

Assignment 5

1.) (5 pts.) Compete in the Kaggle *Titanic: Machine Learning from Disaster* competition using decision tree, random forest, or gradient boosting models.

<https://www.kaggle.com/c/titanic>

To receive full credit, you must score at least 0.79426 correct classification. You may receive one point extra credit for beating the stated benchmark.

Grading rubric:

3 pts. for any logical code or diagram

1 pt. for Kaggle Submission

1 pt. for meeting benchmark

1 pt. for beating benchmark

The screenshot displays the SAS Enterprise Miner interface for a project named "assignment_4". The main workspace shows a workflow diagram with the following steps: "TITANIC_TRAIN" (Data Source) → "Decision Tree" (Model) → "Score" (Assess) → "SAS Code" (Output). A second data source, "TITANIC_TEST", is also connected to the "Decision Tree" model. The left sidebar contains a tree view of the project structure, including "Data Sources", "Diagrams", "Log", and "Model Packages". The "Properties" pane on the left shows the configuration for the "Decision Tree" model, including various splitting rules and node settings. A "Variables - Tree" dialog box is open in the foreground, showing a table of variables and their roles in the model.

Name	Use	Report	Role	Level
Age	Default	No	Input	Interval
Cabin	Default	No	Input	Nominal
Embarked	Default	No	Input	Nominal
Fare	Default	No	Input	Interval
Name	No	No	Text	Nominal
Parch	Default	No	Input	Interval
PassengerId	No	No	ID	Interval
Pclass	Default	No	Input	Nominal
Sex	Default	No	Input	Nominal
SibSp	Default	No	Input	Interval
Survived	Yes	No	Target	Binary
Ticket	Default	No	Input	Nominal

Assignment 5

2.) (5 pts.) Compete in the Kaggle *Digit Recognizer* competition using neural network models.

<https://www.kaggle.com/c/digit-recognizer>

To receive full credit, you must score at least 0.95343 correct classification. You may receive one point extra credit for beating the stated benchmark.

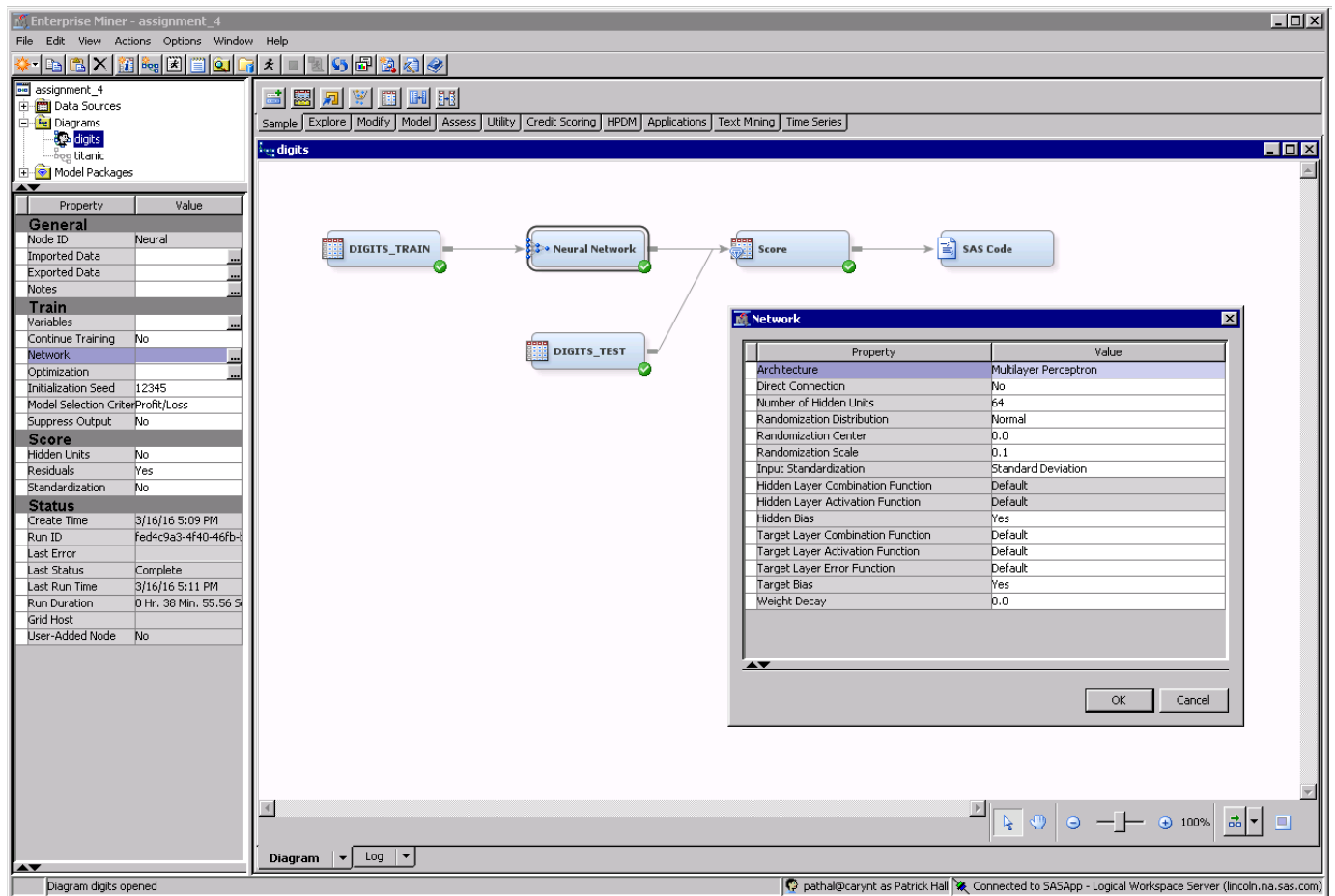
Grading rubric:

3 pts. for any logical code or diagram

1 pt. for Kaggle Submission

1 pt. for meeting benchmark

1 pt. for beating benchmark



Turn in only one document to black board.

If you choose to use SAS Enterprise Miner, this will be a Word document with the screenname you choose for each competition and a screenshot of your modeling diagram(s).

If you choose to use Python, R, or SAS code, turn in your code with your chosen screennames in the comments.

Assignment 5

For those of you using SAS Enterprise Miner, you can use the following code to create your submissions automatically after the Score node in a SAS Code node.

```
/* titanic */
%let submission_file = ;
data _null_;
    length line $25;
    file "&submission_file";
    if _n_ = 1 then do;
        line = 'PassengerId, Survived';
        put line;
    end;
    set &EM_IMPORT_SCORE;
    line = PassengerId||','||EM_CLASSIFICATION;
    put line;
run;

/* digits */
%let submission_file = ;
data _null_;
    length line $25;
    file "&submission_file";
    if _n_ = 1 then do;
        line = 'ImageId, Label';
        put line;
    end;
    set &EM_IMPORT_SCORE;
    line = put(_n_, best.)||','||EM_CLASSIFICATION;
    put line;
run;
```