Experiment-2.2

Student Name: Ashish Kumar
UID: 23MAI10008
Branch: CSE AIML
Section/Group: 23MAI-1
Semester: 01
Date of Performance: 09/10/2023
Subject Name: Artificial Intelligence Lab
Subject Code: 23CSH-621

Aim of the Experiment :

Aim of the Experiment is to apply different testing techniques for the UCI dataset and compare the performance of the best prediction model with the following techniques:

a) 70/30 Split

b) K-fold cross validation

Objective of the Experiment :

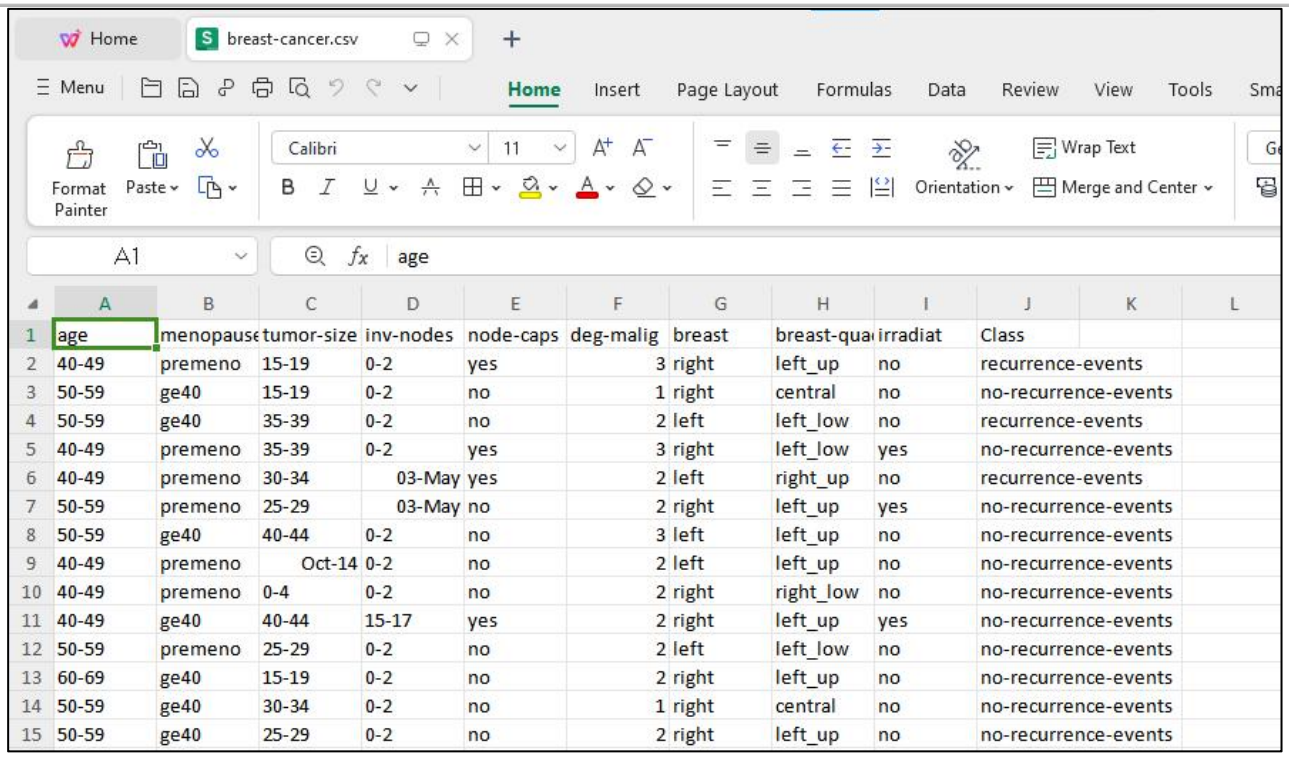
Task to be done for this experiment is that we have to use the UCI dataset and find the best prediction model. Then we have to apply the different testing techniques for the UCI dataset and compare the performance of the best prediction model with the help of the following techniques:

a) 70//30 Split

b) K-fold cross validation

Algorithm/ Steps for Experiment :

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



Home | breast-cancer.csv

Menu | Home | Insert | Page Layout | Formulas | Data | Review | View | Tools

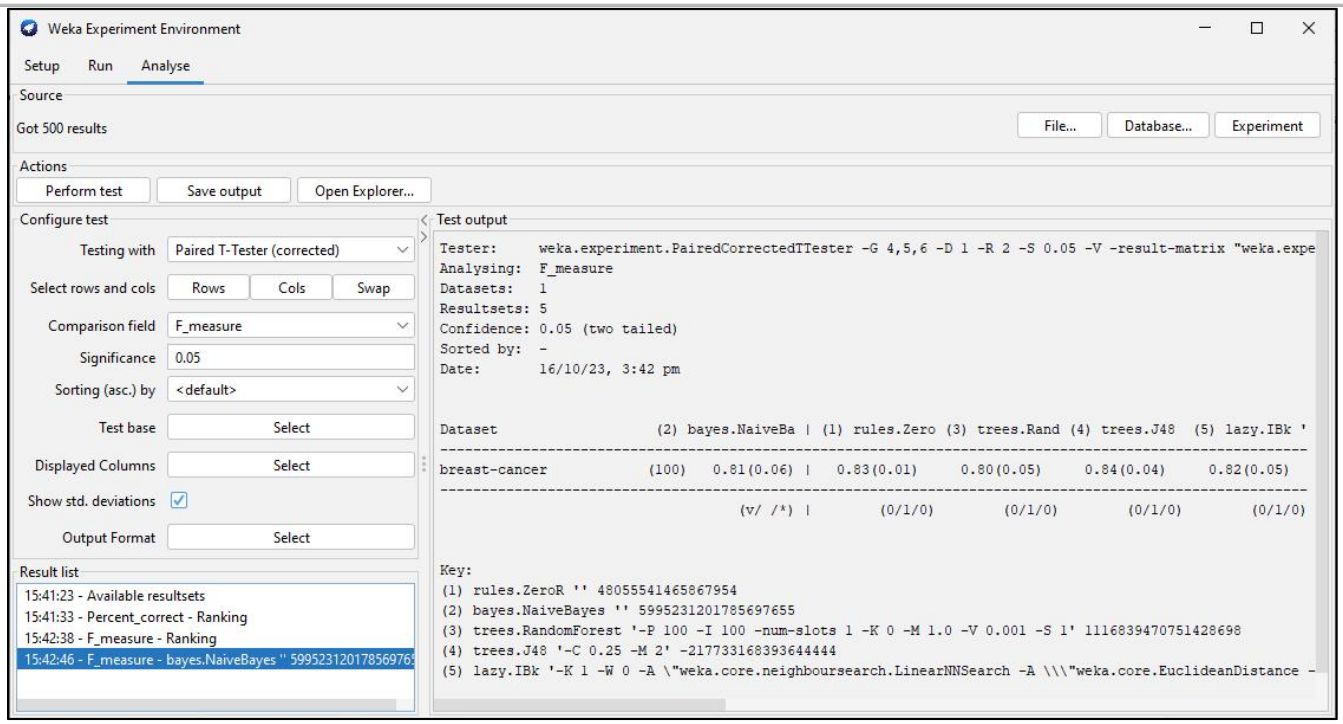
Format Painter | Paste | Calibri | 11 | Wrap Text | Orientation | Merge and Center

A1 | age

	A	B	C	D	E	F	G	H	I	J	K	L
1	age	menopause	tumor-size	inv-nodes	node-caps	deg-malig	breast	breast-qua	irradiat	Class		
2	40-49	premeno	15-19	0-2	yes	3	right	left_up	no	recurrence-events		
3	50-59	ge40	15-19	0-2	no	1	right	central	no	no-recurrence-events		
4	50-59	ge40	35-39	0-2	no	2	left	left_low	no	recurrence-events		
5	40-49	premeno	35-39	0-2	yes	3	right	left_low	yes	no-recurrence-events		
6	40-49	premeno	30-34	03-May	yes	2	left	right_up	no	recurrence-events		
7	50-59	premeno	25-29	03-May	no	2	right	left_up	yes	no-recurrence-events		
8	50-59	ge40	40-44	0-2	no	3	left	left_up	no	no-recurrence-events		
9	40-49	premeno	Oct-14	0-2	no	2	left	left_up	no	no-recurrence-events		
10	40-49	premeno	0-4	0-2	no	2	right	right_low	no	no-recurrence-events		
11	40-49	ge40	40-44	15-17	yes	2	right	left_up	yes	no-recurrence-events		
12	50-59	premeno	25-29	0-2	no	2	left	left_low	no	no-recurrence-events		
13	60-69	ge40	15-19	0-2	no	2	right	left_up	no	no-recurrence-events		
14	50-59	ge40	30-34	0-2	no	1	right	central	no	no-recurrence-events		
15	50-59	ge40	25-29	0-2	no	2	right	left_up	no	no-recurrence-events		

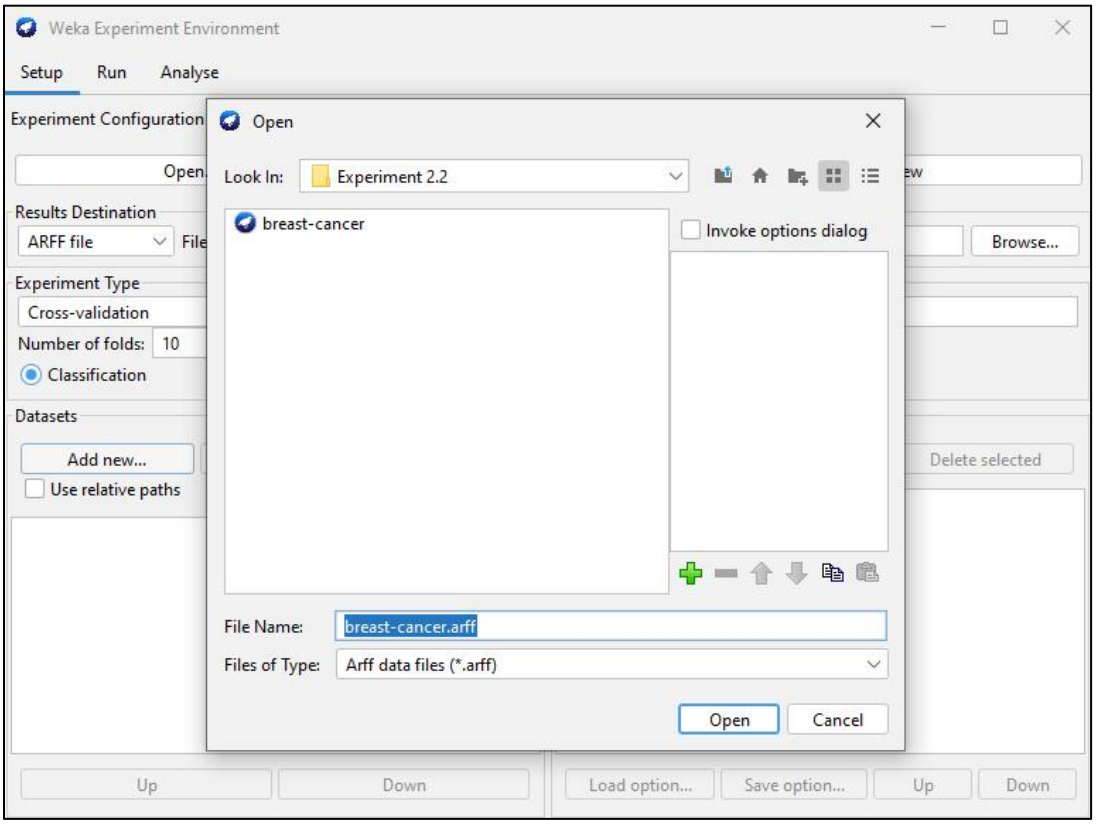
Step 2: Open WEKA and instead of 'Explorer' tab, open the **'Experimenter'** tab.

Step 3: Find the best prediction model out of the applied 5 classifiers which is J48 Classifier.

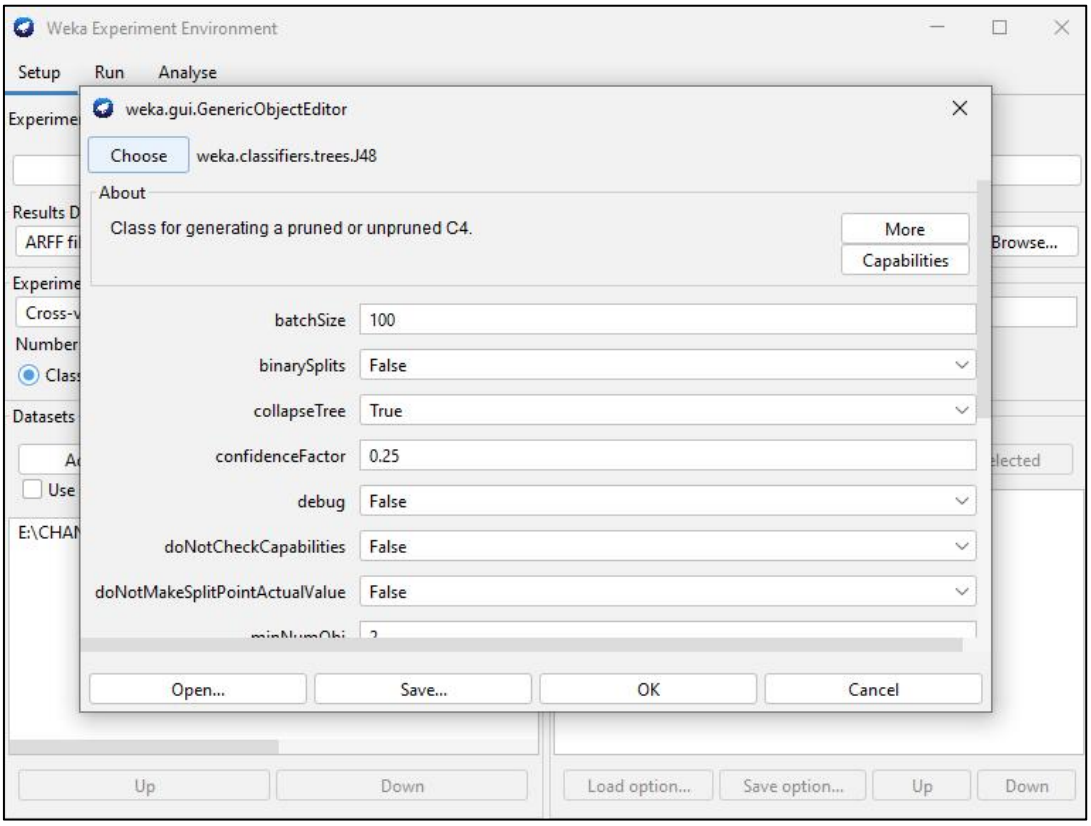


Step 4: In **'Setup'** tab, Click on the **'New'** option to initiate the process of adding data.

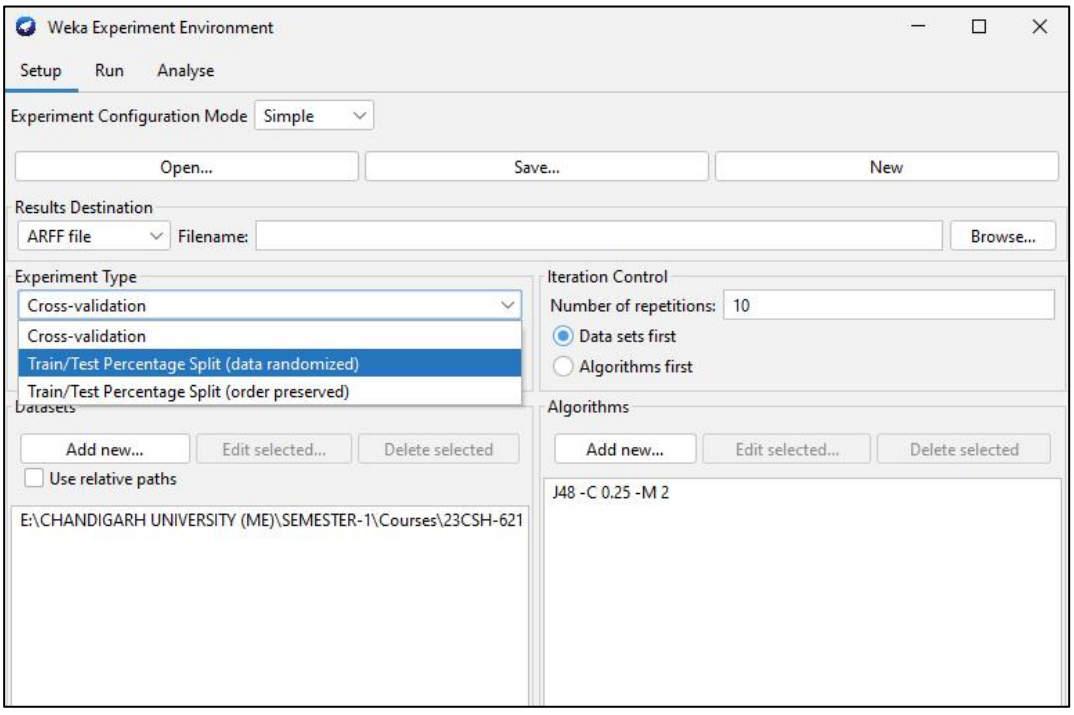
Step 5: In **'Datasets'** section, click on **'Add New'** option and select arff file to import.



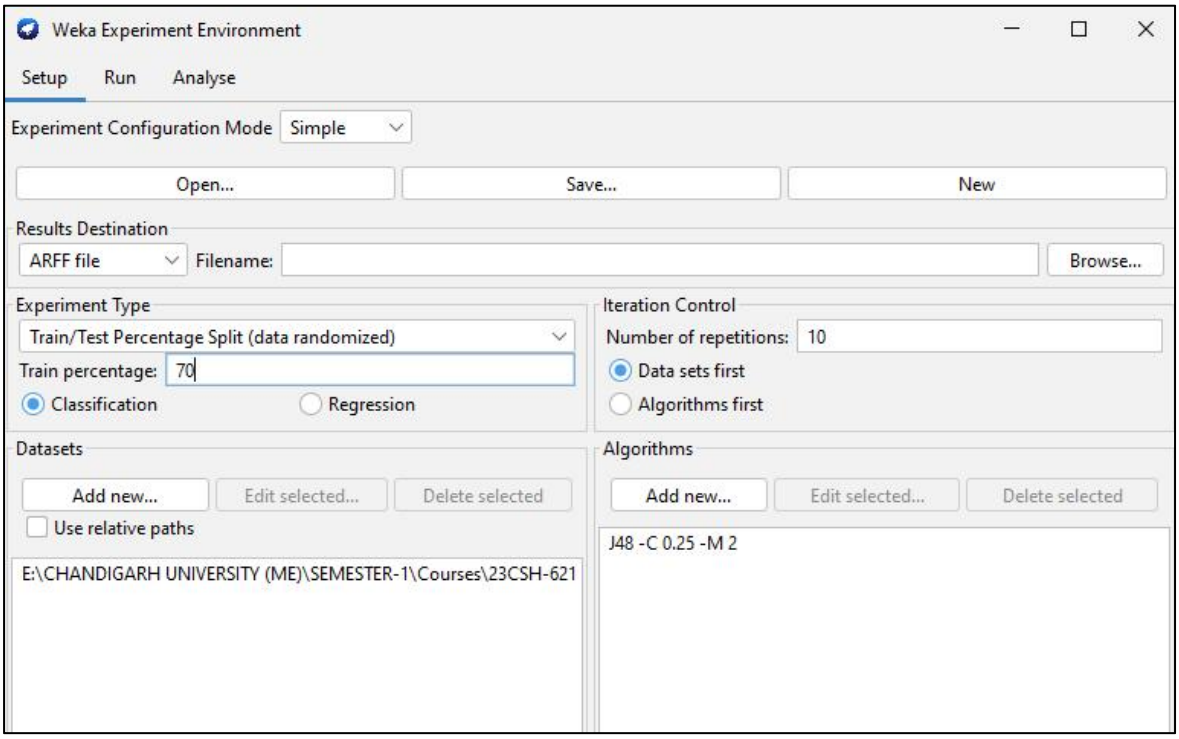
Step 6: In the **Algorithms** section,click on choose add new> choose>>**J48**.



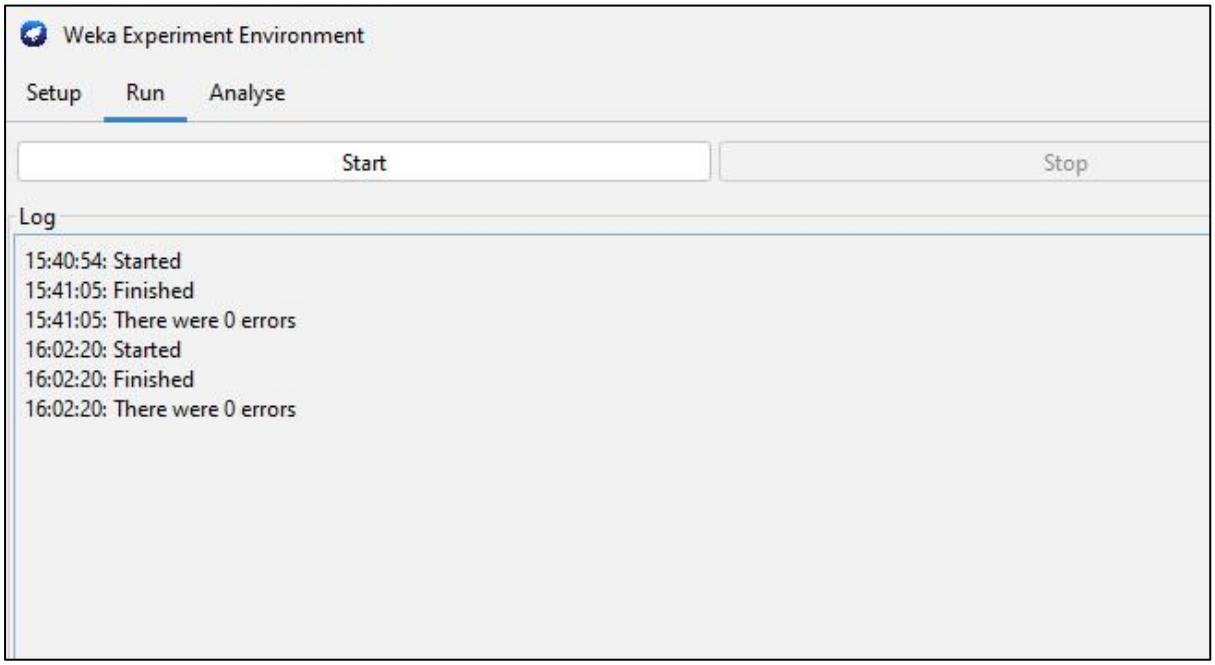
Step 7: In **'Experiment Type'** section, click on **'Train/Test Percentage Split (data randomized)'**.



Step 8: In **'Train Percentage'** option, add the value 70 for 70/30 Split.

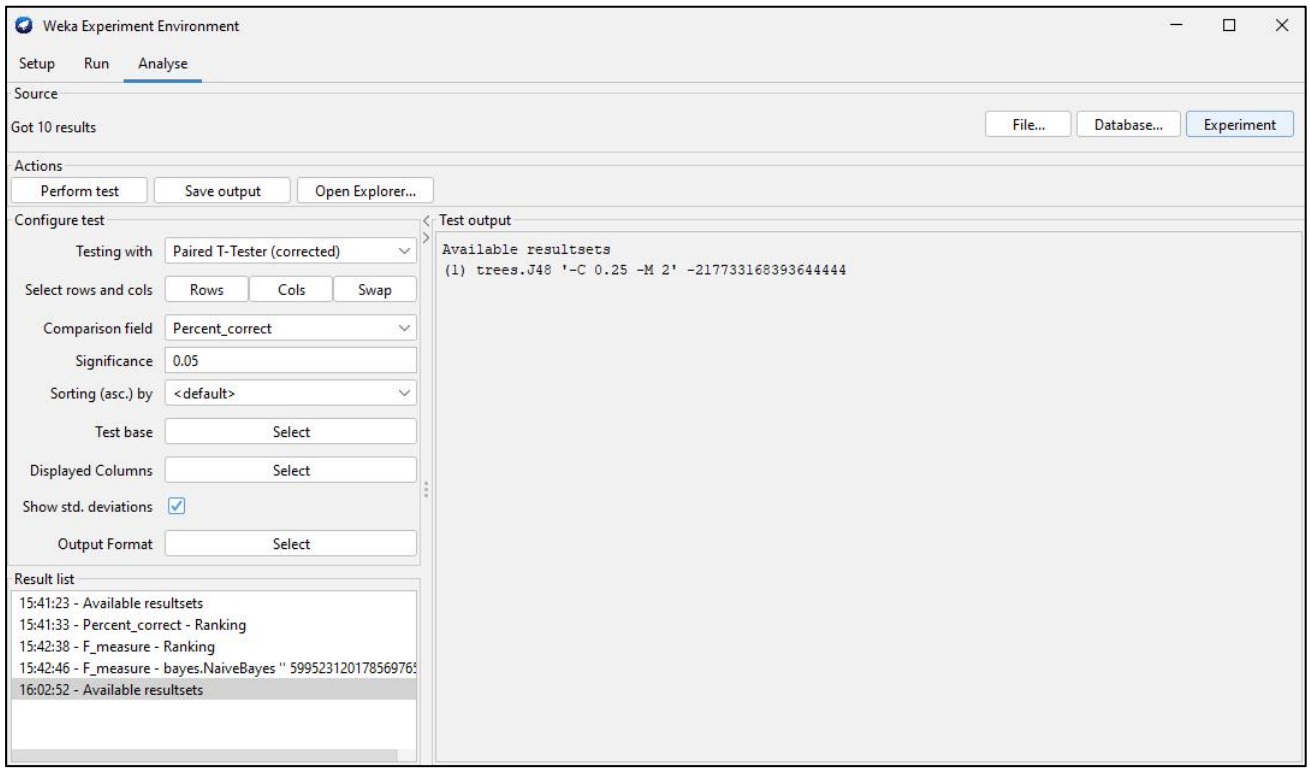


Step 9: Click on **'Run'** tab >> click **'Start'**.

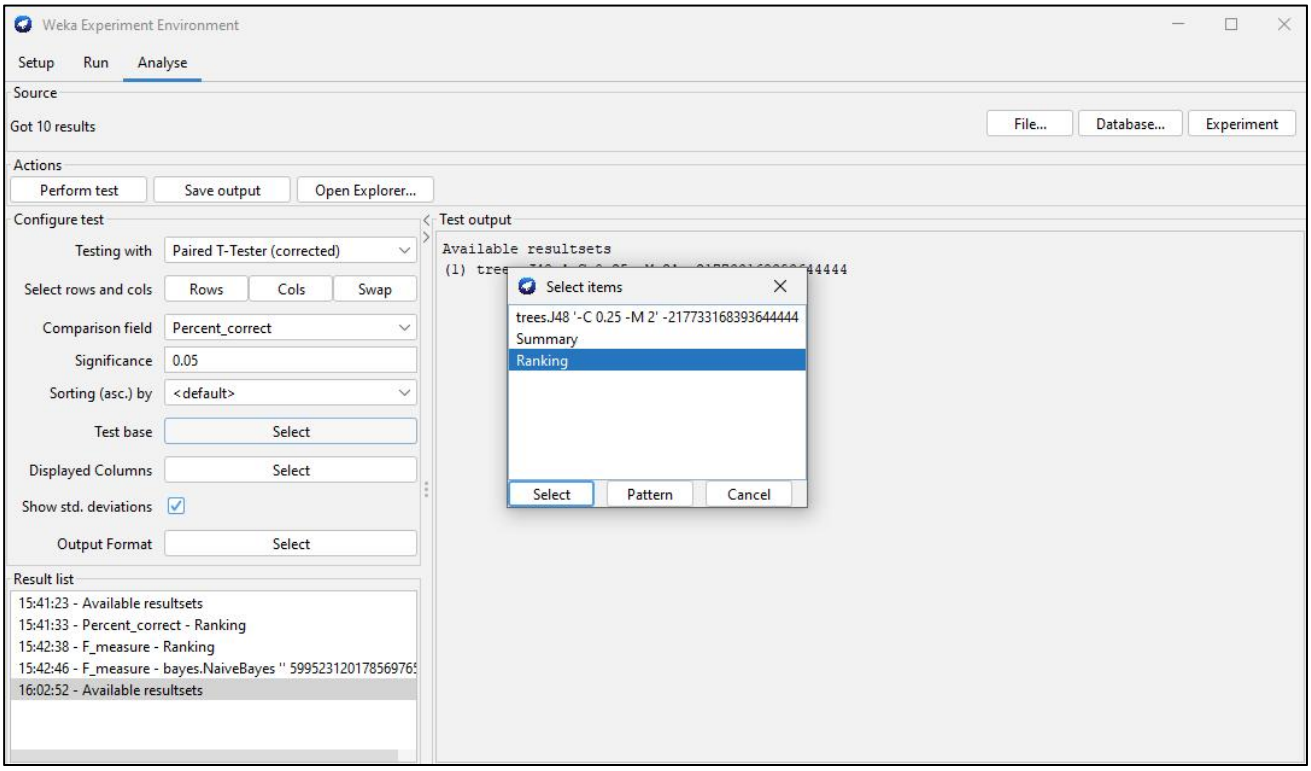


Step 10: After the execution is completed in 'Run' tab >>click on **'Analyse'** tab.

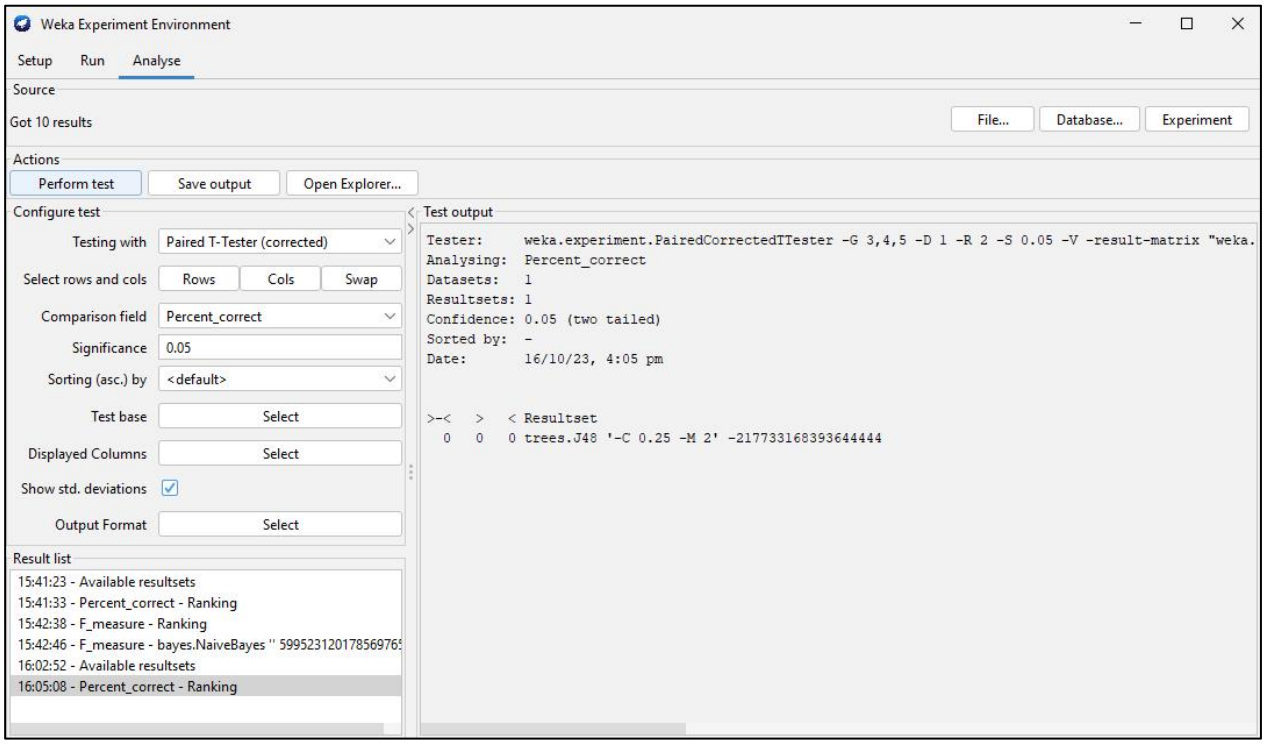
Step 11: In the **'Analyse'** tab >> click on **'Experiment'** tab >> all the options will become active.



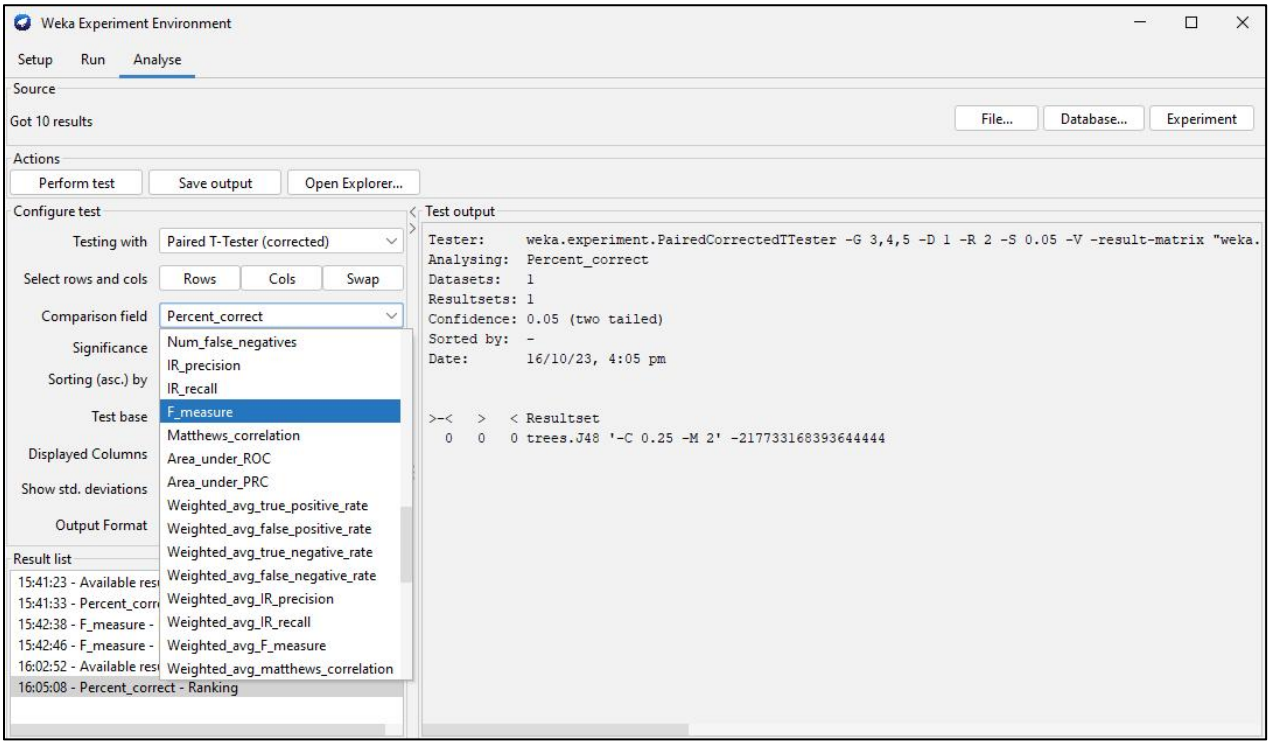
Step 12: In **Configure test** >> click on **Test base**>> and select **Ranking**>>click on **Perform test**.

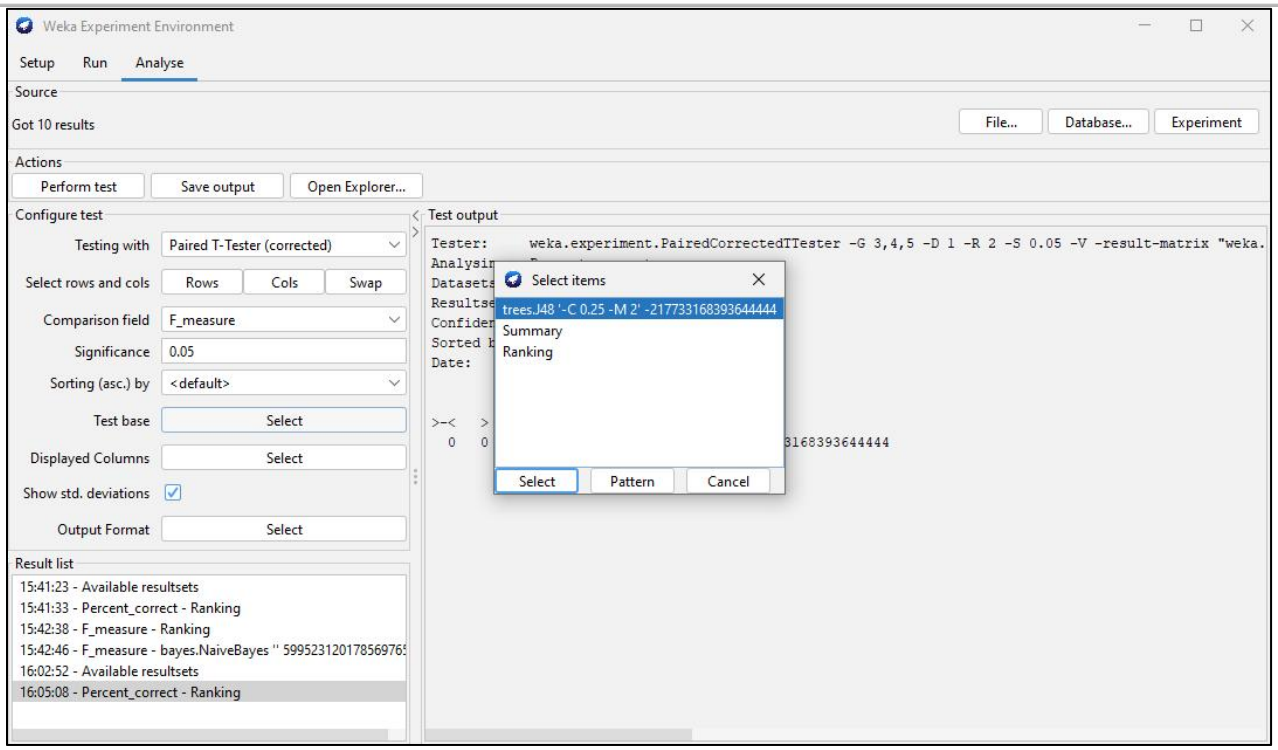


Step 13: In **Test area** ,it will show algorithms which have performed better and will display ranking.

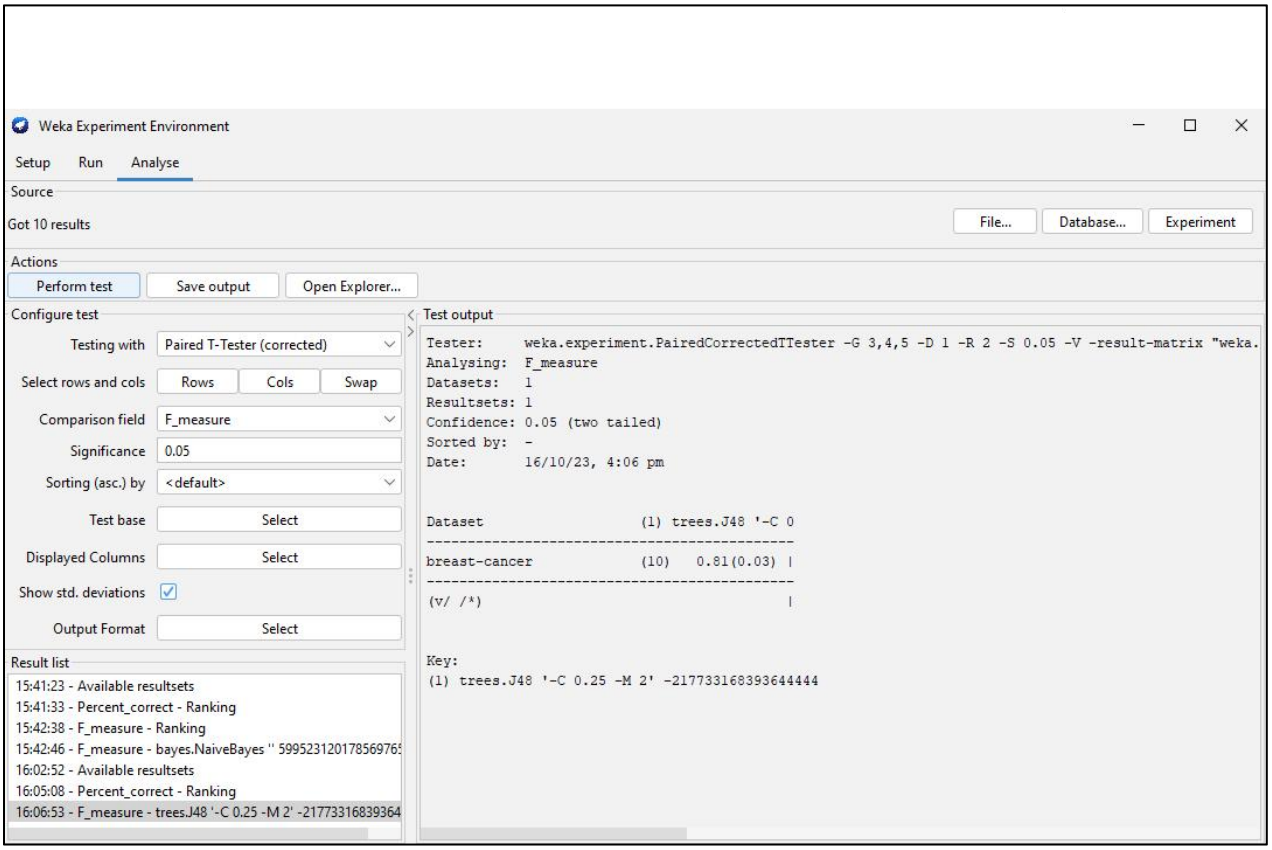


Step 14: To check the **f-measure** and std_dev of an particular algorithm >>click on std_tab >>then select the algorithm which you want to test >> select **f-measure** from the **comparison_field** tab.

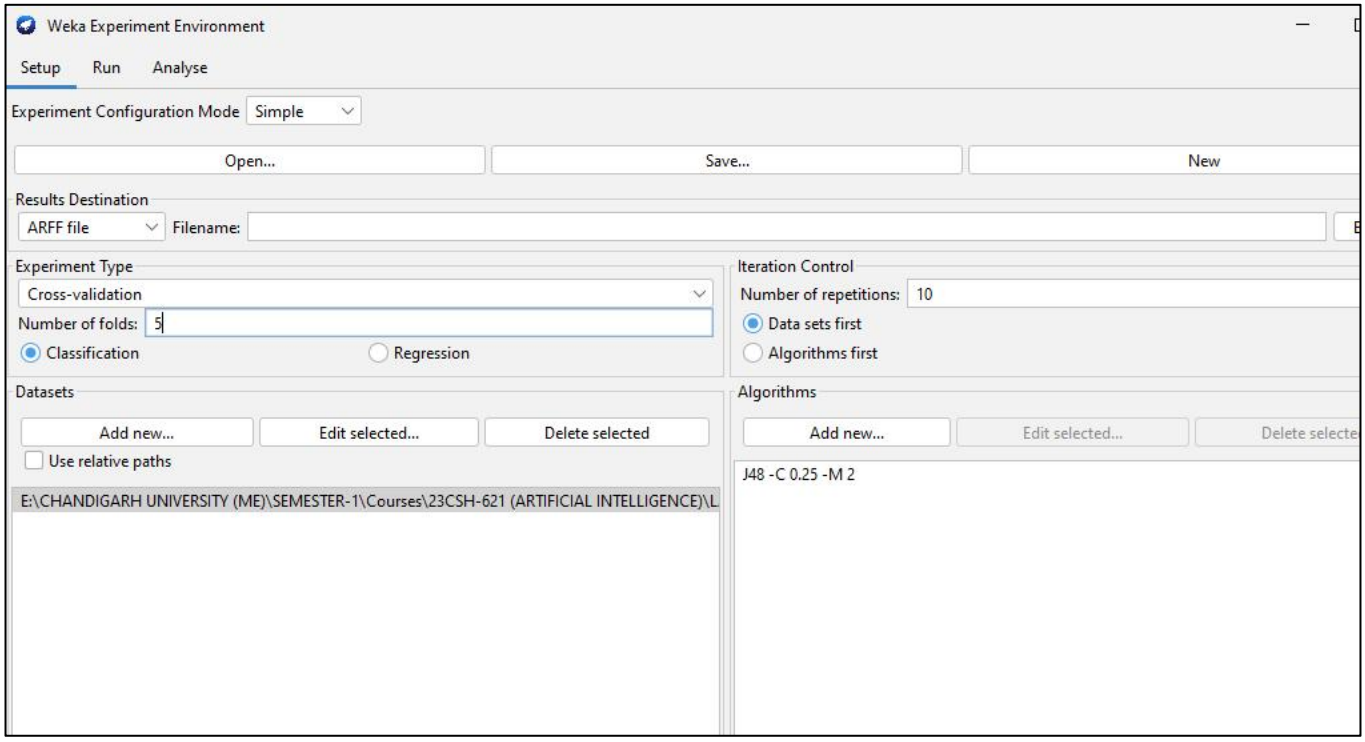




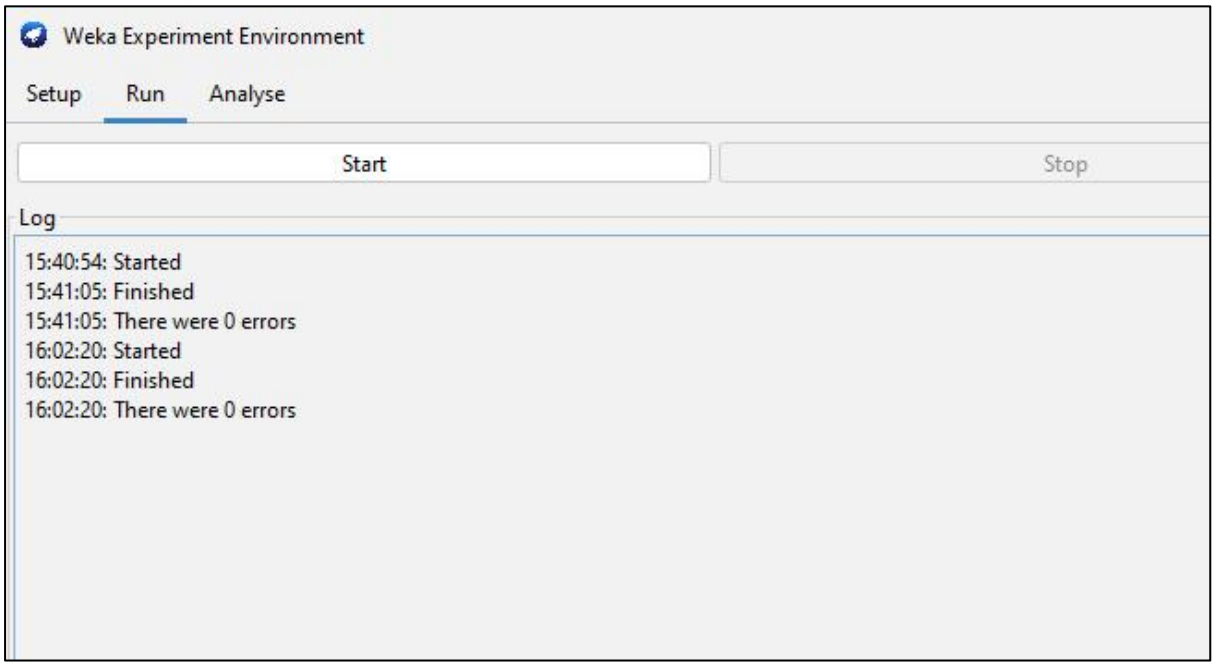
Step 15: Click on **Perform test**.



Step 16: In **'Setup'** tab >> **'Experiment Type'**, click on **'Cross-validation'** and set no of folds as 5.

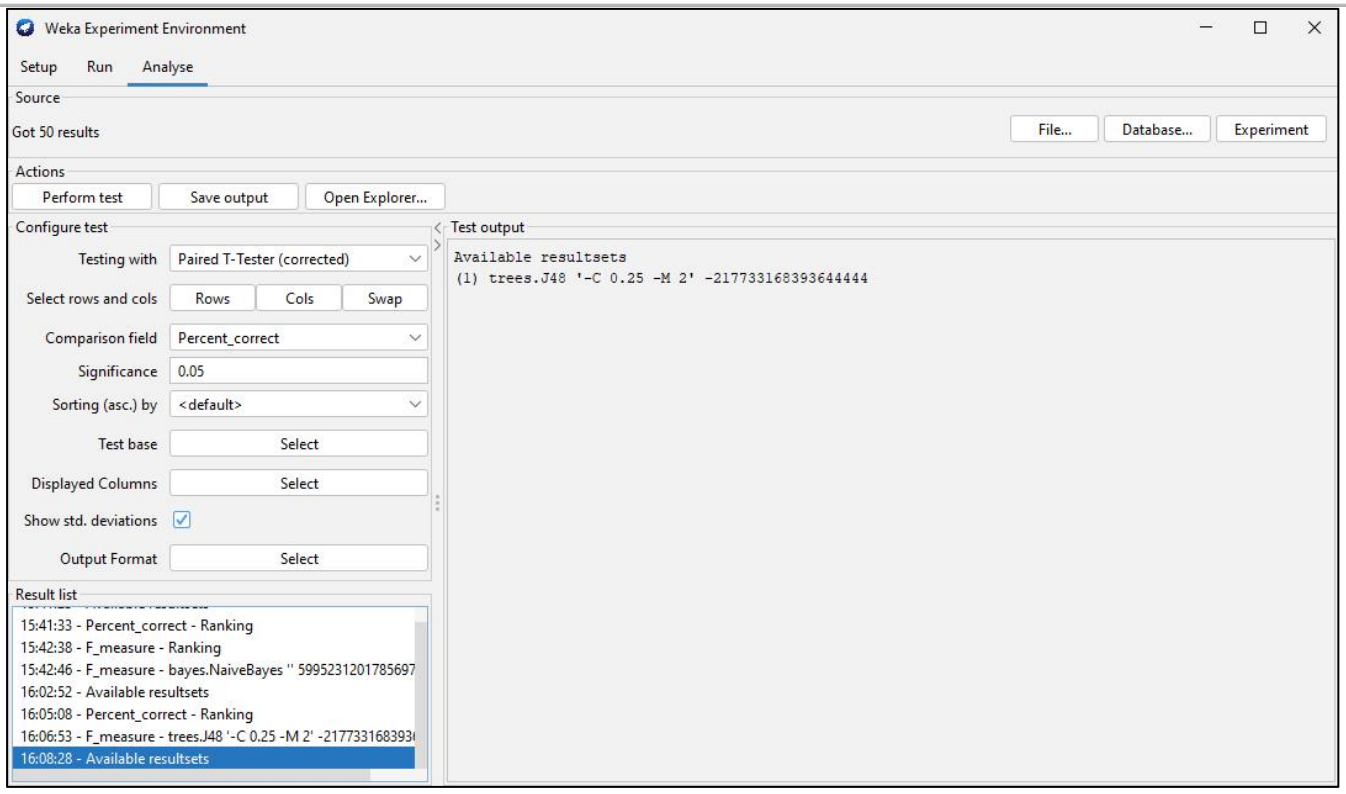


Step 17: Click on **'Run'** tab >> click **'Start'**.

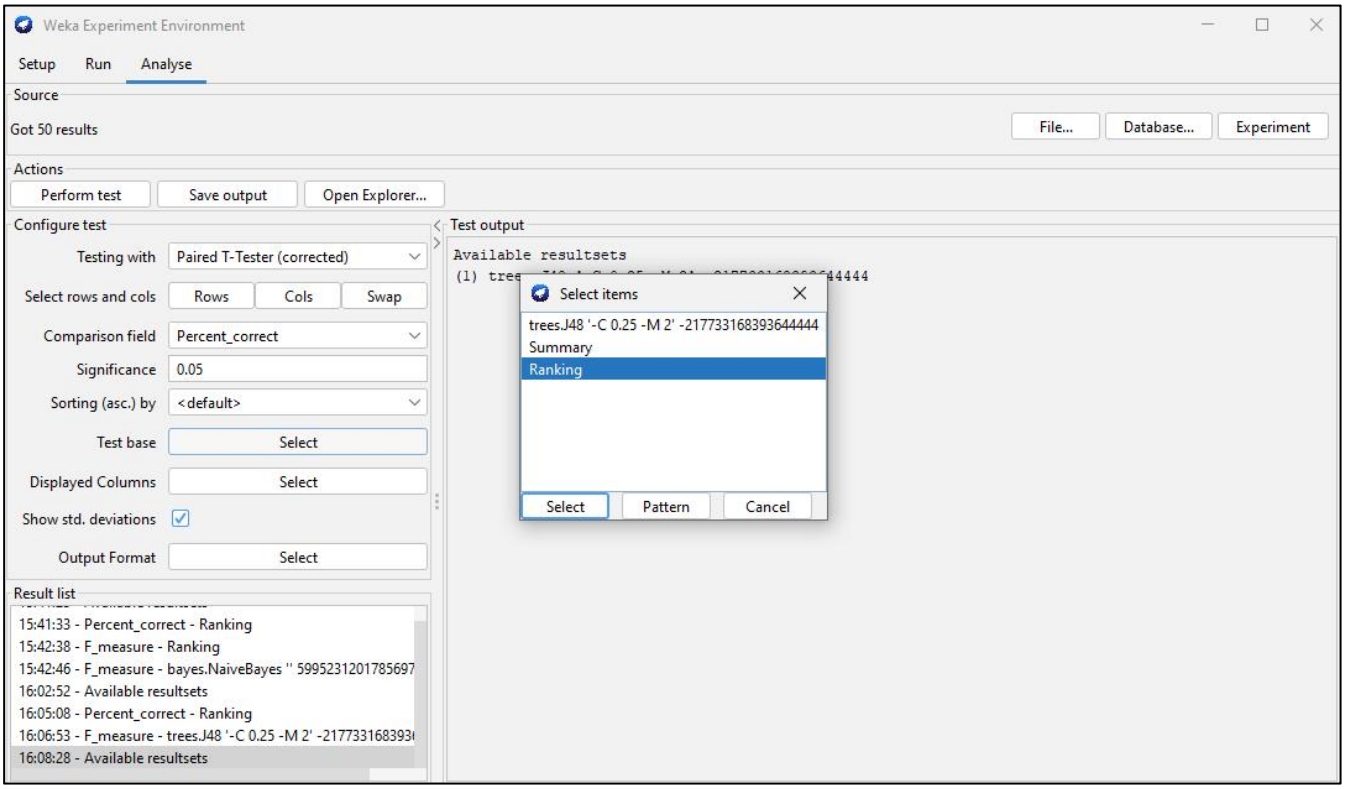


Step 18: After the execution is completed in 'Run' tab >>click on **'Analyse'** tab.

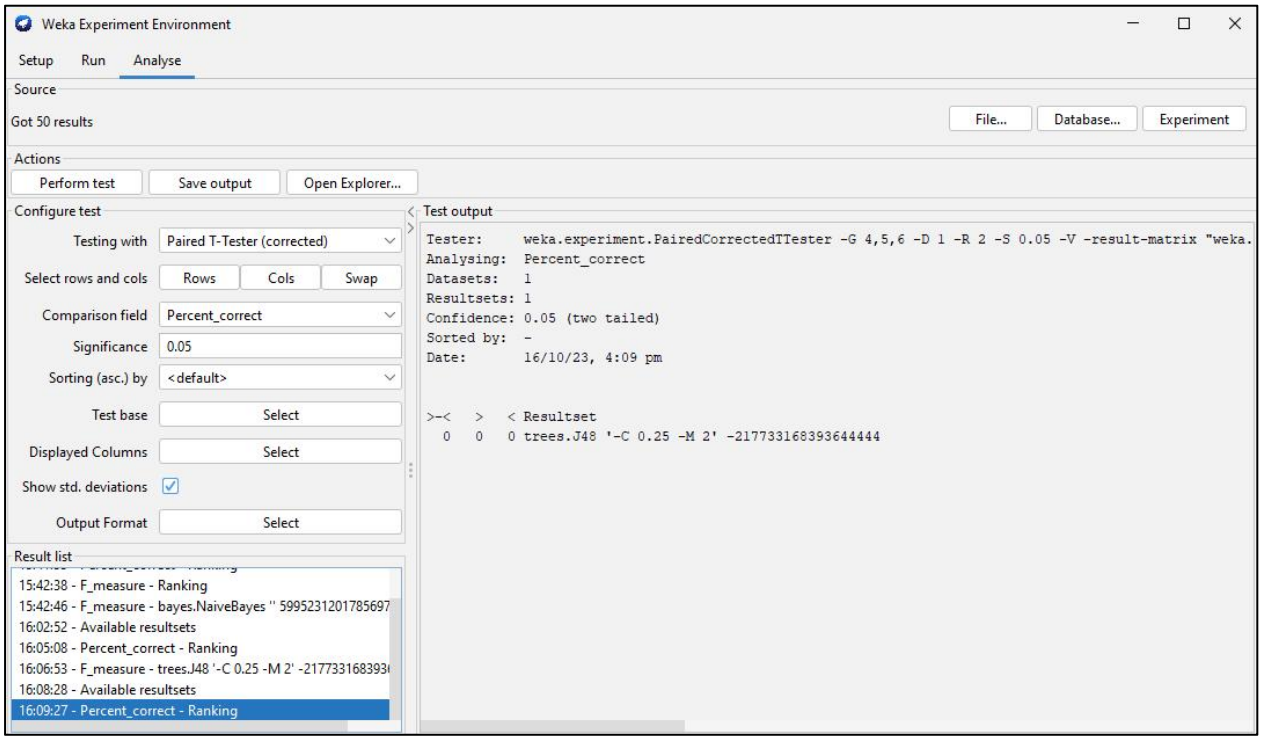
Step 19: In the **'Analyse'** tab >> click on **'Experiment'** tab >> all the options will become active.



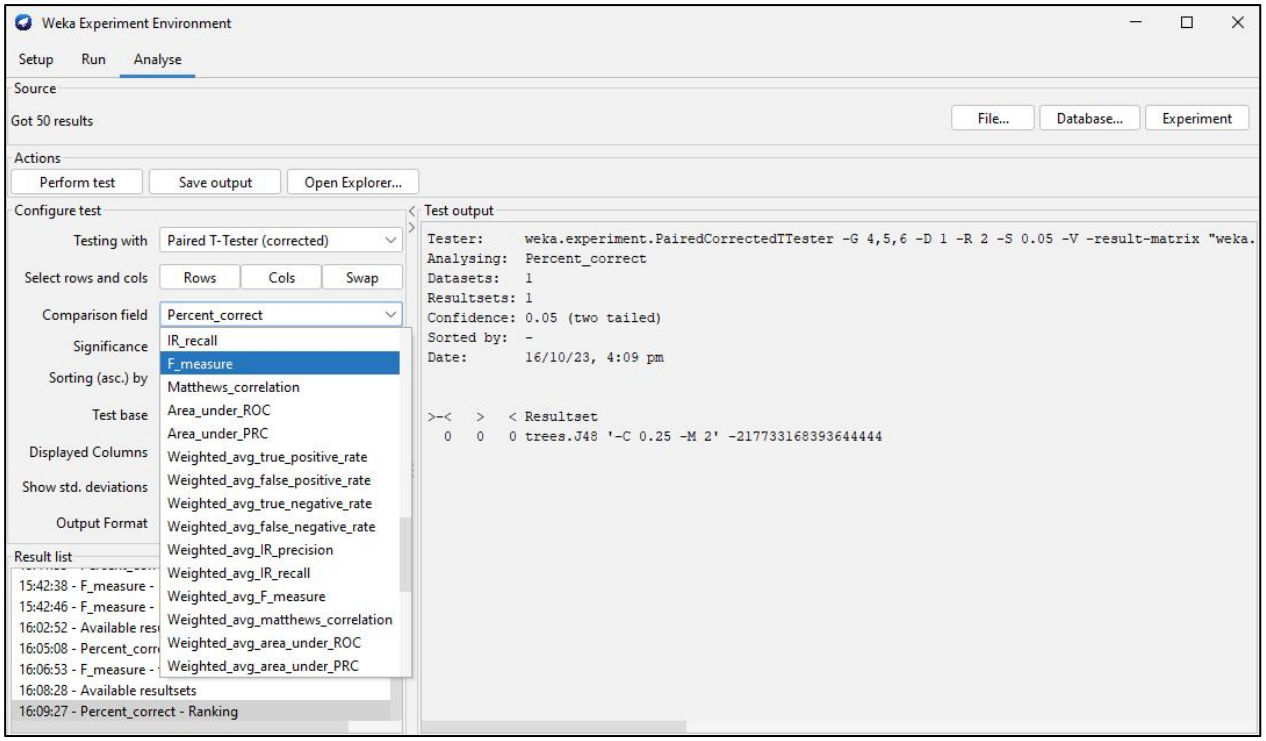
Step 20: In **Configure test** >> click on **Test base**>> and select **Ranking**>>click on **Perform test**.



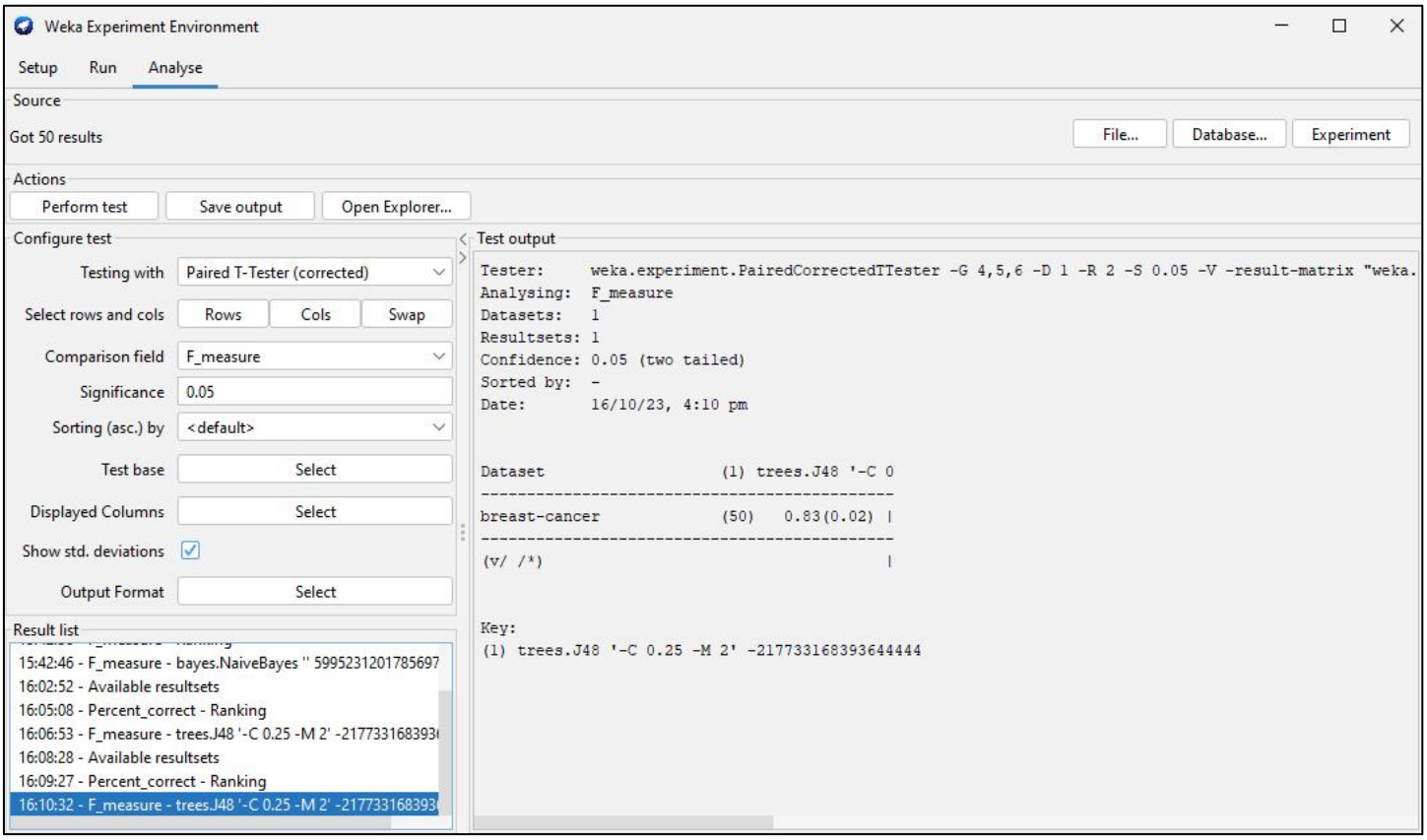
Step 21: In **Test area** ,it will show algorithms which have performed better and will display ranking.



Step 22: To check the **f-measure** and std_dev of an particular algorithm >>click on std_tab >>then select the algorithm which you want to test >> select **f-measure** from the **comparison_field** tab.



Step 23: Click on **Perform test**.



Learning outcomes (What I have learnt):

1. I learnt about the WEKA Tool and its applications.
2. I learnt about different machine learning classifiers in WEKA Tool.
3. I learnt about how to find the best prediction model in WEKA.
4. I learnt about the Train/Test Percentage Split Method to check model performance.
5. I learnt about the Cross-validation method to check model performance.