# Lab Artificial Neural Networks

Please submit your answers in this word document, including detailed discussion for each question with supportive screenshots of your analysis in software.

1. Please provide a screen shot of your entire SAS EM window with opened lab handout walkthrough (Decision Trees and Neural Networks Diagram) similar to the following. This should include your first and last name in the project name.

A screenshot of a computer

Description automatically generated

1. Using the same DMAGECR data (good\_bad as target) and data partition, after variable transformation, build a Neural Network with 5 hidden units and Softmax activation function for target layer (modify the settings in the ‘Network’). Also, use the backpropagation training technique with learning rate of 0.01 (settings in Optimization). Leave all other settings as default.
   1. What is the misclassification rate on the training and validation data, respectively?

Training data misclassification rate = 0.24

Validation data misclassification rate = 0.291

A screenshot of a computer

Description automatically generated



* 1. What is the Sensitivity and specificity on the training and validation data?

A screenshot of a computer

Description automatically generated



From the above results

Training data metrics:

|  |  |
| --- | --- |
| Sensitivity | 0.8866 |
| Specificity | 0.4833 |

Sensitivity: TP/(TP+FN); Specificity: TN/(FP+TN)

Sensitivity: TP/(TP+FN); Specificity: TN/(FP+TN)

Validation data metrics:

|  |  |
| --- | --- |
| Sensitivity | 0.8619 |
| Specificity | 0.3483 |

* 1. Based on above a and b, do you think there is overfitting issue?

1. Sensitivity (true positive rate) is larger, and Specificity (true negative rate) is lower. This means, the model is predicting true for actual true values and lower specificity means model do not classify negative results more correctly.
2. Also, the validation data misclassification rate is greater than training data misclassification rate.

Hence there is a problem of overfitting in Neural Network 5.

1. Using the same DMAGECR data (good\_bad as target) and data partition, build a new AutoNeural model with Maximum Iterations of 15 and Number of Hidden Units of 3. Compare this model with the Neural Network created in Question 2. Which model is better in terms of the ROC on the validation data?

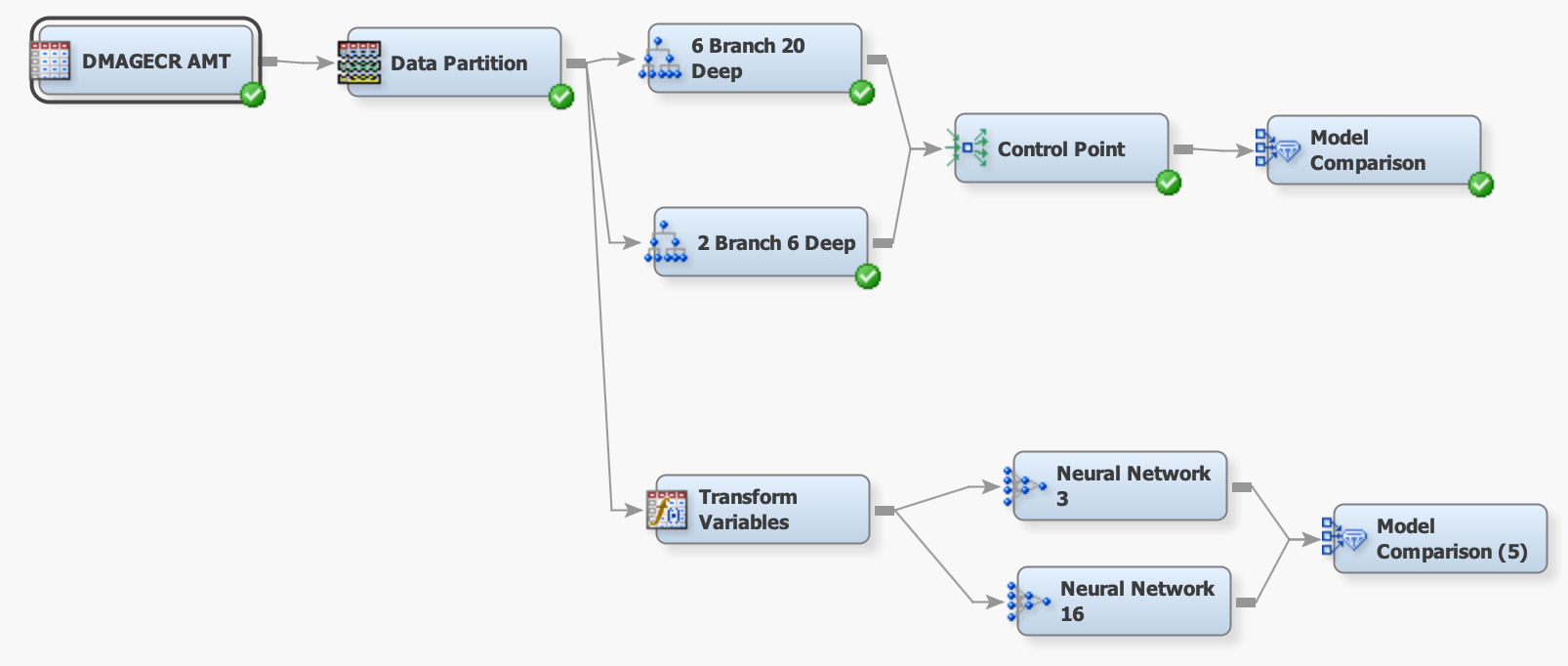
Using the DMAGECR with good\_bad target variable. Based on the results from Neural Network 5 and Auto Neural Network, by performing comparison on validation data ROC using model comparison node, we see that Neural Network 5 have greater ROC of 67.6%. However the margin of comparison is significantly low.

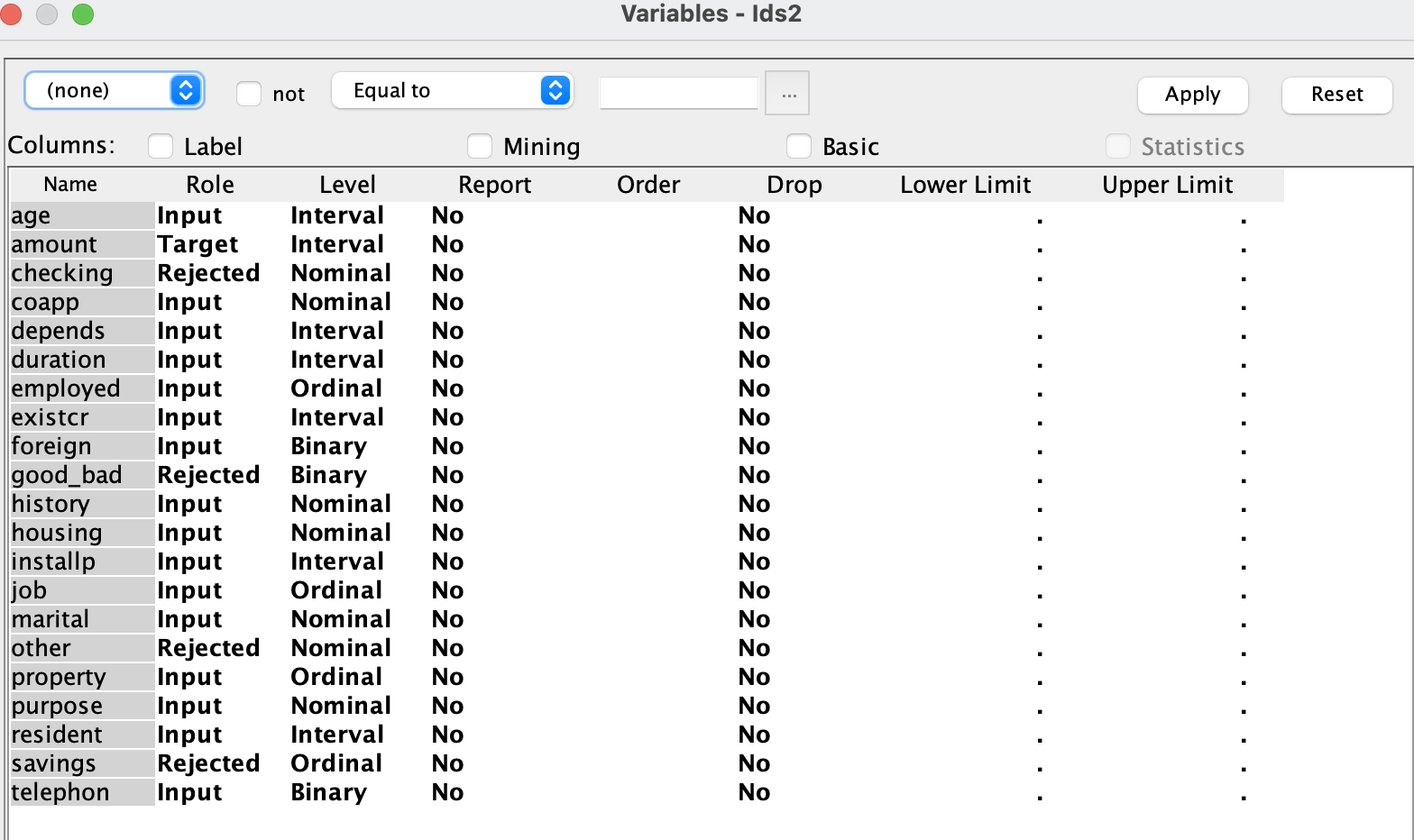
A screenshot of a computer

Description automatically generated



1. Copy the nodes of Transform Variables, Neural Network 3, Neural Network 16 from the classification portion to the prediction portion of Decision Trees and Neural Nets Walkthrough. Change the model selection criterion to Average Error for both neural network models. Add model comparison node to compare the two neural networks. Your diagram should be similar to the following and make sure the data DMAGECR AMT has the following variable settings. Which of the two neural networks is better based on average squared error on the validation data?





In the below screenshot, we see the results of comparing Neural Network 3 and Neural Network 16 on validation data Average Squared Error (ASE), it is clear that neural network 3 have least ASE. Which means the same model can produce more accurate results.

A screenshot of a computer

Description automatically generated

