

**CSA1618 DWDM-DE**

**EXPERIMENT-29**

**EVALUATING ACCURACY OF THE CLASSIFIERS**

**AIM:**

To create evaluating accuracy of the classifiers using weka tool.

**PROCEDURE:**

1. Download and install WEKA.
2. Open WEKA and Choose "Explorer" from the main menu.
3. Under Preprocess, Click on the open file button and select the dataset. Ensure that your dataset contains categorical (nominal) attributes.
4. Go to the Classify tab.
5. Click Choose and select a classifier. Examples:

J48 (Decision Tree) → trees > J48

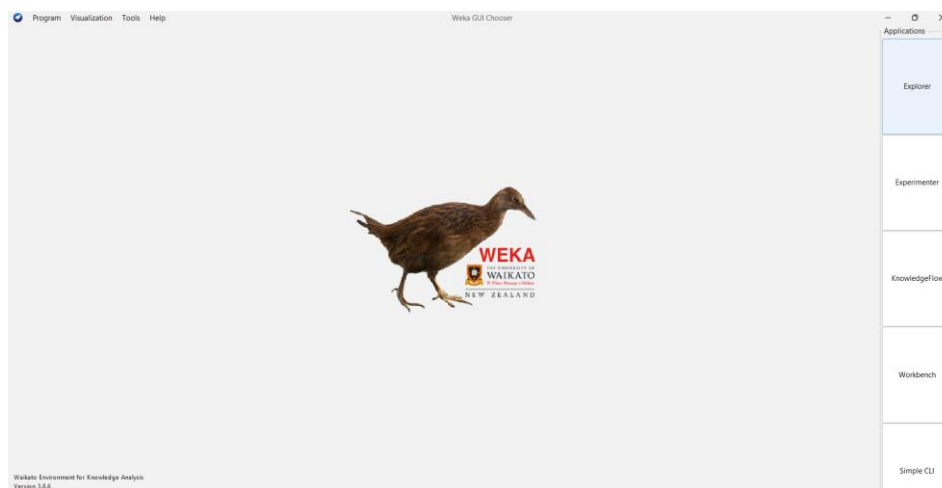
Naïve Bayes → bayes > NaiveBayes

SMO (SVM) → functions > SMO

Random Forest → trees > RandomForest

Logistic Regression → functions > Logistic

6. Select Evaluation Method : Cross-validation.
7. Click Start to begin classification. The Classifier output section will display results.



**Weka Explorer**

Preprocess   **Classify**   Cluster   Associate   Select attributes   Visualize

Open file...   Open URL...   Open DB...   Generate...   Undo   Edit...   Save...

Filter: Choose **None**   Apply   Stop

Current relation  
Relation: weather.symbolic  
Instances: 14

Attributes: 5   Sum of weights: 14

Attributes: All   None   Invert   Pattern

No.	Name
1	<input type="checkbox"/> outlook
2	<input type="checkbox"/> temperature
3	<input type="checkbox"/> humidity
4	<input type="checkbox"/> windy
5	<input type="checkbox"/> play

Remove

Status: OK

Selected attribute  
Name: outlook  
Missing: 0 (0%)   Distinct: 3   Type: Nominal  
Unique: 0 (0%)

No.	Label	Count	Weight
1	sunny	5	5
2	overcast	4	4
3	rainy	5	5

Class: play (Nom)   Visualize All

Log   x 0

## 1.J48 (Decision Tree):

**Weka Explorer**

Preprocess   **Classify**   Cluster   Associate   Select attributes   Visualize

Classifier: Choose **J48 - C 0.25 - M 2**

Test options:  
☐ Use training set   Set...  
☐ Supplied test set   Set...  
☒ Cross-validation   Folds: 10  
☐ Percentage split   %: 66  
 More options...

(Nom) play   Start   Stop

Result list (right-click for options):  
 10/3/20 - trees.J48

Classifier output:  

```

outlook = overcast: yes (4.0)
outlook = rainy
| windy = TRUE: no (2.0)
| windy = FALSE: yes (3.0)

Number of Leaves : 5
Size of the tree : 8

Time taken to build model: 0.01 seconds

=== Stratified cross-validation ===
=== Summary ===
Correctly Classified Instances      7      50 %
Incorrectly Classified Instances    7      50 %
Kappa statistic                    -0.0426
Mean absolute error                 0.4167
Root mean squared error             0.5984
Relative absolute error             87.5 %
Root relative squared error        121.2987 %
Total Number of Instances          14

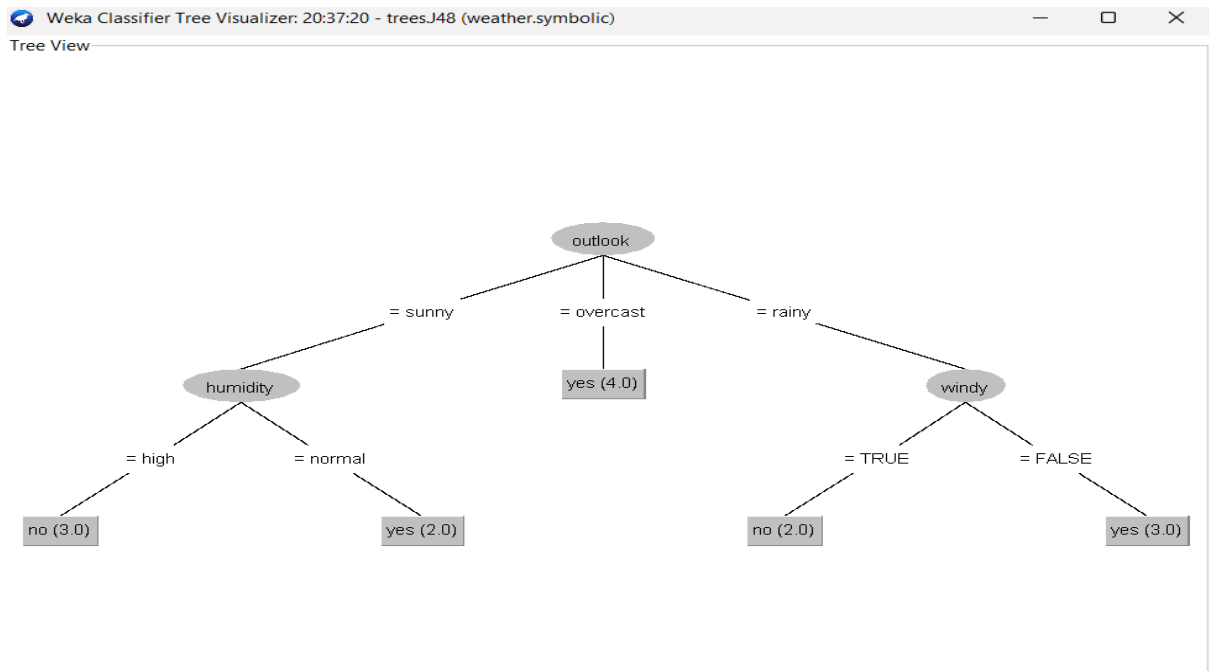
=== Detailed Accuracy By Class ===
               TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
0.556   0.600   0.625   0.556   0.588   -0.043   0.633   0.758   yes
0.400   0.444   0.333   0.400   0.364   -0.043   0.633   0.457   no
Weighted Avg.   0.500   0.544   0.521   0.500   0.508   -0.043   0.633   0.650

=== Confusion Matrix ===
 a b  <-- classified as
 5 4 | a = yes
 3 2 | b = no

```

Status: OK

Log   x 0



### === Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
	0.556	0.600	0.625	0.556	0.588	-0.043	0.633	0.758
	0.400	0.444	0.333	0.400	0.364	-0.043	0.633	0.457
Weighted Avg.	0.500	0.544	0.521	0.500	0.508	-0.043	0.633	0.650

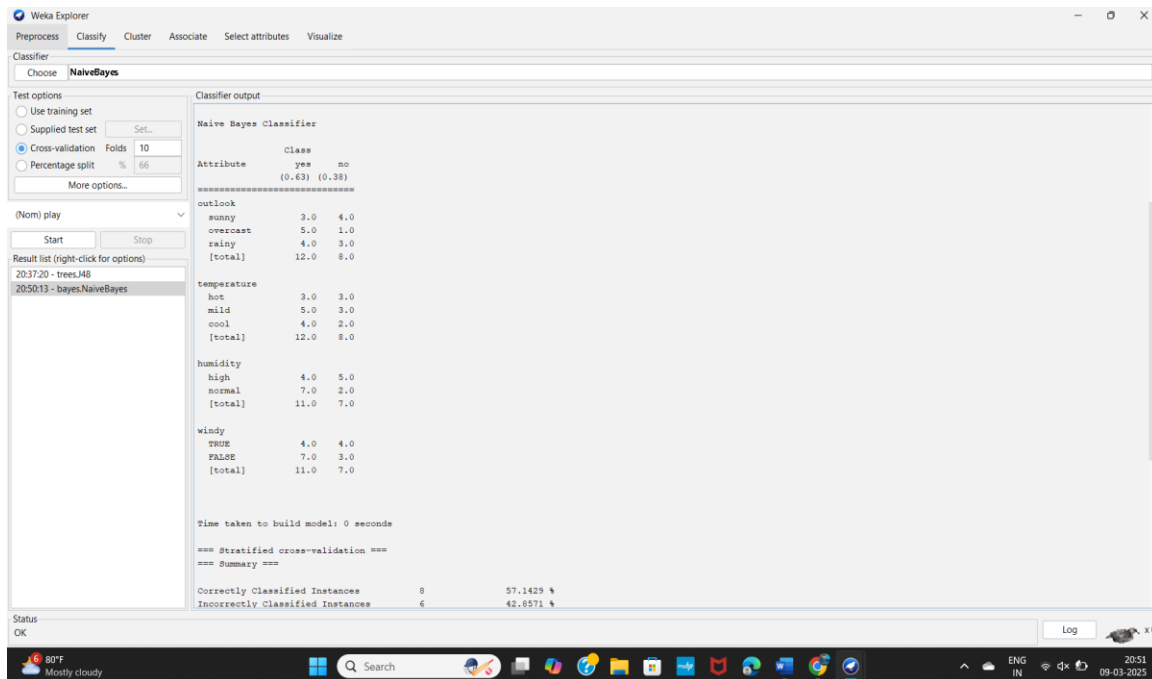
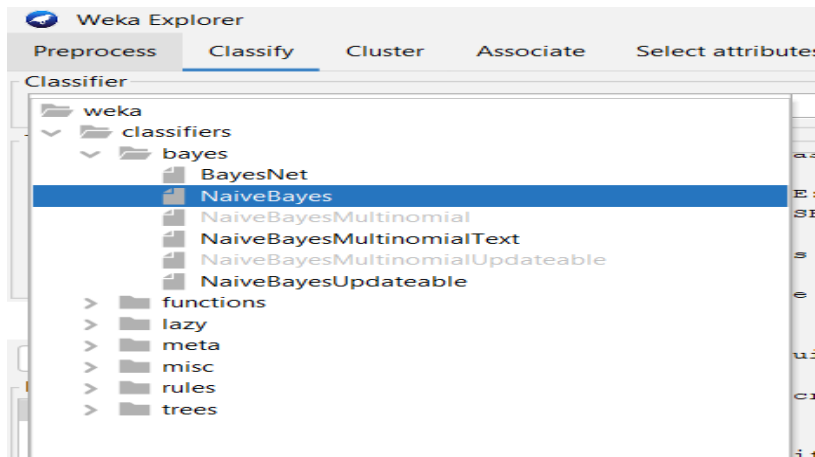
### === Confusion Matrix ===

a b <-- classified as

5 4 | a = yes

3 2 | b = no

## 2.Naive Bayes:



### === Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
yes	0.778	0.800	0.636	0.778	0.700	-0.026	0.578	0.697
no	0.200	0.222	0.333	0.200	0.250	-0.026	0.578	0.557
Wt Avg.	0.571	0.594	0.528	0.571	0.539	-0.026	0.578	0.647

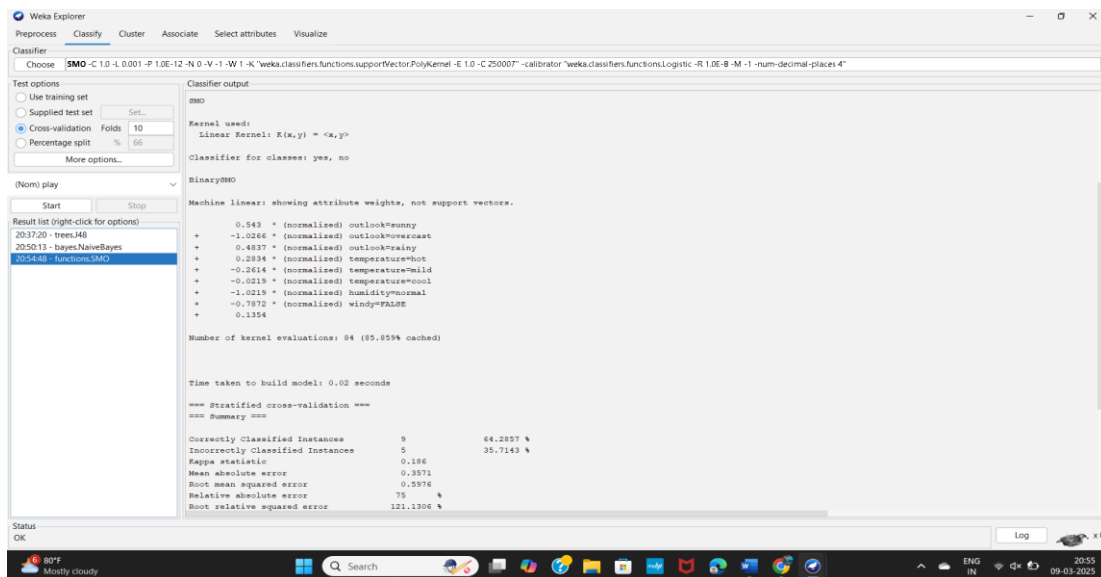
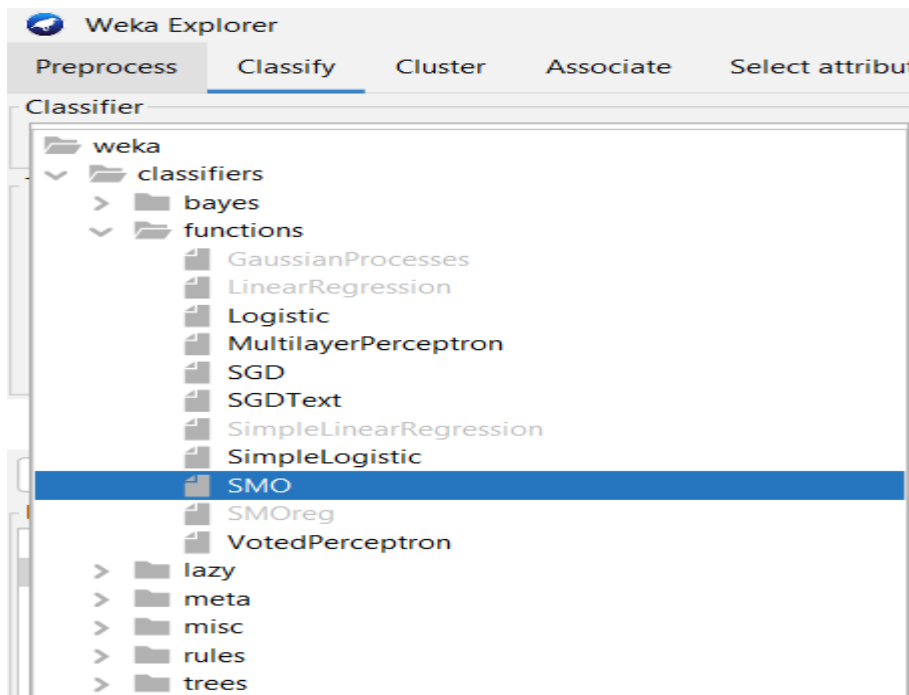
### === Confusion Matrix ===

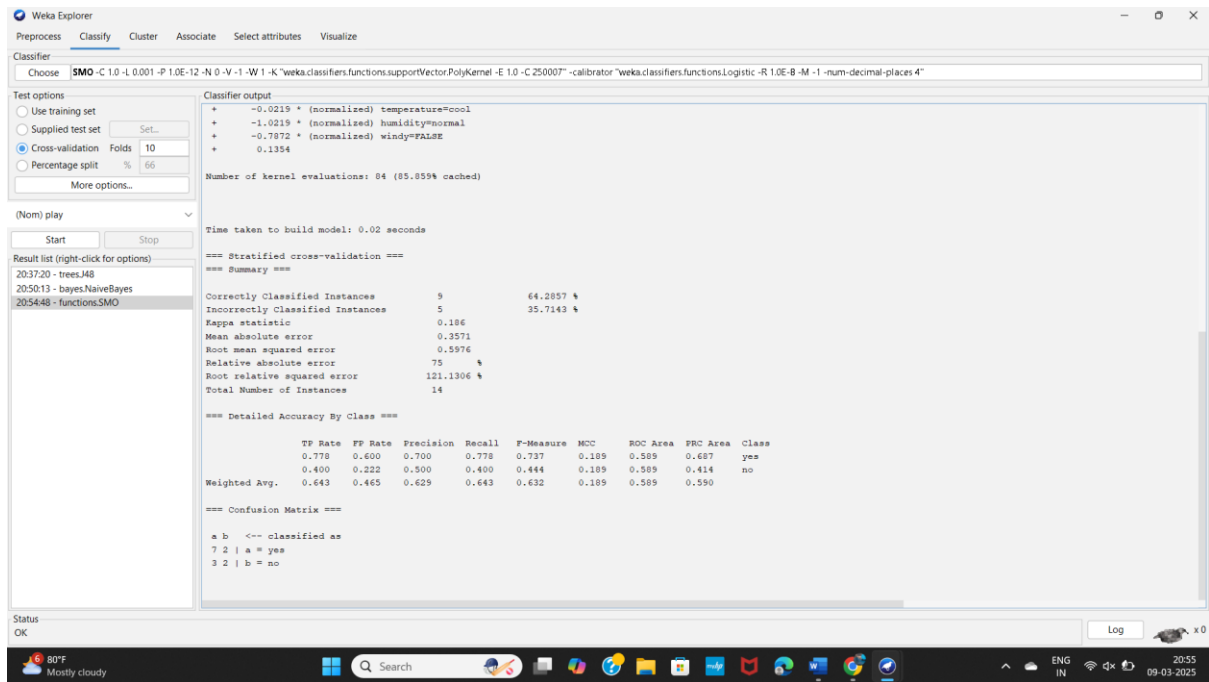
a b <-- classified as

7 2 | a = yes

4 1 | b = no

### 3. SMO:





### === Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	
	0.778	0.600	0.700	0.778	0.737	0.189	0.589	0.687	yes
	0.400	0.222	0.500	0.400	0.444	0.189	0.589	0.414	no
Wt Avg.	0.643	0.465	0.629	0.643	0.632	0.189	0.589	0.590	

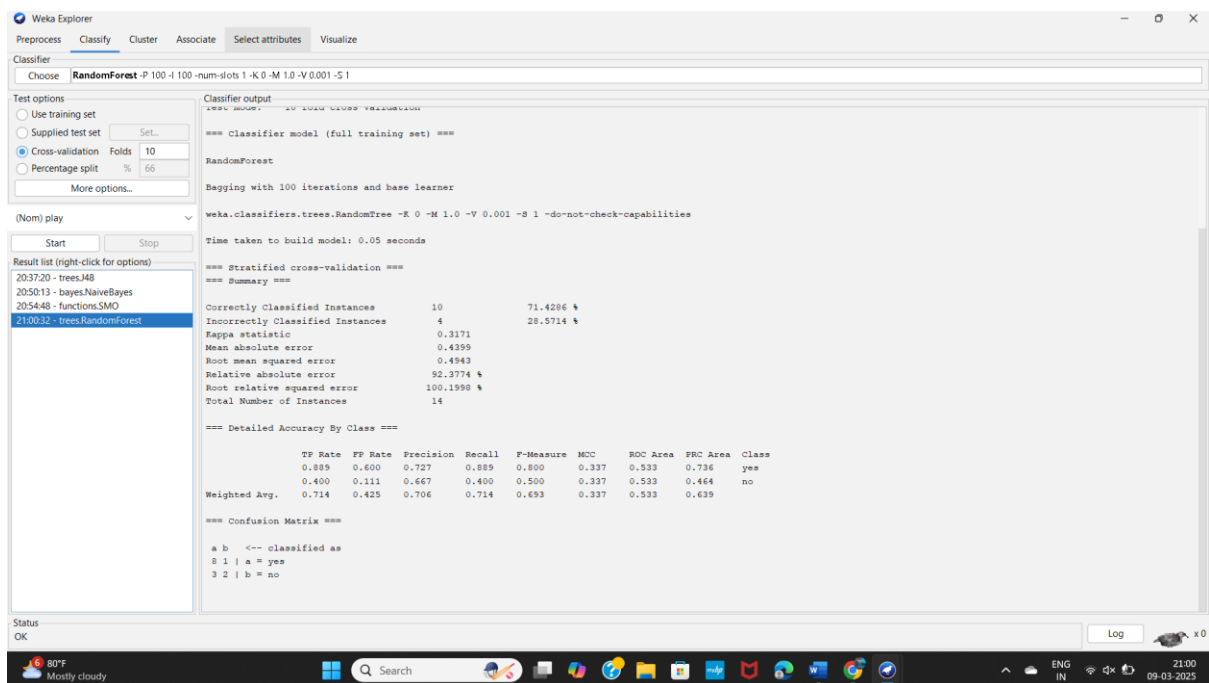
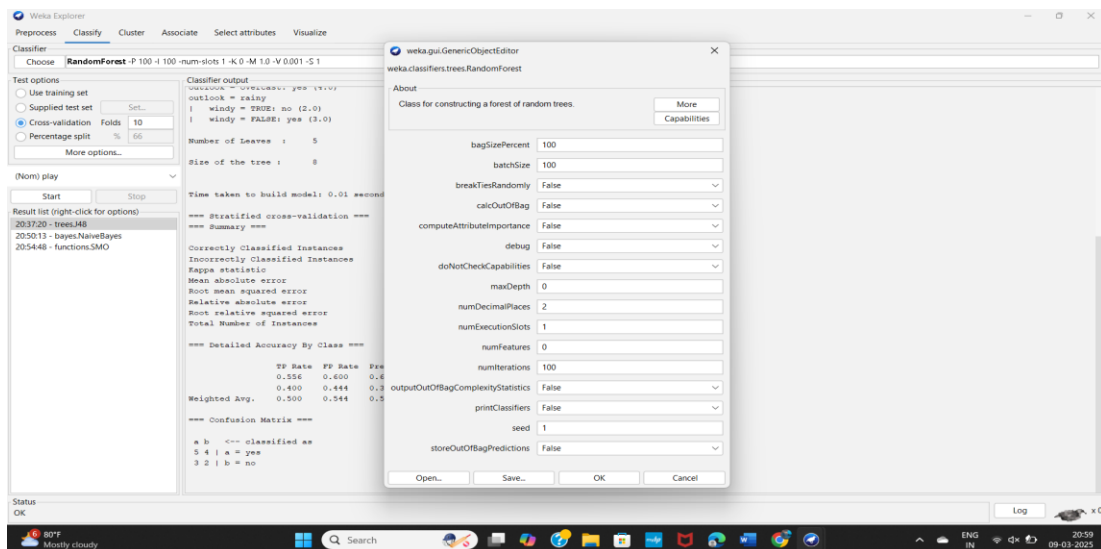
### === Confusion Matrix ===

a b <-- classified as

7 2 | a = yes

3 2 | b = no

## 4. Random Forest:



### === Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	
	0.889	0.600	0.727	0.889	0.800	0.337	0.533	0.736	yes
	0.400	0.111	0.667	0.400	0.500	0.337	0.533	0.464	no
Weighted Avg.	0.714	0.425	0.706	0.714	0.693	0.337	0.533	0.639	

### === Confusion Matrix ===

a b <-- classified as

8 1 | a = yes

3 2 | b = no

### RESULT:

Thus, the comparison of the confusion matrix for all the methods and techniques is performed. Out of the comparing matrix with all the techniques there is a change in instances. The above graphs will show the variations of values in the parameters.