

For this part, you will use JMP Pro to build and test five classifier models – Naïve Bayes, KNN, Partition (decision tree), Boosted Tree, and Neural Network. Follow the instruction given below. You will use *heart-disease-lecture.jmp* dataset, which is posted along with this assignment for all five models. The dataset was downloaded from UCI Machine Learning Repository and slightly modified for this assignment. For more details about how to use each classifier model, refer to *Predictive and Specialized Modeling.pdf*, which you can find from JMP Pro documentations.

(1). Naïve Bayes

- Start JMP Pro
- Open *heart-disease-lecture.jmp*
- Select Analyze > Predictive Modeling > Naïve Bayes
- Select *age* through *thal* and click *X,Factor*
- Select *class* and click *Y,Response*
- Select *Validation* (on the left) and click *Validation* (on the right)
- Click OK.

You will see Naïve Bayes Report (or output) that is similar to Figure 9.1 in *Predictive and Specialized Modeling.pdf* documentation.

- Capture the report screen and paste it in your submission.
- The dataset's class attribute has two possible values – 1 and 2.
- Using the confusion matrix of *Validation* in the report (There are two confusion matrices. Make sure that you use the *Validation* confusion matrix), calculate the following measures for both classes (similar to those in Weka's output window):

TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	Class
						1
						2

(2). KNN

- Start JMP Pro
- Open *heart-disease-lecture.jmp*
- Select Analyze > Predictive Modeling > K Nearest Neighbors
- Select *age* through *thal* and click *X,Factor*
- Select *class* and click *Y,Response*
- Select *Validation* (on the left) and click *Validation* (on the right)
- Click OK.

You will see K Nearest Neighbors Report (or output) that includes, among others, Model Selection, Misclassification Rate for both Training and Validation, and confusion matrices for best *K* value.

- Capture the report screen and paste it in your submission.
- What is the best K value (on the validation set)?
- The dataset's class attribute has two possible values – 1 and 2.
- Using the confusion matrix of *Validation* in the report (There are two confusion matrices. Make sure that you use the *Validation* confusion matrix), calculate the following measures for both classes (similar to those in Weka's output window):

TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	Class
						1
						2

(3). Partition Model (decision tree)

- Start JMP Pro
- Open *heart-disease-lecture.jmp*
- Select Analyze > Predictive Modeling > Partition
- Select *age* through *thal* and click *X,Factor*
- Select *class* and click *Y,Response*
- Select *Validation* (on the left) and click *Validation* (on the right)
- Click OK.
- In the output (this output is called platform report window), click *Go*. You will see a decision tree in the output.
- Click a red triangle next to *Partition for class* and select *Show Fit Details*. Confusion matrices will be added to the output.
- Capture the output screen, which includes a decision tree and confusion matrices, and paste it in your submission. If you cannot capture all of them in a single screenshot, you may use multiple screenshots to capture all.
- The dataset's class attribute has two possible values – 1 and 2.
- Using the confusion matrix of *Validation* in the report (There are two confusion matrices. Make sure that you use the *Validation* confusion matrix), calculate the following measures for both classes (similar to those in Weka's output window):

TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	Class
						1
						2

(4). Boosted Tree

- Start JMP Pro
- Open *heart-disease-lecture.jmp*
- Select Analyze > Predictive Modeling > Boosted Tree
- Select *age* through *thal* and click *X,Factor*
- Select *class* and click *Y,Response*
- Select *Validation* (on the left) and click *Validation* (on the right)

- Click OK.

Gradient-Boosted Tree Specification window appears.

- In the Reproducibility panel, select *Suppress Multithreading* and enter 123 for Random Seed.
- Click OK.
- Capture the output screen, and paste it in your submission.
- The dataset's class attribute has two possible values – 1 and 2.
- Using the confusion matrix of *Validation* in the report (There are two confusion matrices. Make sure that you use the *Validation* confusion matrix), calculate the following measures for both classes (similar to those in Weka's output window):

TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	Class
						1
						2

(5). Neural Network

- Start JMP Pro
- Open *heart-disease-lecture.jmp*
- Select Analyze > Predictive Modeling > Neural
- Select *age* through *thal* and click *X,Factor*
- Select *class* and click *Y,Response*
- Select *Validation* (on the left) and click *Validation* (on the right)
- Click OK.
- In the *Hidden Layer Structure* box, enter 5 in the entry corresponding to *First* and *TanH* as shown below:

Hidden Layer Structure

Number of nodes of each activation type

Activation Sigmoid Identity Radial

Layer	TanH	Linear	Gaussian
First	5	0	0
Second	0	0	0

Second layer is closer to X's in two layer models.

- Click GO.
- In the next screen, click red triangle next to *Model NTanH(5)* and select *Diagram*. You will see confusion matrices and a neural network diagram in the output window.
- Capture the output screen, and paste it in your submission.
- How many hidden layers does the model have?

- The dataset's class attribute has two possible values – 1 and 2.
- Using the confusion matrix of *Validation* in the report (There are two confusion matrices. Make sure that you use the *Validation* confusion matrix), calculate the following measures for both classes (similar to those in Weka's output window):

TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	Class
						1
						2