

## Experiment-3.1

**Student Name: Milan Sharma**

**Branch: CSE AIML**

**Semester: 01**

**Subject Name: Artificial Intelligence Lab**

**UID: 23MAI10003**

**Section/Group: 23MAI-1**

**Date of Performance: 16/10/2023**

**Subject Code: 23CSH-621**

### **Aim of the Experiment :**

Aim of the Experiment is to train and test the prediction dataset (from UCI ML repository) using the 10 popular ML models and build your own ensemble using Majority voting and Stacking methods.

### **Objective of the Experiment :**

Task to be done for this experiment is that we have to perform following tasks:

- a) Train and test the prediction dataset (from UCI ML repository) using 10 popular ML models.
- b) Build your own ensemble using Majority Voting and Stacking.

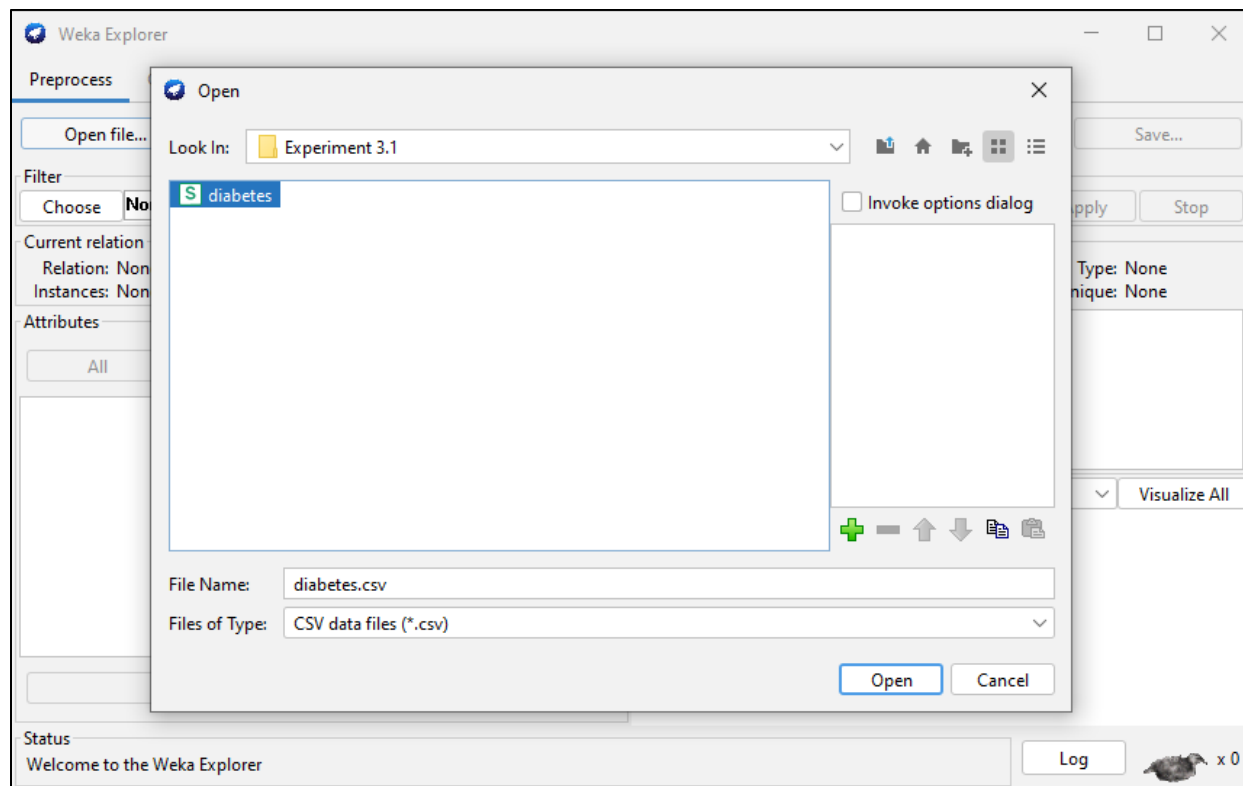
### **Algorithm/ Steps for Experiment :**

Step 1: Download the Diabetes dataset from UCI Machine Learning repository.

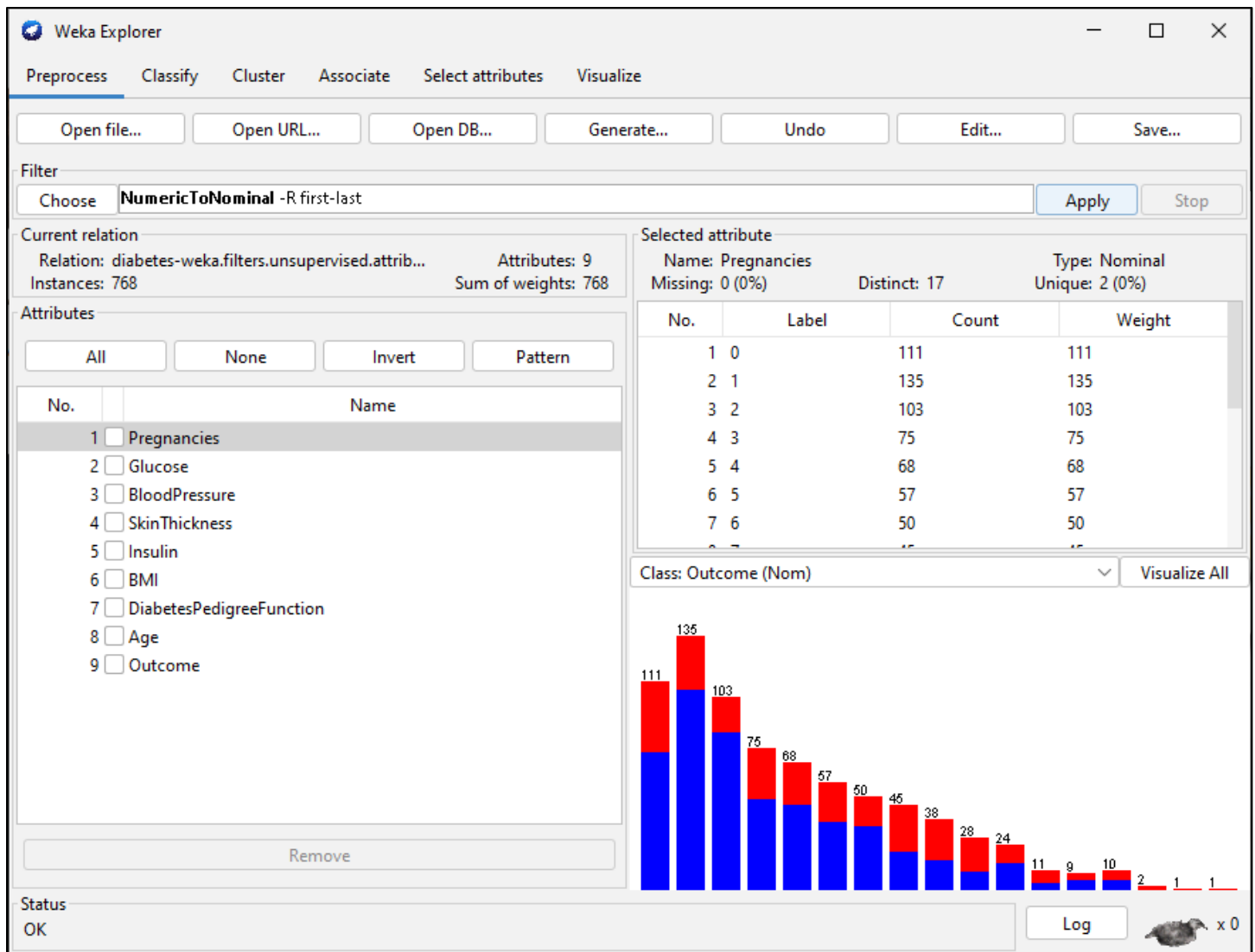
Step 2: Open the WEKA Tool and open the 'Explorer' tab.



Step 3: Click on the 'Open file' Option >> Select Diabetes dataset >> Click on Open.



Step 4: Change the dataset from Numeric to Nominal.



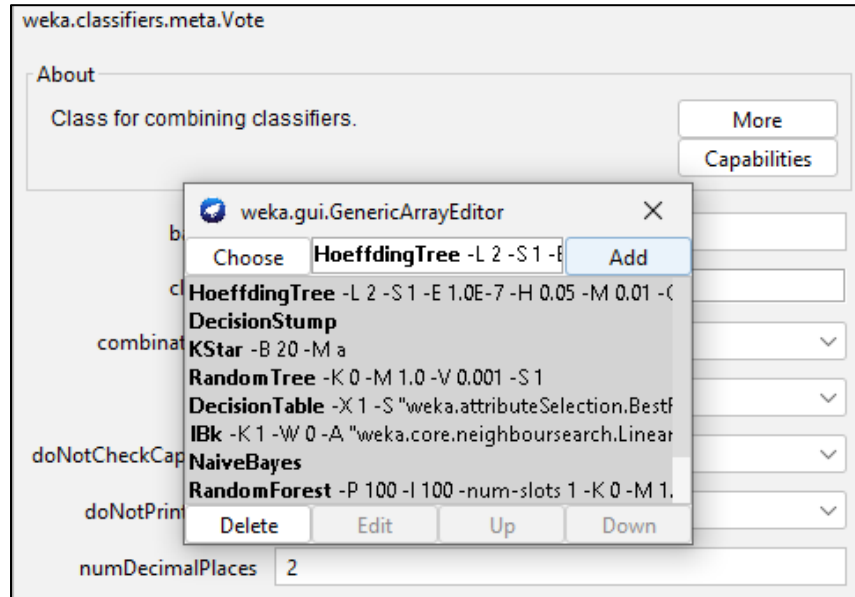
The screenshot shows the Weka Explorer interface with the 'Preprocess' tab selected. The 'Filter' dropdown is set to 'NumericToNominal -R first-last'. The 'Current relation' is 'diabetes-weka.filters.unsupervised.attrib...', with 9 attributes and 768 instances. The 'Attributes' list on the left includes 'Pregnancies', 'Glucose', 'BloodPressure', 'SkinThickness', 'Insulin', 'BMI', 'DiabetesPedigreeFunction', 'Age', and 'Outcome'. The 'Selected attribute' table on the right shows the distribution for 'Pregnancies' (Nominal, 17 distinct values, 2 unique values). Below this, a bar chart visualizes the distribution of the 'Outcome' attribute (Nominal), showing counts for each class (0 and 1).

No.	Label	Count	Weight
1	0	111	111
2	1	135	135
3	2	103	103
4	3	75	75
5	4	68	68
6	5	57	57
7	6	50	50
8	7	45	45
9	8	38	38
10	9	28	28
11	10	24	24
12	11	11	11
13	12	9	9
14	13	10	10
15	14	2	2
16	15	1	1
17	16	1	1

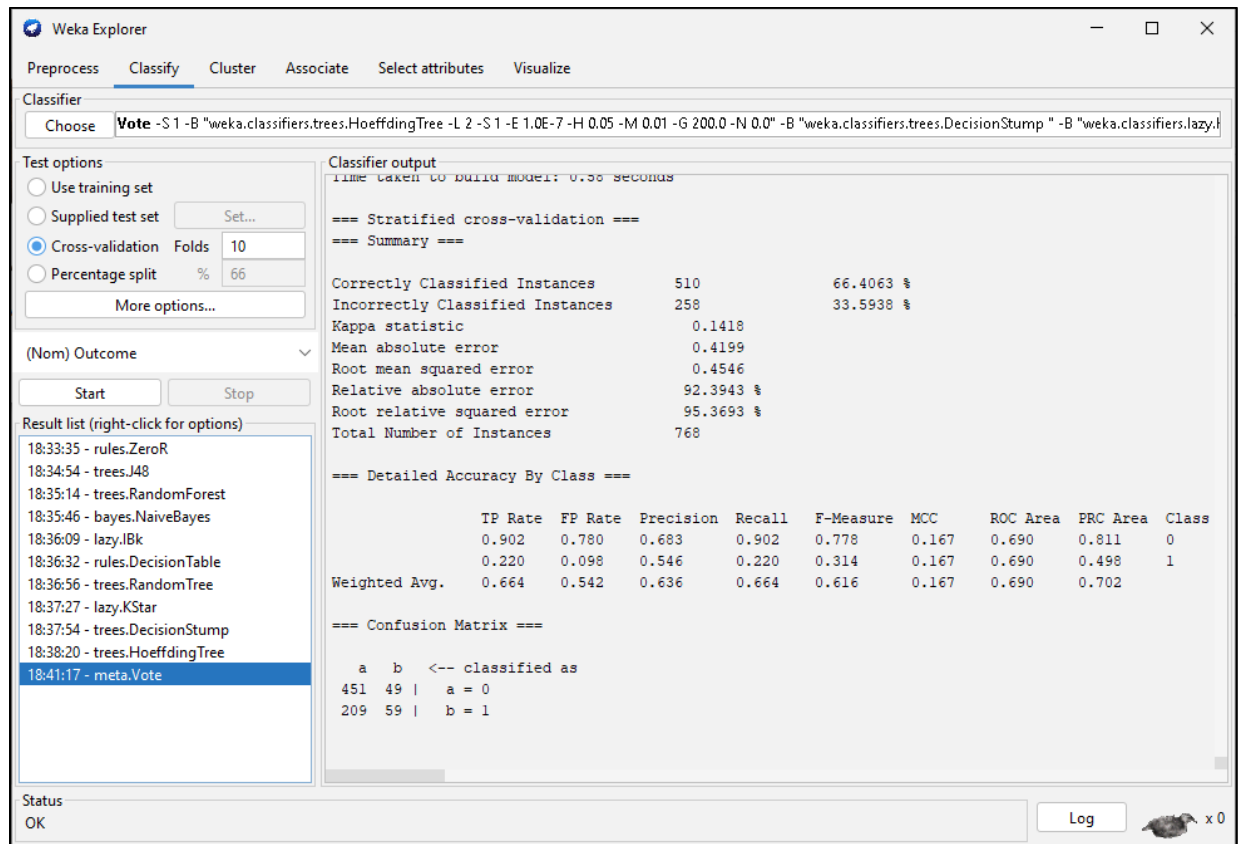
Class	Count	Weight
0	111	111
1	135	135

Step 5: Click on the 'Classify' Tab >> Choose and select different models from 'Classifier' Section.

Step 6: Select the “Vote” model for majority voting. Click on the model then click on Classifiers to select 10 different classifier models.



Step 7: Click Start to build the model. Classifier output shows the model evaluation parameters.



**Weka Explorer**

Preprocess **Classify** Cluster Associate Select attributes Visualize

Classifier: Choose **Vote** -S 1 -B "weka.classifiers.trees.HoeffdingTree -L 2 -S 1 -E 1.0E-7 -H 0.05 -M 0.01 -G 200.0 -N 0.0" -B "weka.classifiers.trees.DecisionStump" -B "weka.classifiers.lazy.IBk"

**Test options**

- ☐ Use training set
- ☐ Supplied test set
- ☒ Cross-validation Folds
- ☐ Percentage split %

(Nom) Outcome

**Result list (right-click for options)**

- 18:33:35 - rules.ZeroR
- 18:34:54 - trees.J48
- 18:35:14 - trees.RandomForest
- 18:35:46 - bayes.NaiveBayes
- 18:36:09 - lazy.IBk
- 18:36:32 - rules.DecisionTable
- 18:36:56 - trees.RandomTree
- 18:37:27 - lazy.KStar
- 18:37:54 - trees.DecisionStump
- 18:38:20 - trees.HoeffdingTree
- 18:41:17 - meta.Vote**

**Classifier output**

Time taken to build model: 0.36 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	510	66.4063 %
Incorrectly Classified Instances	258	33.5938 %
Kappa statistic	0.1418	
Mean absolute error	0.4199	
Root mean squared error	0.4546	
Relative absolute error	92.3943 %	
Root relative squared error	95.3693 %	
Total Number of Instances	768	

=== Detailed Accuracy By Class ===


	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.902	0.780	0.683	0.902	0.778	0.167	0.690	0.811	0
	0.220	0.098	0.546	0.220	0.314	0.167	0.690	0.498	1
Weighted Avg.	0.664	0.542	0.636	0.664	0.616	0.167	0.690	0.702	

=== Confusion Matrix ===

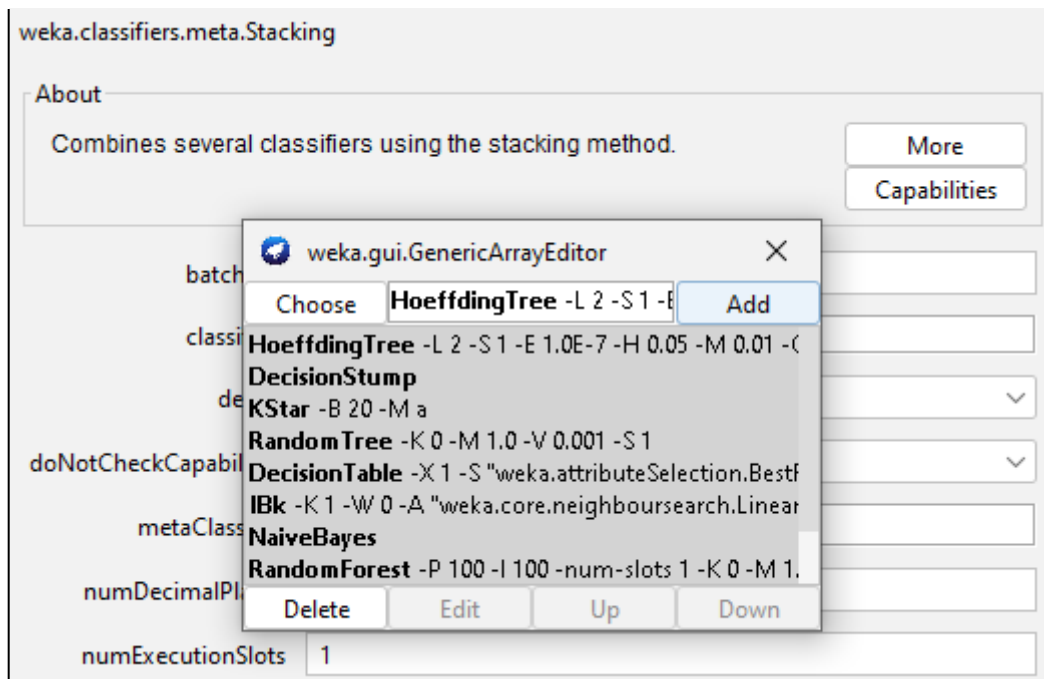
```

a  b  <-- classified as
451 49 | a = 0
209 59 | b = 1

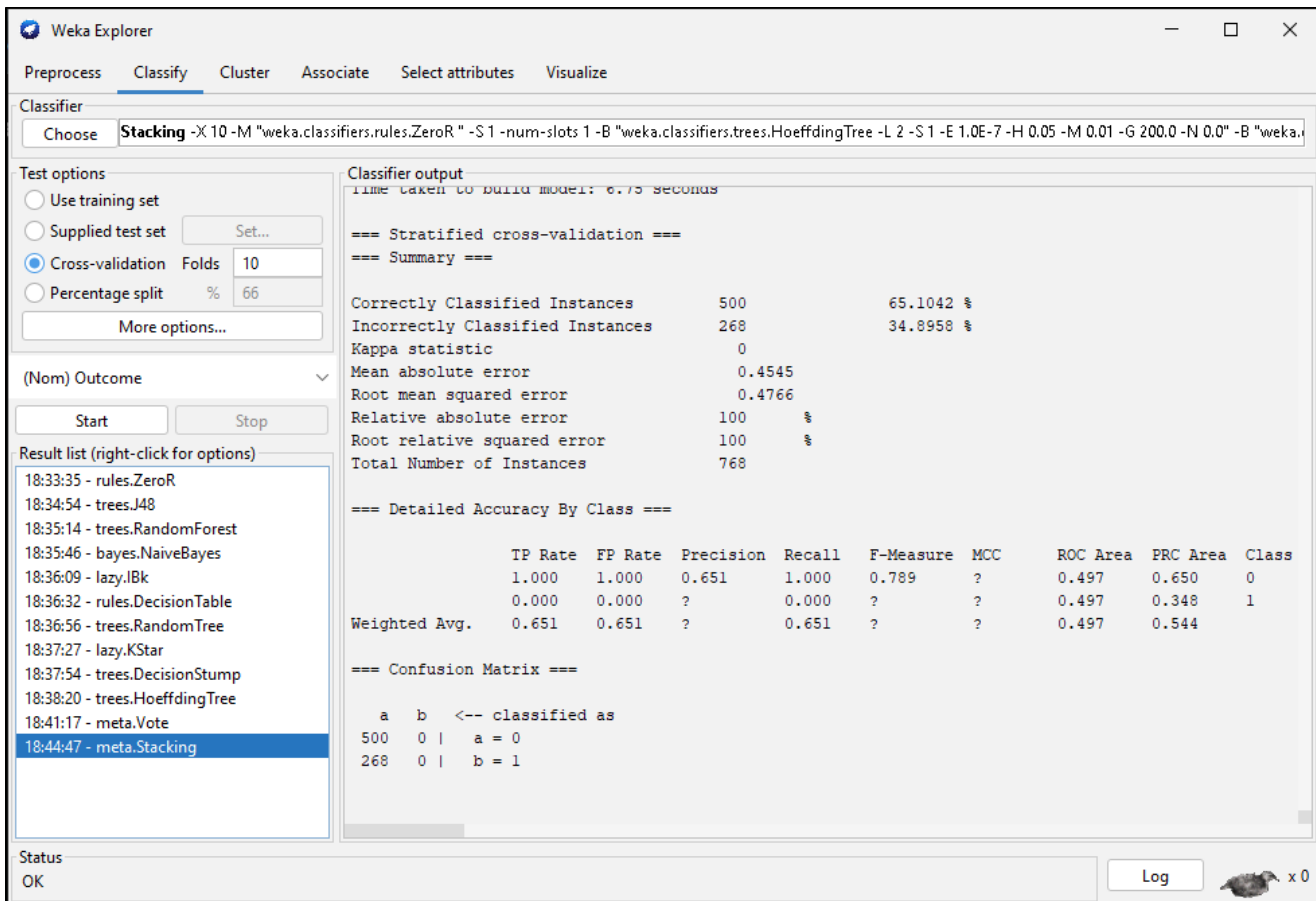
```

Status: OK   x 0

Step 8: Select the “Stacking” model for majority voting. Click on the model then click on Classifiers to select 10 different classifier models.



Step 9: Click Start to build the model. Classifier output shows the model evaluation parameters.



**Weka Explorer**

Preprocess **Classify** Cluster Associate Select attributes Visualize

Classifier: Choose **Stacking -X 10 -M "weka.classifiers.rules.ZeroR" -S 1 -num-slots 1 -B "weka.classifiers.trees.HoeffdingTree -L 2 -S 1 -E 1.0E-7 -H 0.05 -M 0.01 -G 200.0 -N 0.0" -B "weka.classifiers.trees.HoeffdingTree -L 2 -S 1 -E 1.0E-7 -H 0.05 -M 0.01 -G 200.0 -N 0.0"**

**Test options**

☐ Use training set  
☐ Supplied test set  
☒ Cross-validation Folds **10**  
☐ Percentage split % **66**

More options...

(Nom) Outcome **✓**

Start Stop

**Result list (right-click for options)**

- 18:33:35 - rules.ZeroR
- 18:34:54 - trees.J48
- 18:35:14 - trees.RandomForest
- 18:35:46 - bayes.NaiveBayes
- 18:36:09 - lazy.IBk
- 18:36:32 - rules.DecisionTable
- 18:36:56 - trees.RandomTree
- 18:37:27 - lazy.KStar
- 18:37:54 - trees.DecisionStump
- 18:38:20 - trees.HoeffdingTree
- 18:41:17 - meta.Vote
- 18:44:47 - meta.Stacking**

**Classifier output**

Time taken to build model: 6.75 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	500	65.1042 %
Incorrectly Classified Instances	268	34.8958 %
Kappa statistic	0	
Mean absolute error	0.4545	
Root mean squared error	0.4766	
Relative absolute error	100 %	
Root relative squared error	100 %	
Total Number of Instances	768	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
0	1.000	1.000	0.651	1.000	0.789	?	0.497	0.650	0
1	0.000	0.000	?	0.000	?	?	0.497	0.348	1
Weighted Avg.	0.651	0.651	?	0.651	?	?	0.497	0.544	

=== Confusion Matrix ===

a	b	-- classified as
500	0	a = 0
268	0	b = 1

Status: OK

Log x 0

## Learning outcomes (What I have learnt):

1. I learnt about the WEKA Tool and its applications.
2. I learnt about how to use Explorer Tab in WEKA Tool.
3. I learnt about how to change attributes from Numeric to Nominal.
4. I learnt about how to use the Vote and Stacking method in WEKA Tool.
5. I learnt about how to compare the accuracy of different models.