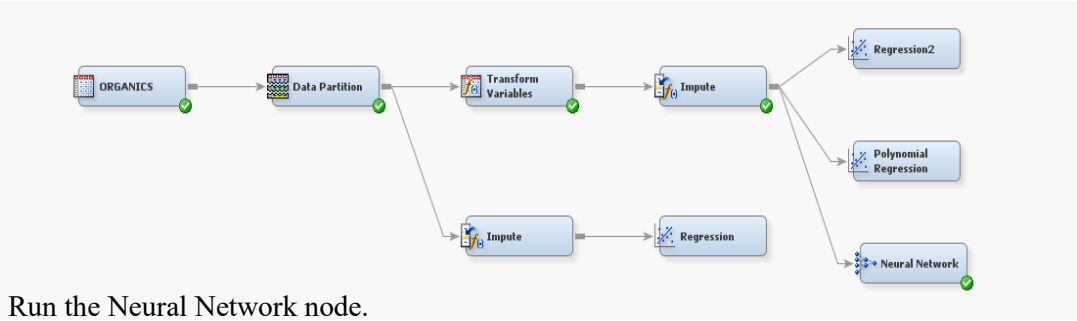
1. Is data transformation and imputation generally needed for a neural network model? Why or why not? (0.5 Point)

Solution: Yes, imputation would be necessary however, transformation may not be. Neural network models, as well as most models relying on a prediction formula, require a complete record for both modeling and scoring. Neural network models create transformations of inputs for use in a regression-like model. However, having input distributions with low skewness and kurtosis values tend to result in more stable models.

NNA1:

- ✓ Use the Organics XML diagram for this exercise (Provided for your reference).
- ✓ Add a Neural Network tool to the Organics XML diagram.
- ✓ Set the model selection criterion to average error.
- ✓ Connect the Impute node to the Neural Network node.

✓ Your diagram should look as follows.



✓ Run the Neural Network node.

2. Examine the results of the models and compare/ contrast the ASE and misclassification rates for all the models in the diagram. (0.5 point)

Solution: The Validation ASE and the validation Misclassification rate for the neural net model is the smallest so far.

Model	Validation ASE	Validation Misclassification	Number of Parameters
Untransformed Regression	0.136696	0.188419	8
Log Transformed Regression	0.137887	0.186469	8
Polynomial Regression	0.133731	0.188569	13
Neural Network	0.132638	0.184068	121

The number of parameters is 121. This is a very large number compared to regression models built earlier.

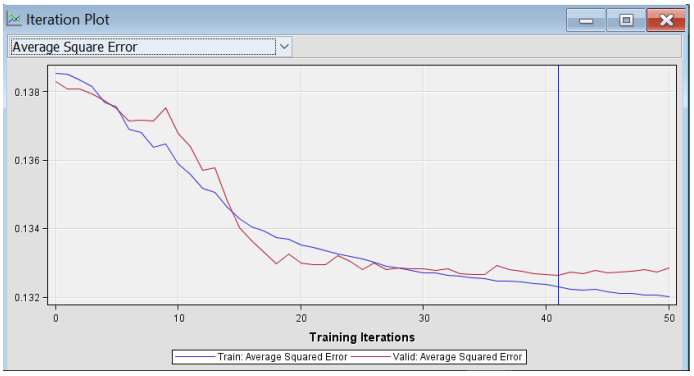
```
Dual Quasi-Newton Optimization

Dual Broyden - Fletcher - Goldfarb - Shanno Update (DBFGS)

Parameter Estimates          121
```

3. **Discuss the optimization process for this model, whether it converged, and what effect convergence has on your selected model. (0.5 point)**

   Solution: The model did not converge but this is not an issue since the optimal number of iterations was below 50.

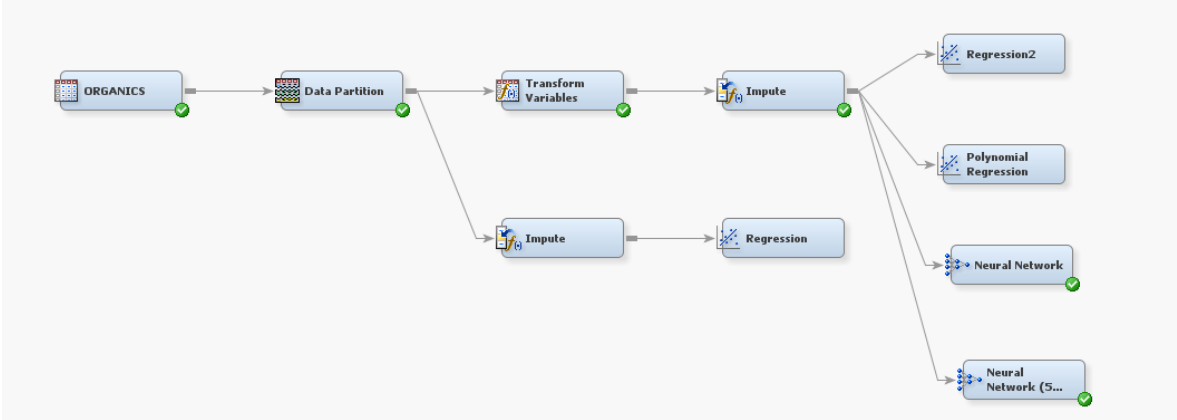



```
QUANEW needs more than 50 iterations or 2147483647 function calls.

WARNING: QUANEW Optimization cannot be completed.
```

NNA2:

- ✓ Copy and paste the Neural network node.
- ✓ Connect the newly pasted node to the Impute node as shown below.
- ✓ Change the hidden units in the Neural Network (rename as Neural Net (5 hidden nodes) to 5.
- ✓ Your diagram should look as follows.



- ✓ Run this Neural Network node.

4. **Compare and contrast the number of parameters across all the models and discuss why examining the number of parameters is important. (1 Point)**

   Solution: The number of parameters is 201. This is a very large number compared to regression models and neural network built earlier.

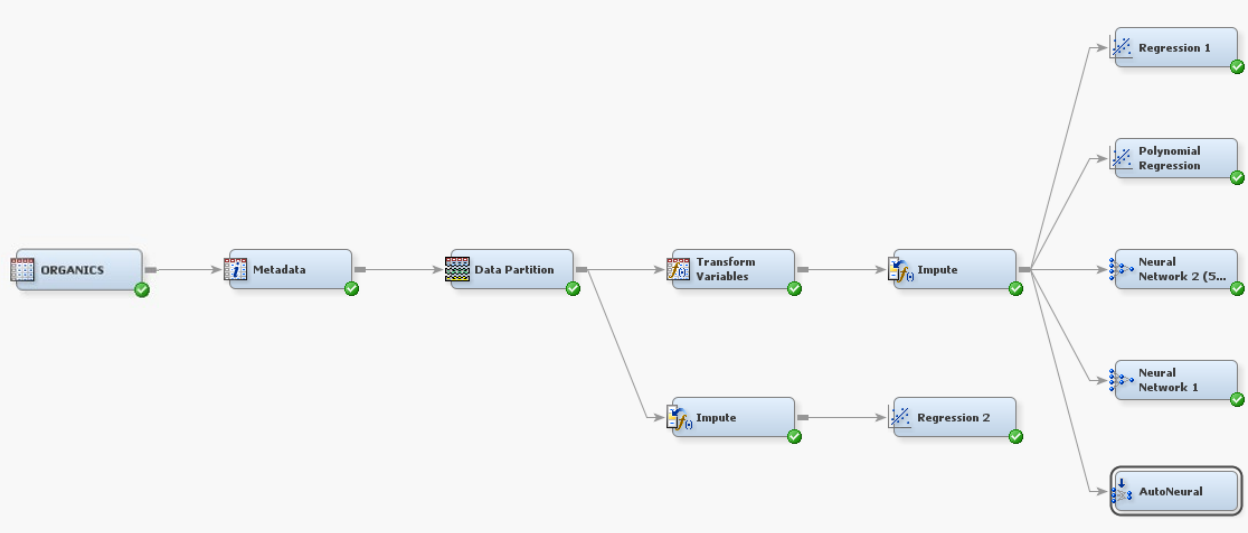
5. **What is the validation average squared error and the validation misclassification rate? Does it change from the previous model (Neural Net Model)? (0.5 Point)**

Solution: The misclassification rate further declined as did the ASE.

Model	Validation ASE	Validation Misclassification	Number of Parameters
Untransformed Regression	0.136696	0.188419	8
Log Transformed Regression	0.137887	0.186469	8
Polynomial Regression	0.133731	0.188569	13
Neural Network	0.132638	0.184068	121
Neural Network 5 nodes	0.131866	0.183618	201

NNA3:

- ✓ Add an Auto Neural Network node.
- ✓ Connect the Impute node to the Auto Neural node as in step c and e.
- ✓ Select Train Action Search.
- ✓ Select Number of Hidden Units 2.
- ✓ Select Tolerance High.
- ✓ In Activation function, select Direct No, Normal No and Sine No.
- ✓ Run the Auto Neural node.



6. **How many neurons are in the autoneural node? How many weights are estimated in this model? (0.5 Point)**

Solution: The auto neural network model has 2 neurons:

```
Final Model
Stopping: Termination criteria were satisfied: overfitting based on _VMISC_

_func_     _AVERR_     _VAVERR_     neurons

 TANH      0.41875      0.41830        2
                                    =======
                                        2
```

There are 81 weights estimated in the model:

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	T
TargetBuy	Organics Pu...	DFT	Total Degre...	15557	.	
TargetBuy	Organics Pu...	DFE	Degrees of ...	15476	.	
TargetBuy	Organics Pu...	DFM	Model Degr...	81	.	
TargetBuy	Organics Pu...	NW	Number of E...	81	.	
TargetBuy	Organics Pu...	AIC	Akaike's Info...	13189.35	.	
TargetBuy	Organics Pu...	SBC	Schwarz's B...	13809.19	.	
TargetBuy	Organics Pu...	ASE	Average Sq...	0.133016	0.132756	
TargetBuy	Organics Pu...	MAX	Maximum Ab...	0.983678	0.978236	
TargetBuy	Organics Pu...	DIV	Divisor for A...	31114	13332	
TargetBuy	Organics Pu...	NOBS	Sum of Freq...	15557	6666	
TargetBuy	Organics Pu...	RASE	Root Averag...	0.364713	0.364358	
TargetBuy	Organics Pu...	SSE	Sum of Squ...	4138.648	1769.909	
TargetBuy	Organics Pu...	SUMW	Sum of Cas...	31114	13332	
TargetBuy	Organics Pu...	FPE	Final Predict...	0.134408	.	
TargetBuy	Organics Pu...	MSE	Mean Squar...	0.133712	0.132756	
TargetBuy	Organics Pu...	RFPE	Root Final P...	0.366617	.	
TargetBuy	Organics Pu...	RMSE	Root Mean ...	0.365666	0.364358	
TargetBuy	Organics Pu...	AVERR	Average Err...	0.418698	0.418184	
TargetBuy	Organics Pu...	ERR	Error Function	13027.35	5575.225	
TargetBuy	Organics Pu...	MISC	Misclassifica...	0.184611	0.184818	
TargetBuy	Organics Pu...	WRONG	Number of ...	2872	1232	

7. **Examine the results of the models and compare/ contrast the ASE and misclassification rates for all the models in the diagram.** **(1 Point)**

   Solution: Average square error and misclassification rate for validation dataset has increased compared to the other Neural Net models.

Model	Validation ASE	Validation Misclassification	Number of Parameters
Untransformed Regression	0.136696	0.188419	8
Log Transformed Regression	0.137887	0.186469	8
Polynomial Regression	0.133731	0.188569	13
Neural Network	0.132638	0.184068	121

Neural Network 5 nodes	0.131866	0.183618	201
Auto Neural	0.132756	0.184818	81

8. Which model would you select as the best model? (0.5 Point)

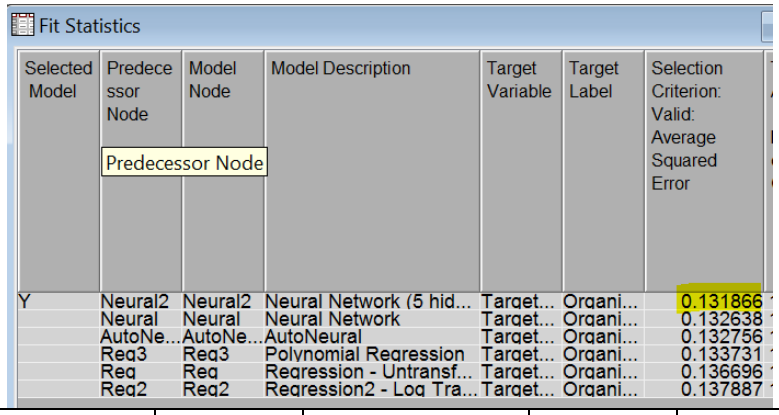
Solution: This auto neural node does not improve upon previous neural net models. It would not be the selected model of choice. At this point, I would select the Neural Network with 5 hidden units as the best model.

NNA4:

- ✓ Add a Model Comparison node
- ✓ Use Validation ASE
- ✓ Connect it with all the above models.
- ✓ Run the model.
- ✓ Create a table in which you are comparing all your models built based on validation ASE, AIC, ROC Index and Sensitivity.

9. Discuss the outcome of the Model Comparison node optimized on Validation ASE compared to the results of your table in relation to the other measures requested. (1 Point)

Solution: The second neural network model (NN- 5 hidden units) is selected by the model comparison node because it has the lowest validation average square error. It also gives the best misclassification rate but this model is by far the most complex model. Sensitivity indicates the other two neural networks do a slightly better job with classifying the data. AIC One thing is definite, the neural networks outperform the regression analysis.



Model	Validation ASE	Validation Misclassification	AIC	Validation ROC Index	Validation Sensitivity	Number of Parameters
Untransformed Regression	0.136696	0.188419	13553.18	0.809	0.3901	8
Log Transformed Regression	0.137887	0.186469	13644.59	0.806	0.4010	8
Polynomial Regression	0.133731	0.188569	13267.52	0.822	0.4046	13

Neural Network	0.132638	0.184068	13193.58	0.825	0.4397	121
Neural Network 5 nodes	0.131866	0.183618	13333.53	0.825	0.4313	201
Auto Neural	0.132756	0.184818	13189.35	0.828	0.4343	81

B. Neural Network Optimized on Misclassification Rate

NNB1:

- ✓ Create a new diagram
- ✓ Copy all the nodes in your current diagram and paste it in the new diagram.
- ✓ In the new diagram, use Misclassification rate to optimize all of the above models.
- ✓ Once you have ran all the models, bring in a model comparison node and connect it with the models built using misclassification rate for optimization.
- ✓ Use validation misclassification rate as criterion in Model Comparison node.

10. Discuss the model that was selected by this model comparison node, how it compared to the previous model comparison and why you obtained these results. (1 Point)

Solution: The first neural network model was selected by the model comparison node. We can see from the previous table that the misclassification rate was the second smallest of all models. When optimized on misclassification, this model was able to surpass the 2nd neural network model.

Fit Statistics

Selected Model	Predecessor Node	Model Node	Model Description	Target Variable	Target Label	Selection Criterion: Valid: Misclassification Rate
Y	Neural	Neural	Neural ...	Target...	Organi...	0.182568
	Neural2	Neural2	Neural ...	Target...	Organi...	0.183318
	AutoNe...	AutoNe...	Neural ...	Target...	Organi...	0.184818
	Reg2	Reg2	Regres...	Target...	Organi...	0.186469
	Reg3	Reg3	Regres...	Target...	Organi...	0.187219
	Reg	Reg	Regres...	Target...	Organi...	0.188419

C. Neural Networks with Differing Imputation Methods

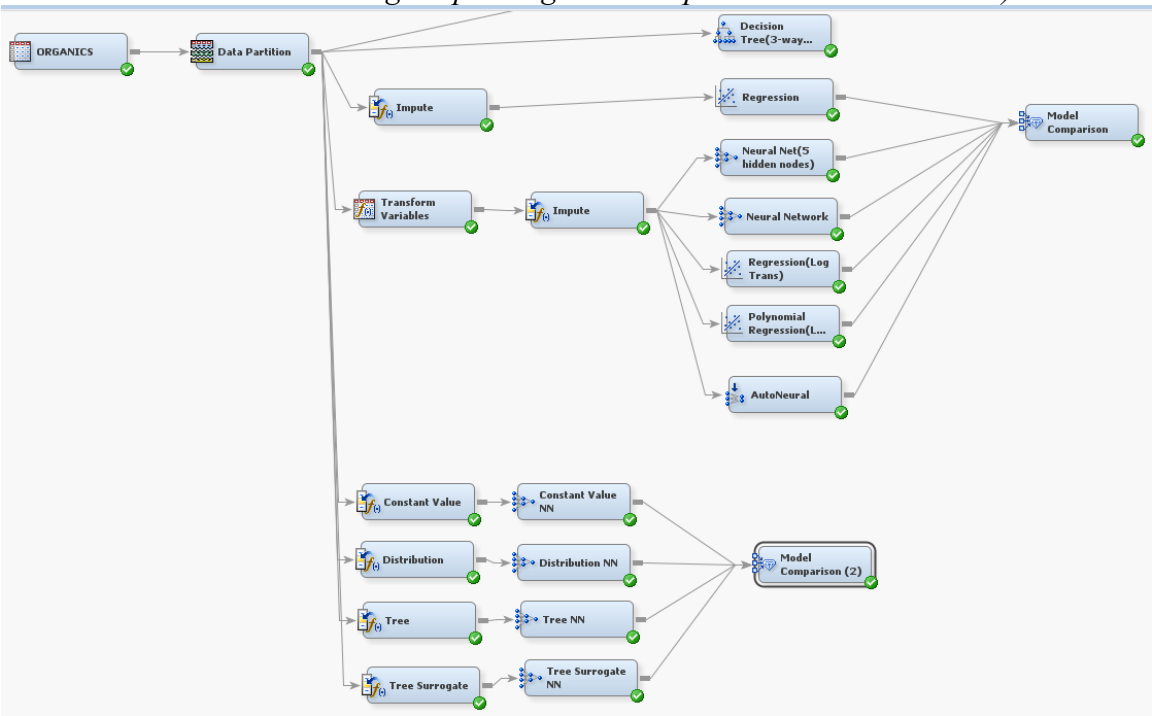
NNC1:

- ✓ In your original diagram, connect 4 different impute nodes with different Default Input Method for interval variables (See the property panel of impute node) after the data partition node.
- ✓ Connect a Neural Network node to each of the impute nodes.
- ✓ Use the setting as listed above in NNA1 instructions to build the neural network models.

✓ Compare the four different neural network models.

11. Which model is performing the best (use Validation ASE). Provide your comments. Attach appropriate screenshots of your work. (0.5 Point)

Solution: According to the model comparison node, the Neural Network built from the distribution imputation had the best validation ASE.
Note: *This answer could change depending on the imputation methods chosen)*



Fit Statistics

Selected Model	Predecessor Node	Model Node	Model Description	Target Variable	Target Label	Selection Criterion: Valid: Average Squared Error
Y	Neural4	Neural4	Distribution NN	Target...	Organi...	0.133692
	Neural5	Neural5	Tree NN	Target...	Organi...	0.13921
	Neural6	Neural6	Tree Surrogate NN	Target...	Organi...	0.139498
	Neural3	Neural3	Constant Value NN	Target...	Organi...	0.143461

NNC2:

✓ From the best performing impute node (connected to best performing neural net model in question 12) connect a decision tree, a PLS node and a LARS node, for variable selection.

✓ Use the default settings.

- ✓ Connect a neural network model with each of the decision tree, PLS and the LARS nodes.

12. Which neural network model is working best (Use Validation ASE). Is the best performing neural network performing better than the best model in question Why or why not? Attach appropriate screenshots. (0.5 Point)

Solution: The best model from this set is the neural network that had distribution imputation and a decision tree for variable selection with a validation ASE of 0.13151. This ASE did improve over the traditional neural network model. The misclassification and ROC index also improved over the first neural network model. The sensitivity is 45.0%, which is higher than the original neural networks.

Fit Statistics

Selected Model	Predecessor Node	Model Node	Model Description	Target Variable	Target Label	Selection Criterion: Valid: Average Squared Error
Y	Neural7	Neural7	Constant Value DT NN	Target...	Organi...	0.13151
	Neural8	Neural8	Constant Value LARS NN	Target...	Organi...	0.132482
	Neural9	Neural9	Constant Value PLS NN	Target...	Organi...	0.133692

13. Compare and contrast the results from Viya and Enterprise Miner. What worked and what didn't? Were you able to make the same option changes in both environments? Did your results turn out the same? Which did you like more? What else did you notice? (2 points)

Solution: Answers will vary