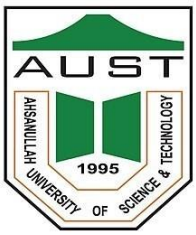


Ahsanullah University of Science & Technology

Department of Computer Science & Engineering

Course No : **CSE4142**

Course Title : **Data Warehousing and Mining Lab**

Assignment No : **03**

Date of Submission : **23.06.2024**

Submitted To : **Mr. Saha Reno & Mr. Raiyan Jahangir**

Submitted By-

Name: Nabila Rahman

ID: 20200204065

Section: B1

(i) Create a Custom Dataset Which Will Have 8 Attributes: 4 Numeric, 3 Nominal & 1 Class (3 Class Values)

=>

```
@relation vehicle

@attribute engine_size numeric
@attribute horsepower numeric
@attribute weight numeric
@attribute fuel_efficiency numeric
@attribute color {red,green,blue}
@attribute body_type {sedan,SUV,truck}
@attribute transmission_type {manual,automatic,CVT}
@attribute class {economy,standard,luxury}
```

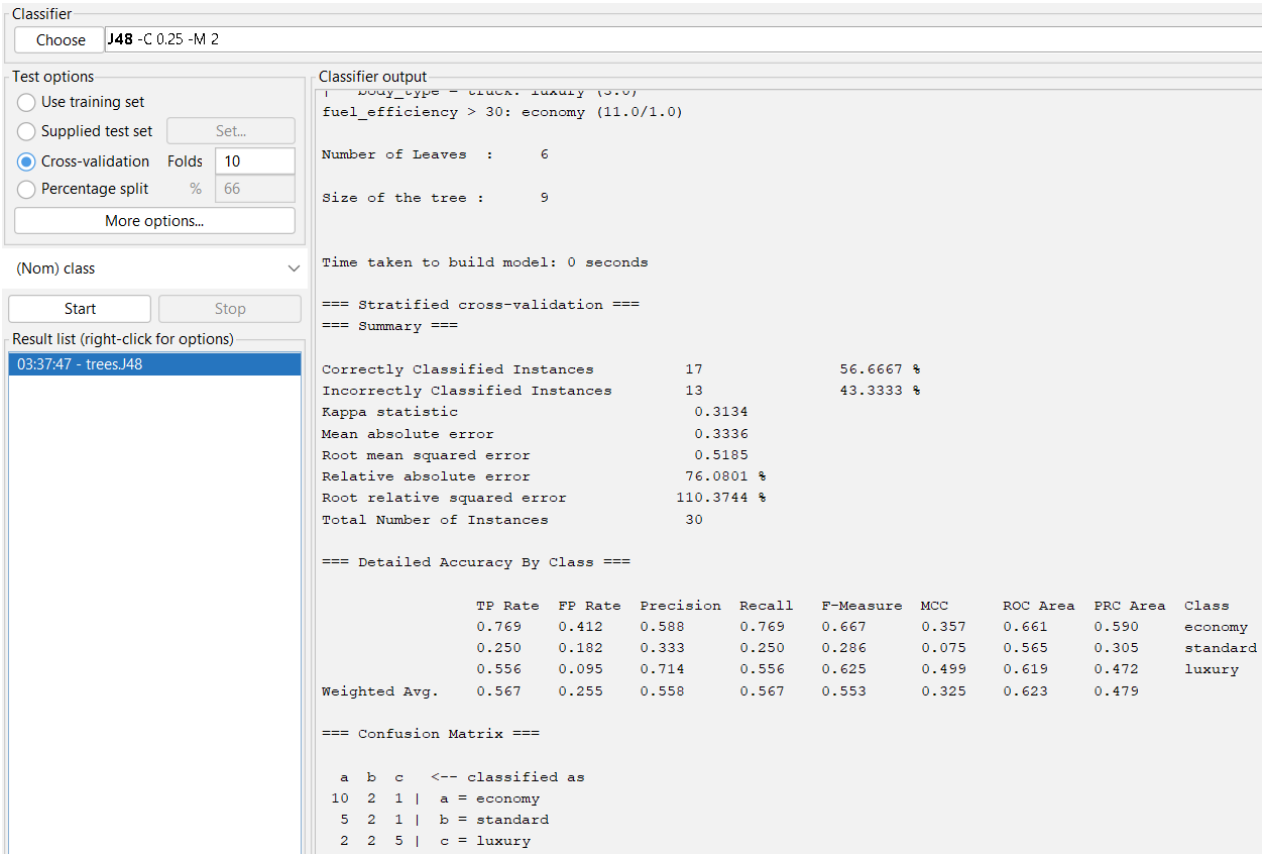
(ii) Create 30 Instances of That Dataset.

=>

```
@data
1.6, 120, 2800, 35.5, red, sedan, manual, economy
2.0, 150, 3000, 30.0, green, SUV, automatic, standard
3.5, 250, 4000, 25.0, blue, truck, CVT, luxury
1.8, 140, 2900, 32.5, red, sedan, manual, economy
2.2, 160, 3200, 28.0, green, SUV, automatic, standard
4.0, 300, 4200, 22.0, blue, truck, CVT, luxury
1.5, 110, 2700, 37.0, red, sedan, manual, economy
2.5, 180, 3300, 26.5, green, SUV, automatic, standard
3.8, 280, 4100, 23.5, blue, truck, CVT, luxury
1.7, 130, 2850, 34.0, red, sedan, manual, economy
2.1, 155, 3100, 29.5, green, SUV, automatic, standard
4.2, 320, 4300, 21.0, blue, truck, CVT, luxury
1.9, 145, 2950, 31.5, red, sedan, manual, economy
2.3, 170, 3250, 27.5, green, SUV, automatic, standard
3.6, 260, 4050, 24.5, blue, truck, CVT, luxury
1.4, 100, 2600, 38.5, red, sedan, manual, economy
2.4, 175, 3350, 26.0, green, SUV, automatic, standard
4.1, 310, 4250, 22.5, blue, truck, CVT, luxury
1.8, 135, 2750, 33.0, red, sedan, manual, economy
2.6, 190, 3400, 25.0, green, SUV, automatic, standard
3.7, 270, 4150, 23.0, blue, truck, CVT, luxury
1.7, 125, 2800, 36.0, red, sedan, manual, economy
2.2, 165, 3150, 28.5, green, SUV, automatic, standard
4.3, 330, 4350, 20.5, blue, truck, CVT, luxury
1.6, 120, 2900, 35.0, red, sedan, manual, economy
2.0, 150, 3100, 30.5, green, SUV, automatic, standard
3.9, 290, 4200, 24.0, blue, truck, CVT, luxury
1.5, 110, 2750, 37.5, red, sedan, manual, economy
2.5, 185, 3300, 27.0, green, SUV, automatic, standard
4.0, 300, 4400, 21.5, blue, truck, CVT, luxury
```

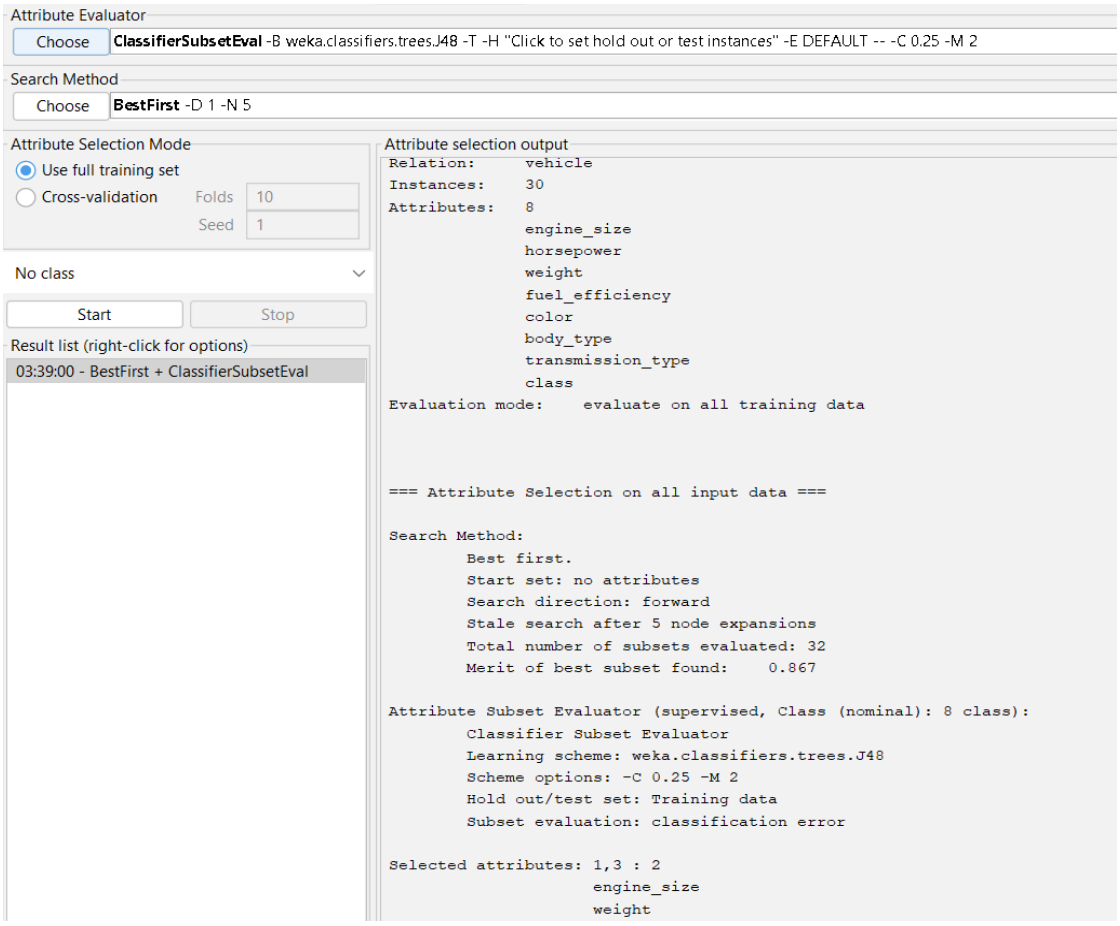
(iii) Construct a Classification Model using J48 Decision Tree Algorithm, Use 10-Fold Cross Validation.

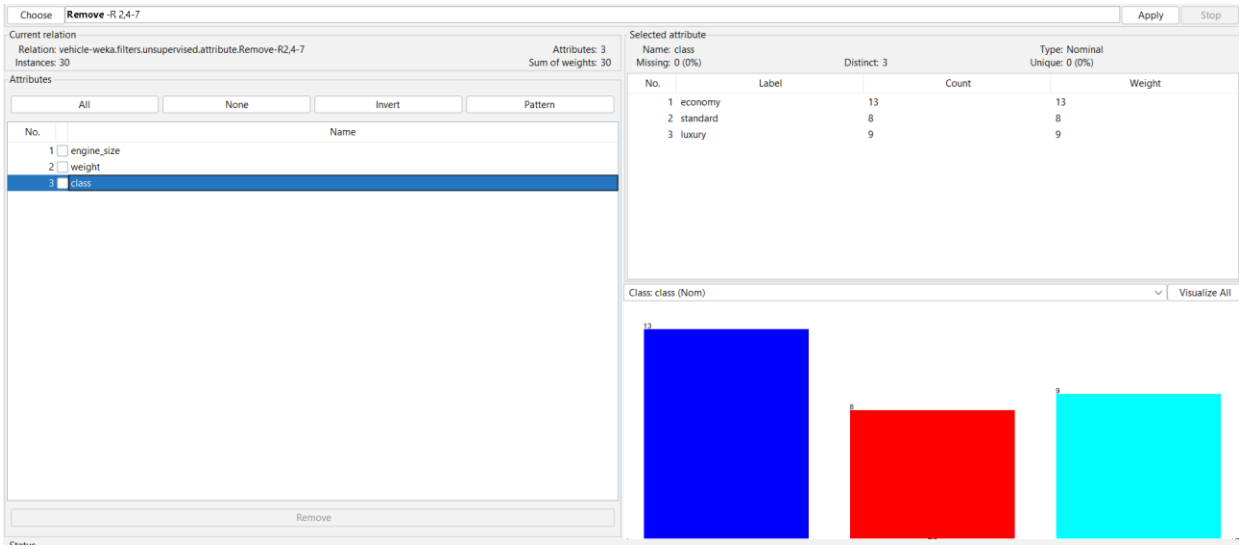
=>



(iv) Using Filter Method, Find Out Those Attributes for Which the J48 Model Performs the Best (Least Number of "Inaccurately Classified Instances"). All Other Attributes Should Be Removed.

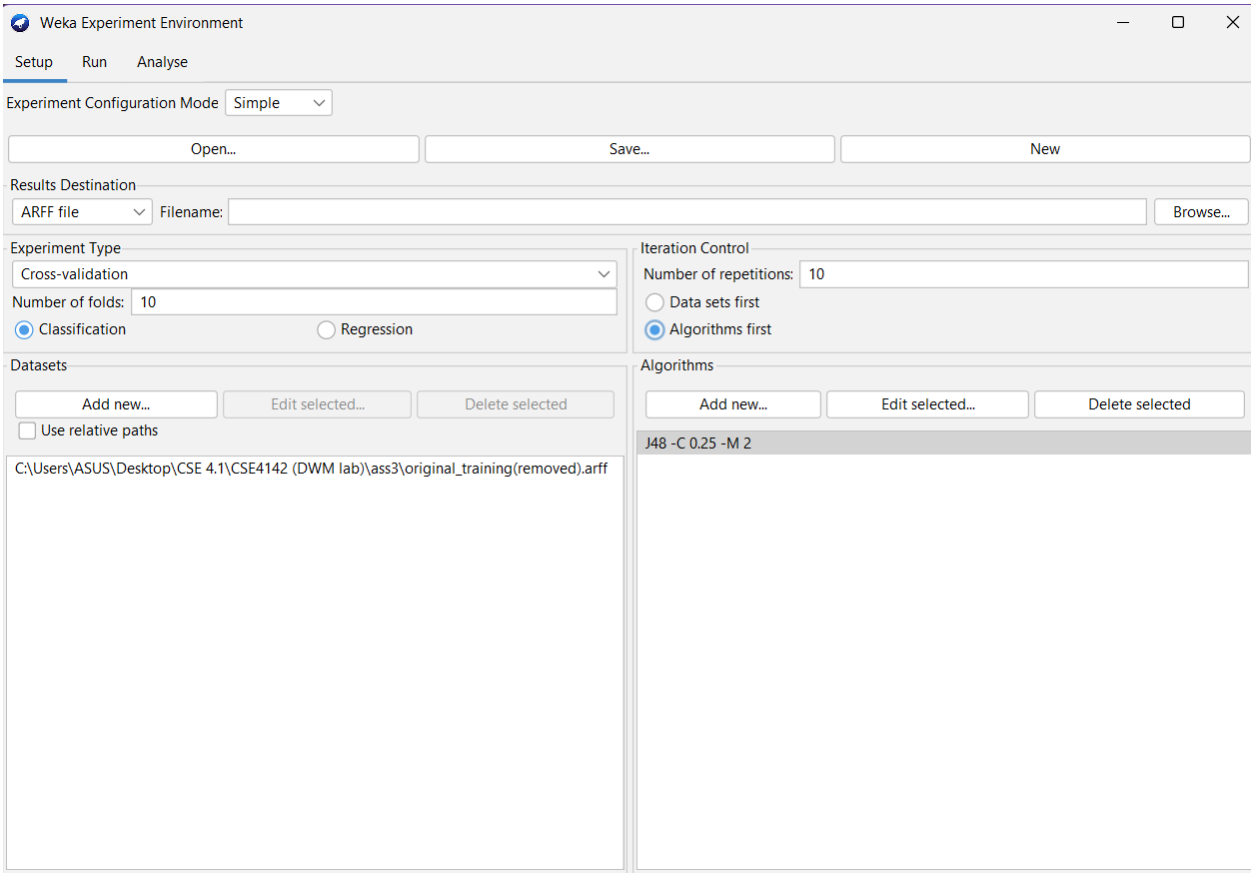
=>

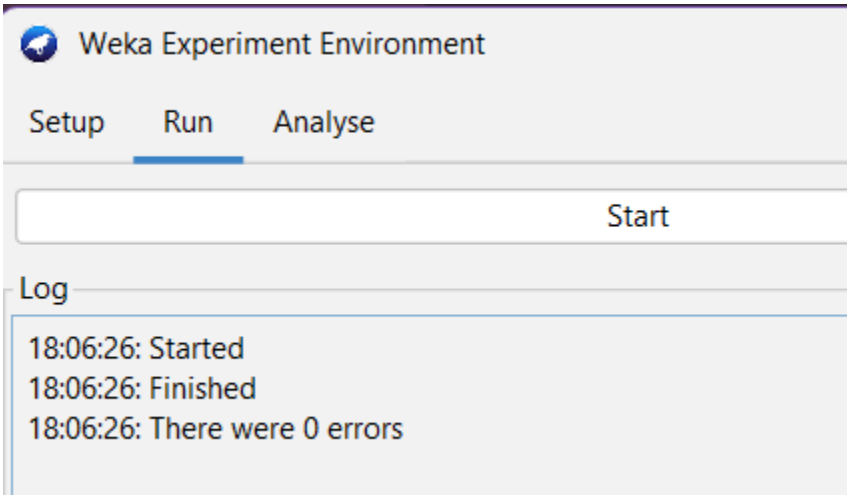


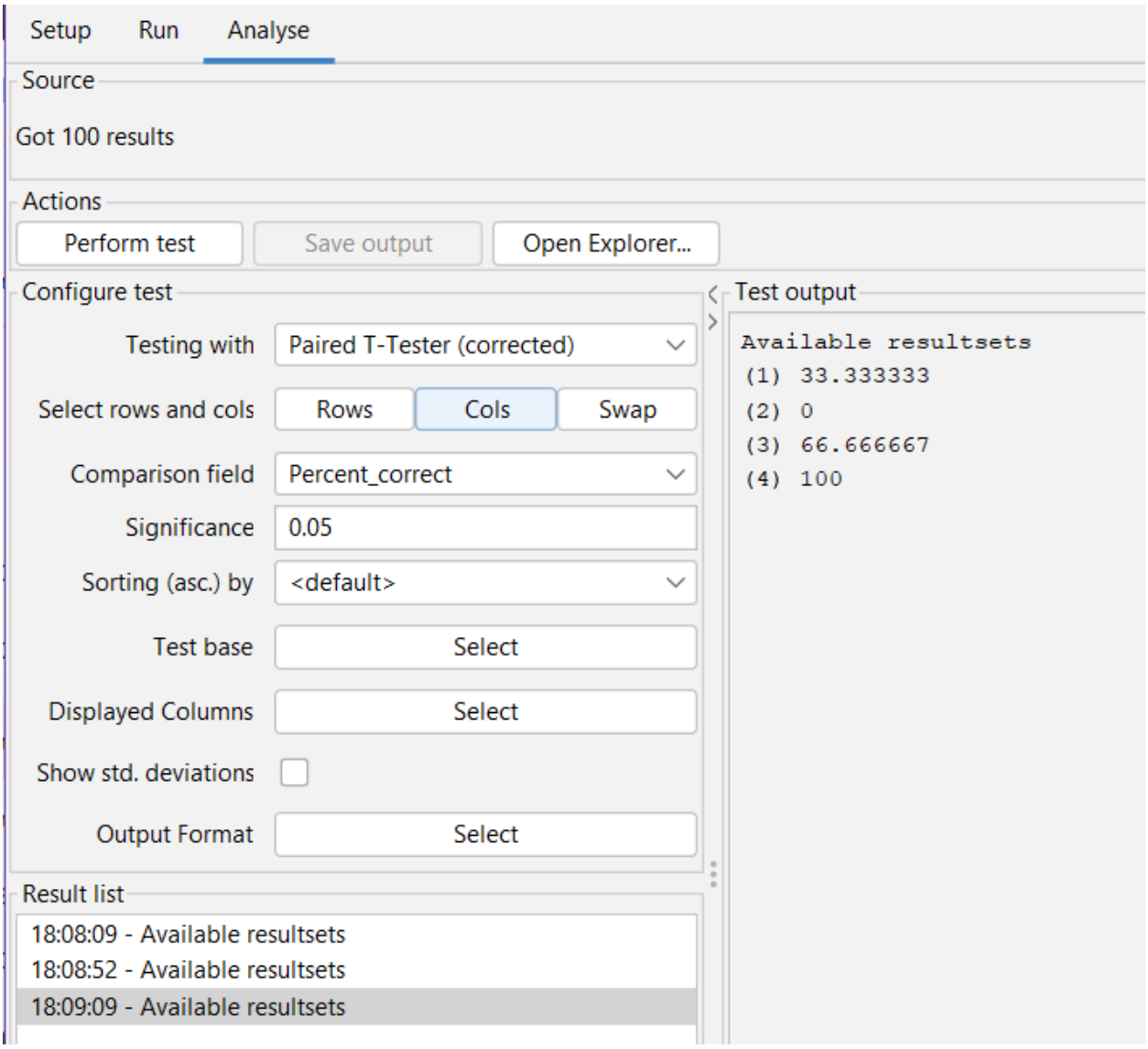


(v) Also, Find Out The Percentage of Incorrect Classification for Each Folds using Weka Experiment Environment.

=>






Setup
Run
Analyse
Source
Got 100 results
Actions
Perform test
Save output
Open Explorer...
Configure test
Testing with
Paired T-Tester (corrected)
Select rows and cols
Rows
Cols
Swap
Comparison field
Percent_correct
Significance
0.05
Sorting (asc.) by
<default>
Test base
Select
Displayed Columns
Select
Show std. deviations
Output Format
Select
Result list
18:08:09 - Available resultsets
18:08:52 - Available resultsets
18:09:09 - Available resultsets
Test output
Available resultsets
(1) 33.333333
(2) 0
(3) 66.666667
(4) 100