Hindi Vidya Prachar Samiti's

RAMNIRANJAN JHUNJHUNWALA COLLEGE OF ARTS, SCIENCE & COMMERCE (AUTONOMOUS)

Python for Data Science Journal

Name : Shraddha Anand Dwivedi
Roll No: 733
Class : MSc Data Science and Artificial Intelligence part I

Ramniranjan Jhunjhunwala College

of Arts, Science and Commerce

Department of Data Science and Artificial Intelligence

CERTIFICATE

This is to certify Shraddha Dwivedi of Msc. Data Science and Artificial Intelligence Roll No 733 has successfully completed the practical of Artificial Intelligence for Data Science during the Academic Year 2023-2024.

Date :

(Prof. Sujata Kotian)
Prof-In-Charge

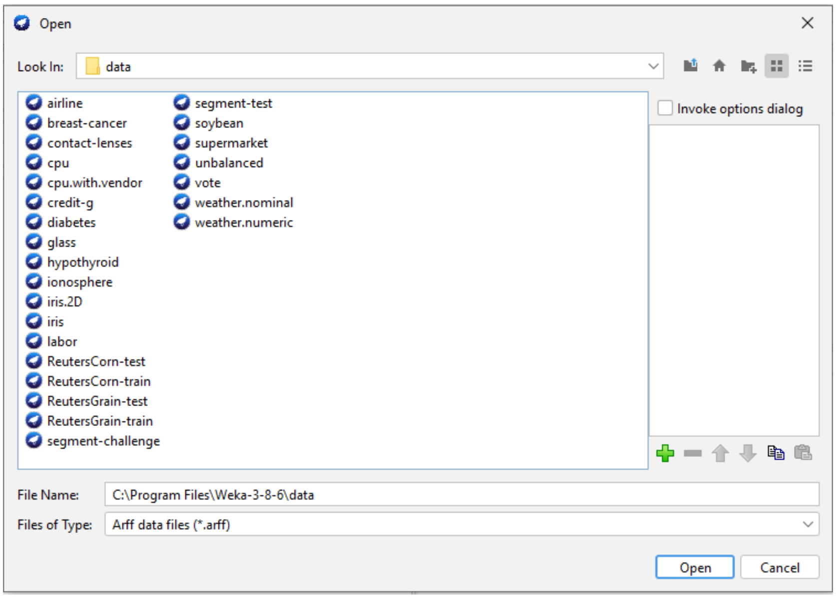
External Examiner

Practical 1 Supervised Learning using Weka tool

1) Decision tools

Open Weka tool

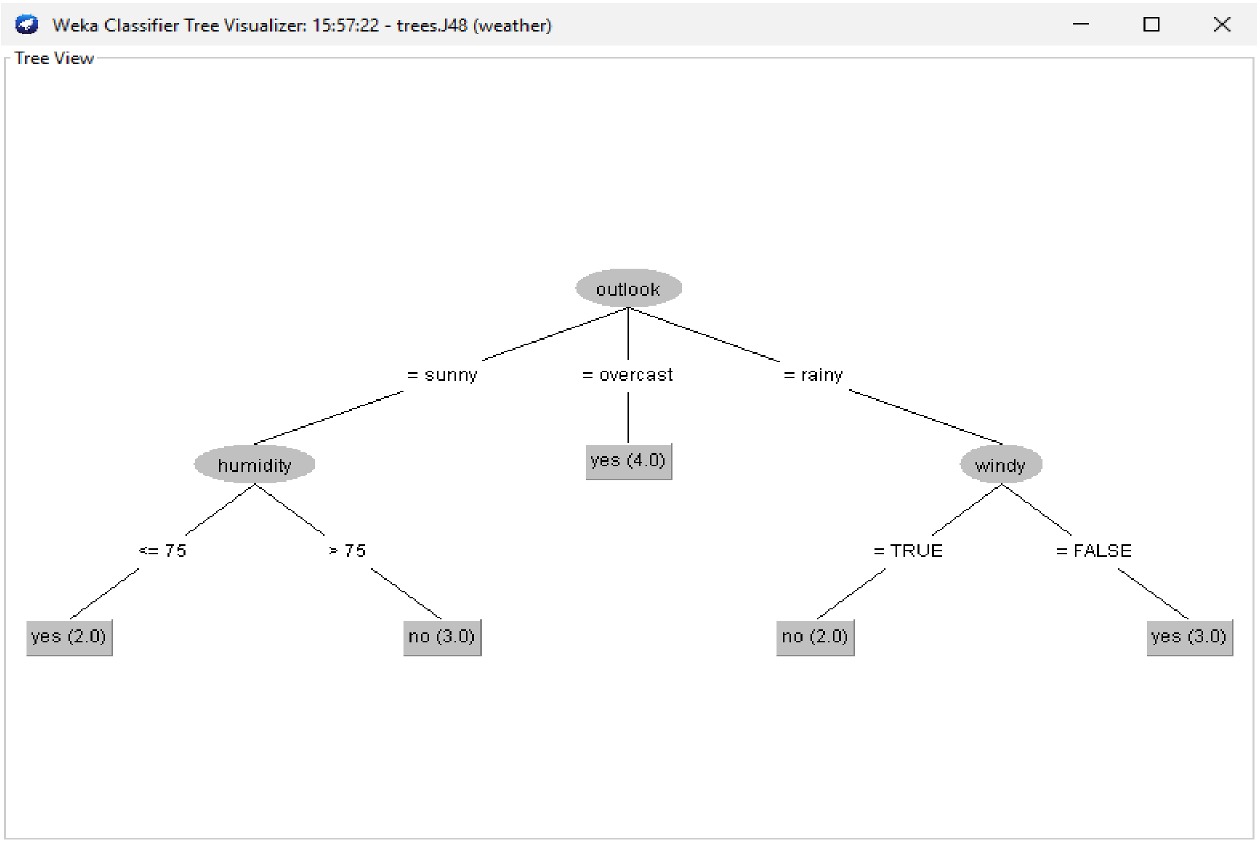
Select data set



Select classifier - J48 from trees

```
=== Summary ===

Correctly Classified Instances           9               64.2857 %
Incorrectly Classified Instances         5               35.7143 %
Kappa statistic                          0.186
Mean absolute error                      0.2857
Root mean squared error                  0.4818
Relative absolute error                 60      %
Root relative squared error             97.6586 %
Total Number of Instances               14

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.778    0.600    0.700      0.778    0.737      0.189    0.789     0.847     yes
                 0.400    0.222    0.500      0.400    0.444      0.189    0.789     0.738     no
Weighted Avg.    0.643    0.465    0.629      0.643    0.632      0.189    0.789     0.808

=== Confusion Matrix ===

 a b   <-- classified as
 7 2 | a = yes
 3 2 | b = no
```

Right click on trees.J48 and select ‘visualise tree’

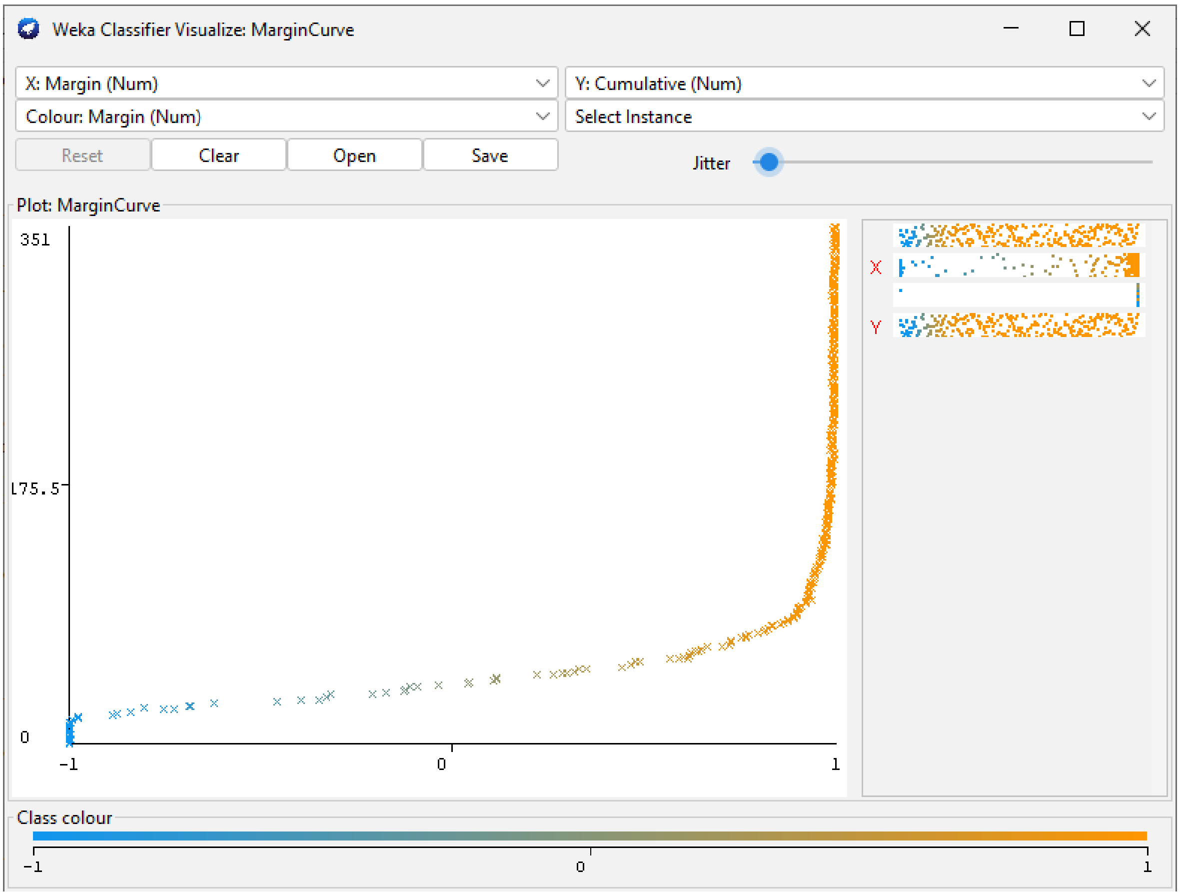


2) Logistics Regression

Select data ionosphere from preprocess
Select logistics from functions classifier

```
=== Summary ===

Correctly Classified Instances         312               88.8889 %
Incorrectly Classified Instances        39               11.1111 %
Kappa statistic                          0.753
Mean absolute error                      0.1283
Root mean squared error                  0.3035
Relative absolute error                 27.8593 %
Root relative squared error             63.2601 %
Total Number of Instances              351

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.794    0.058    0.885      0.794    0.837      0.756    0.870     0.896     b
                 0.942    0.206    0.891      0.942    0.916      0.756    0.870     0.832     g
Weighted Avg.    0.889    0.153    0.889      0.889    0.887      0.756    0.870     0.855

=== Confusion Matrix ===

   a   b   <-- classified as
 100  26 |   a = b
  13 212 |   b = g
```

Margin curve

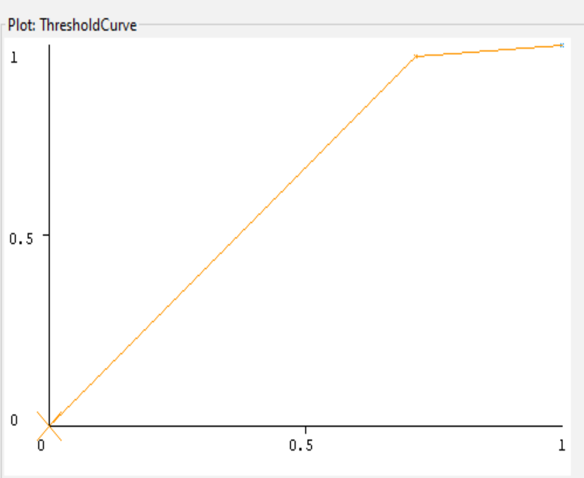


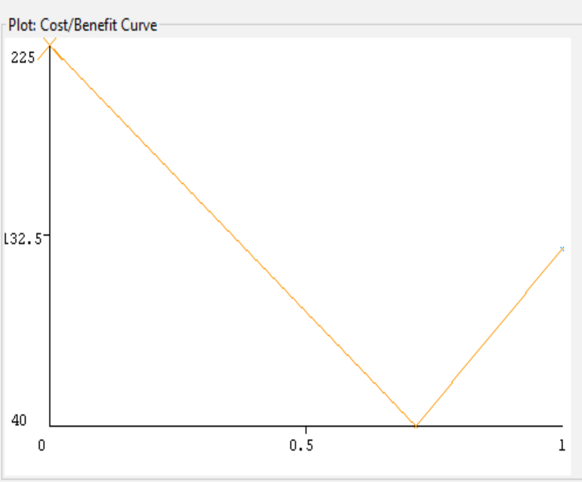
3) Support vector Machine

Select dataset ionosphere
Select SMO from functions classifier

```
=== Summary ===

Correctly Classified Instances         311               88.604  %
Incorrectly Classified Instances        40               11.396  %
Kappa statistic                          0.7406
Mean absolute error                      0.114
Root mean squared error                  0.3376
Relative absolute error                 24.7463 %
Root relative squared error             70.3666 %
Total Number of Instances              351

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.738    0.031    0.930      0.738    0.823      0.751    0.853     0.780     b
                 0.969    0.262    0.869      0.969    0.916      0.751    0.853     0.861     g
Weighted Avg.    0.886    0.179    0.891      0.886    0.883      0.751    0.853     0.832

=== Confusion Matrix ===

   a   b   <-- classified as
  93  33 |   a = b
   7 218 |   b = g
```

cost / benefit graph



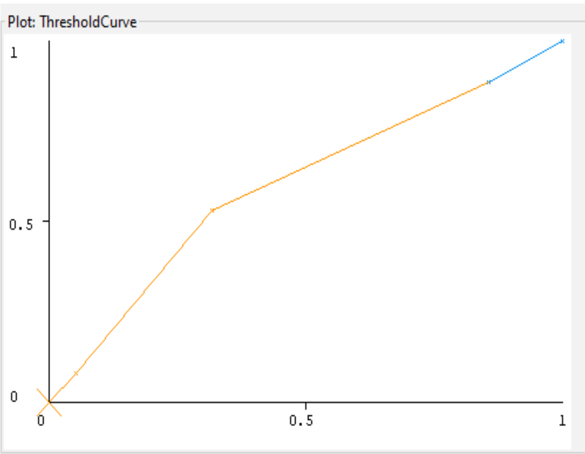


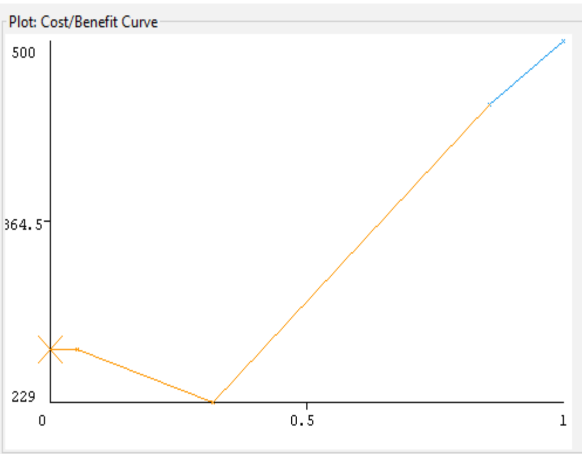
4) Linear Search

Select dataset diabetes
Select IBk from lazy classifier

```
=== Summary ===

Correctly Classified Instances         539               70.1823 %
Incorrectly Classified Instances       229               29.8177 %
Kappa statistic                          0.3304
Mean absolute error                      0.2988
Root mean squared error                  0.5453
Relative absolute error                 65.7327 %
Root relative squared error            114.3977 %
Total Number of Instances              768

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.794    0.470    0.759      0.794    0.776      0.331    0.650     0.732     tested_negative
                 0.530    0.206    0.580      0.530    0.554      0.331    0.650     0.469     tested_positive
Weighted Avg.    0.702    0.378    0.696      0.702    0.698      0.331    0.650     0.640

=== Confusion Matrix ===

   a   b   <-- classified as
 397 103 |   a = tested_negative
 126 142 |   b = tested_positive
```

Positive cost benefit curve





Signature: ____________

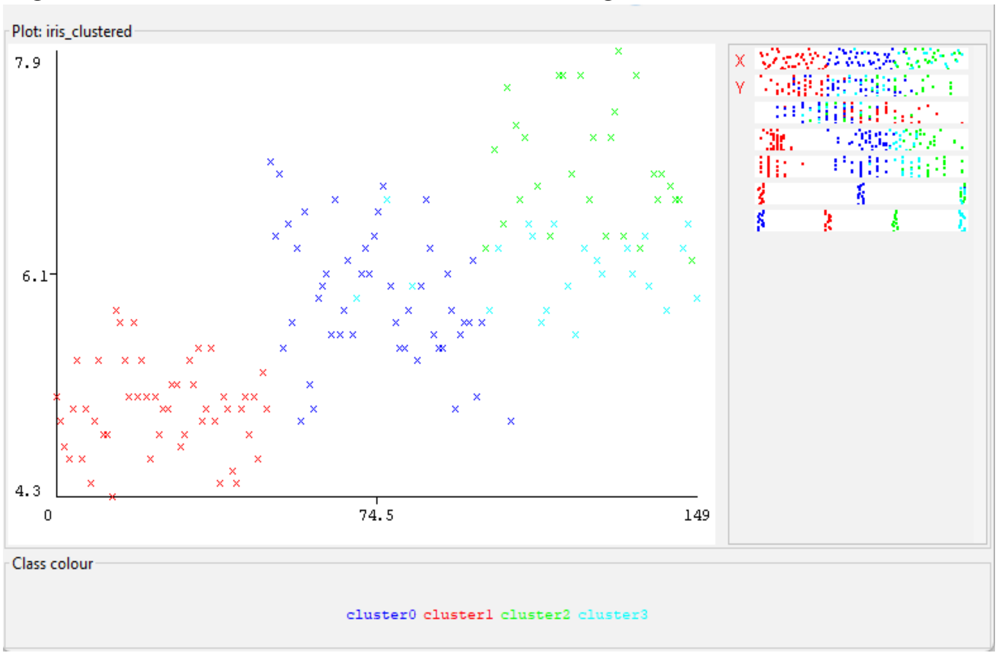
Practical 2 Unsupervised Learning using Weka tool

1)Clustering

Select dataset iris from preprocess
Select EM from cluster

```
                        Cluster
Attribute                     0        1        2        3
                         (0.32)   (0.33)    (0.2)   (0.14)
====================================================
sepallength
  mean                    5.897    5.006   6.9426   6.1304
  std. dev.              0.5279   0.3489    0.498   0.2943

sepalwidth
  mean                   2.7519    3.418   3.1103   2.8088
  std. dev.              0.3103   0.3772   0.2952   0.2361

petallength
  mean                   4.2267    1.464   5.8559   5.0993
  std. dev.               0.445   0.1718   0.4626   0.2462

petalwidth
  mean                   1.3134    0.244   2.1495   1.8254
  std. dev.              0.1864   0.1061    0.232   0.2152

class
  Iris-setosa                 1       51        1        1
  Iris-versicolor       48.1125        1   1.0182   3.8693
  Iris-virginica         2.0983        1  31.0375  19.8641
  [total]               51.2108       53  33.0557  24.7335


Time taken to build model (full training data) : 0.2 seconds

=== Model and evaluation on training set ===

Clustered Instances

0       48 ( 32%)
1       50 ( 33%)
2       29 ( 19%)
3       23 ( 15%)


Log likelihood: -2.03504
```

Right click and select visualize cluster assignments



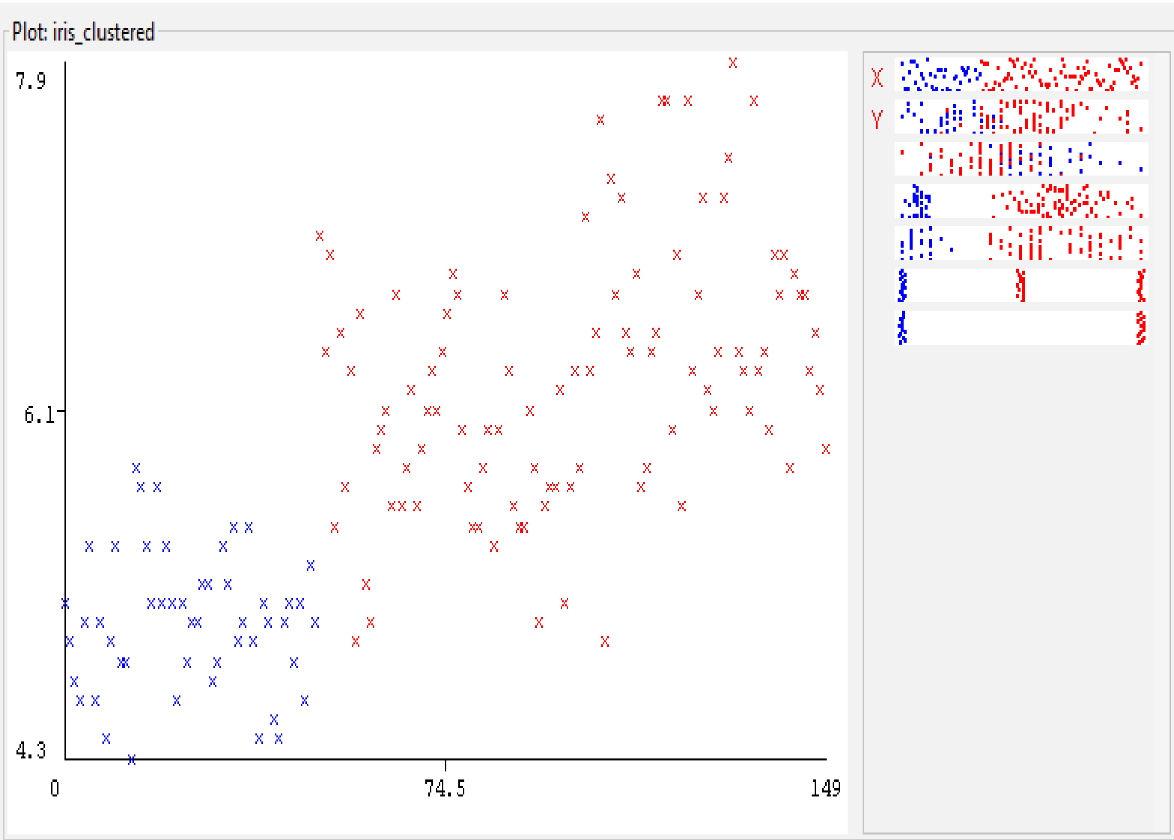
2)Hierarchical Clustering

Select iris dataset from preprocess
Select cluster>clusterer>Hirarchical clustering>start

Clusterer output

```
=== Run information ===

Scheme:       weka.clusterers.HierarchicalClusterer -N 2 -L SINGLE -P -A "weka.core.EuclideanDistance -R first-last"
Relation:     iris
Instances:    150
Attributes:   5
              sepallength
              sepalwidth
              petallength
              petalwidth
              class
Test mode:    evaluate on training data


=== Clustering model (full training set) ===

Cluster 0
((((((((((((((((((((0.0:0.03254,0.0:0.03254):0.00913,(0.0:0.03254,0.0:0.03254):0.00913):0.00332,((0.0:0.02778,0.0:0.02778):0.00476,0.0:0.03254):(

Cluster 1
(((((((((((((((((1.0:0.07344,(((1.0:0.06508,1.0:0.06508):0.00066,(1.0:0.05008,1.0:0.05008):0.01566):0.00224,1.0:0.06798):0.00546):0.00188,(1.0:0.



Time taken to build model (full training data) : 0.03 seconds

=== Model and evaluation on training set ===

Clustered Instances

0       50 ( 33%)
1      100 ( 67%)
```

Right click on 'Hirarchical clusterer' and select visualize cluster assignments

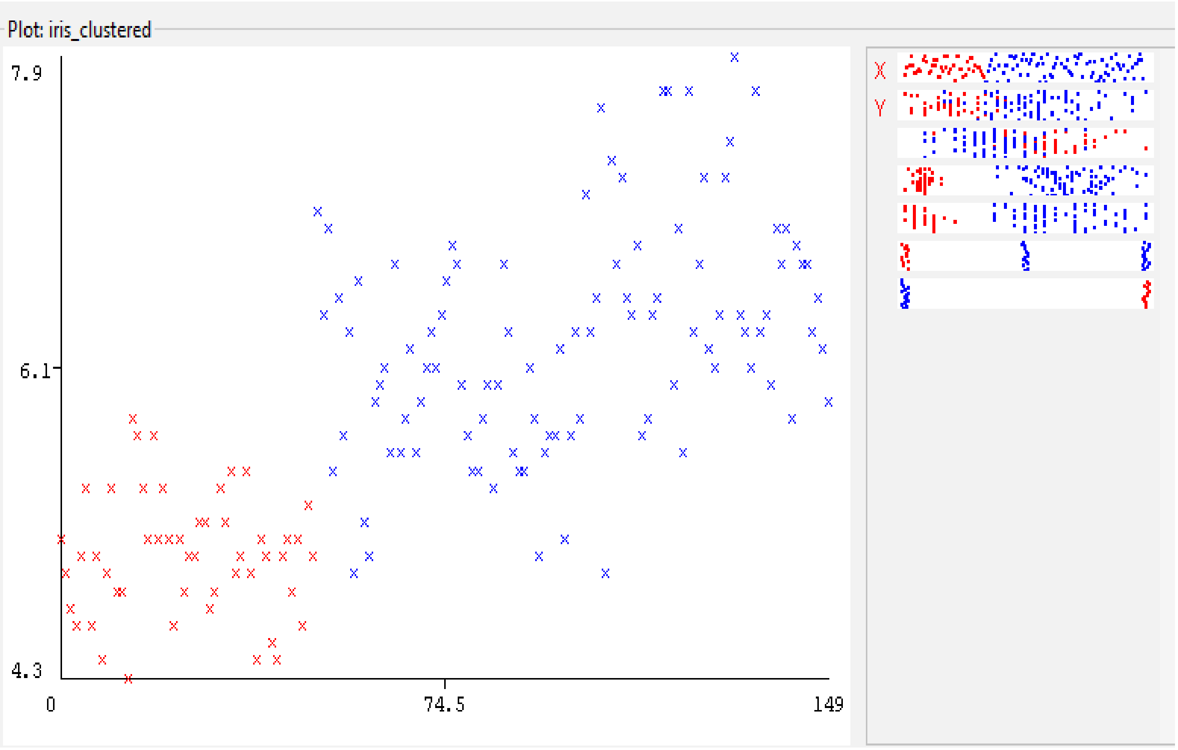


3)Density Based Clustering

Select iris dataset from preprocess
Select cluster>clusterer>MakeDensityBasedCluster>start

```
Attribute: sepallength
Normal Distribution. Mean = 6.262 StdDev = 0.6595
Attribute: sepalwidth
Normal Distribution. Mean = 2.872 StdDev = 0.3311
Attribute: petallength
Normal Distribution. Mean = 4.906 StdDev = 0.8214
Attribute: petalwidth
Normal Distribution. Mean = 1.676 StdDev = 0.4226
Attribute: class
Discrete Estimator. Counts =  1 51 51  (Total = 103)

Cluster: 1 Prior probability: 0.3355

Attribute: sepallength
Normal Distribution. Mean = 5.006 StdDev = 0.3489
Attribute: sepalwidth
Normal Distribution. Mean = 3.418 StdDev = 0.3772
Attribute: petallength
Normal Distribution. Mean = 1.464 StdDev = 0.1718
Attribute: petalwidth
Normal Distribution. Mean = 0.244 StdDev = 0.1061
Attribute: class
Discrete Estimator. Counts =  51 1 1  (Total = 53)


Time taken to build model (full training data) : 0 seconds

=== Model and evaluation on training set ===

Clustered Instances

0       100 ( 67%)
1        50 ( 33%)


Log likelihood: -3.06315
```

Right click on MakeDensityBasedCluster>Visualise cluster assignments

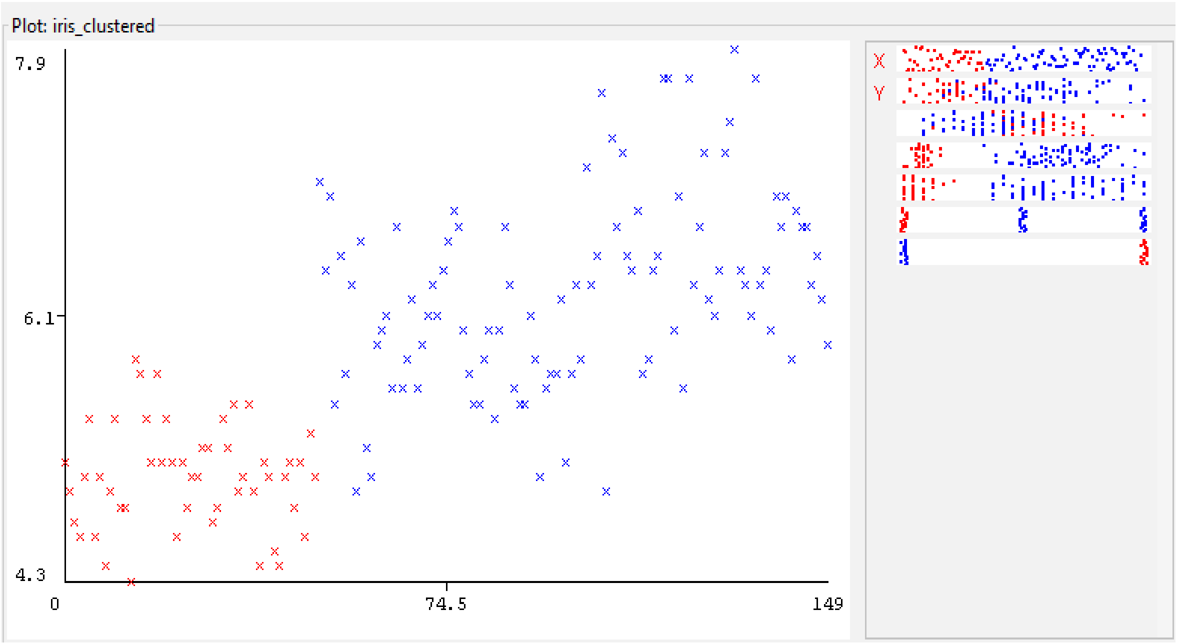


4) Simple K-means clustering

Select iris dataset from preprocess
Select cluster>clusterer>MakeDensityBasedCluster>start

```
kMeans
======

Number of iterations: 7
Within cluster sum of squared errors: 62.1436882815797

Initial starting points (random):

Cluster 0: 6.1,2.9,4.7,1.4,Iris-versicolor
Cluster 1: 6.2,2.9,4.3,1.3,Iris-versicolor

Missing values globally replaced with mean/mode

Final cluster centroids:
                                         Cluster#
Attribute                Full Data              0              1
                           (150.0)        (100.0)         (50.0)
==================================================================
sepallength                 5.8433          6.262          5.006
sepalwidth                   3.054          2.872          3.418
petallength                 3.7587          4.906          1.464
petalwidth                  1.1987          1.676          0.244
class                  Iris-setosa Iris-versicolor   Iris-setosa




Time taken to build model (full training data) : 0 seconds

=== Model and evaluation on training set ===

Clustered Instances

0      100 ( 67%)
1       50 ( 33%)
```

Right click on 'SimpleKmeans'>visualize cluster assignments



Signature: ____________

Practical 3 Association Algorithm

1) Association algorithm

Select iris dataset from preprocess
select> associate>associater>Apriori

```
Apriori
=======

Minimum support: 0.15 (694 instances)
Minimum metric <confidence>: 0.9
Number of cycles performed: 17

Generated sets of large itemsets:

Size of set of large itemsets L(1): 44

Size of set of large itemsets L(2): 380

Size of set of large itemsets L(3): 910

Size of set of large itemsets L(4): 633

Size of set of large itemsets L(5): 105

Size of set of large itemsets L(6): 1

Best rules found:

 1. biscuits=t frozen foods=t fruit=t total=high 788 ==> bread and cake=t 723    <conf:(0.92)> lift:(1.27) lev:(0.03
 2. baking needs=t biscuits=t fruit=t total=high 760 ==> bread and cake=t 696    <conf:(0.92)> lift:(1.27) lev:(0.03
 3. baking needs=t frozen foods=t fruit=t total=high 770 ==> bread and cake=t 705    <conf:(0.92)> lift:(1.27) lev:(
 4. biscuits=t fruit=t vegetables=t total=high 815 ==> bread and cake=t 746    <conf:(0.92)> lift:(1.27) lev:(0.03)
 5. party snack foods=t fruit=t total=high 854 ==> bread and cake=t 779    <conf:(0.91)> lift:(1.27) lev:(0.04) [164
 6. biscuits=t frozen foods=t vegetables=t total=high 797 ==> bread and cake=t 725    <conf:(0.91)> lift:(1.26) lev:
 7. baking needs=t biscuits=t vegetables=t total=high 772 ==> bread and cake=t 701    <conf:(0.91)> lift:(1.26) lev:
 8. biscuits=t fruit=t total=high 954 ==> bread and cake=t 866    <conf:(0.91)> lift:(1.26) lev:(0.04) [179] conv:(3
 9. frozen foods=t fruit=t vegetables=t total=high 834 ==> bread and cake=t 757    <conf:(0.91)> lift:(1.26) lev:(0.
10. frozen foods=t fruit=t total=high 969 ==> bread and cake=t 877    <conf:(0.91)> lift:(1.26) lev:(0.04) [179] con
```

2)Feature Extraction

Select soybean dataset from preprocess
Select select attributes>Attribute Evaluator>CfsSubsetEval
>Search Method>BestFirst

```
=== Attribute Selection on all input data ===

Search Method:
	Best first.
	Start set: no attributes
	Search direction: forward
	Stale search after 5 node expansions
	Total number of subsets evaluated: 552
	Merit of best subset found:    0.702

Attribute Subset Evaluator (supervised, Class (nominal): 36 class):
	CFS Subset Evaluator
	Including locally predictive attributes

Selected attributes: 1,3,4,5,7,8,9,10,11,12,13,15,17,18,19,22,23,24,26,28,30,35 : 22
                     date
                     precip
                     temp
                     hail
                     area-damaged
                     severity
                     seed-tmt
                     germination
                     plant-growth
                     leaves
                     leafspots-halo
                     leafspot-size
                     leaf-malf
                     leaf-mild
                     stem
                     canker-lesion
                     fruiting-bodies
                     external-decay
                     int-discolor
                     fruit-pods
                     seed
```

3)Name based Classifier/ Bayesian Classifier

Select diabetes dataset from preprocess
Select classify>bayes>NaiveBayes>Start

```
=== Summary ===

Correctly Classified Instances         586               76.3021 %
Incorrectly Classified Instances       182               23.6979 %
Kappa statistic                          0.4664
Mean absolute error                      0.2841
Root mean squared error                  0.4168
Relative absolute error                 62.5028 %
Root relative squared error             87.4349 %
Total Number of Instances              768

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.844    0.388    0.802      0.844    0.823      0.468    0.819     0.892     tested_negative
                 0.612    0.156    0.678      0.612    0.643      0.468    0.819     0.671     tested_positive
Weighted Avg.    0.763    0.307    0.759      0.763    0.760      0.468    0.819     0.815

=== Confusion Matrix ===

   a   b   <-- classified as
 422  78 |   a = tested_negative
 104 164 |   b = tested_positive
```

Signature: _____________

Practical 4

Q.1.a) Write a prolog program to implement simple facts in Queries

```
ram likes mangoes.
seema is a girl.
bill likes cindy.
rose is red.
john owns gold.

%clause

likes(ram,mangoes).
girl(seema).
red(rose).
likes(bill,cindy).
owns(john,gold).

 ?- lkes(ram,What).
 Correct to: "likes(ram,What)"?
 Please answer 'y' or 'n'? yes
 What = mangoes.

 ?- likes(ram,What).
 What = mangoes.

 ?- red(What).
 What = rose.

 ?- owns(Who,What).
 Who = john,
 What = gold.

 ?- likes(bill,Whom).
 Whom = cindy.

 ?- likes(Who,Whom).
 Who = ram,
 Whom = mangoes ;
 Who = bill,
 Whom = cindy.



#FOOD
%facts

food(burger).
food(sandwich).
food(pizza).
lunch(sandwich).
dinner(pizza).

%Rules

meal(X):- food(X).
```

```
.
% c:/Users/User38/Desktop/ShraddhaD/food.pl compiled 0.00 sec, 6 clauses
?- meal(pizza).
true.

?- meal(X).
X = burger ;
X = sandwich ;
X = pizza.

?- food(sandwich).
true.

?- lunch(pizza).
false.

?- dinner(burger).
false.
```

c) Student_Teacher
%facts

studies(charlie, csc135).
studies(olivia,csc135).
studies(jack,csc131).
studies(arthur,csc134).

teaches(kirke,csc135).
teaches(collins,csc131).
teaches(collins,csc171).
teaches(juniper,csc134).

%rules

professor(X,Y):- teaches(X,C),studies(Y,C).

```
.
% c:/Users/User38/Desktop/ShraddhaD/studies.pl compiled 0.00 sec, 9 clauses
?- studies(charlie,What).
What = csc135.

?- professor(kirke,Students).
Students = charlie ;
Students = olivia.

?- professor(Teaches,arthur).
Teaches = juniper.
```

d)
%facts

owns(jack,car(bmw)).
owns(john,car(chevy)).
owns(olivia,car(civic)).
owns(jane,car(chevy)).

sedan(car(bmw)).
sedan(car(civic)).
truck(car(chevy)).

```
?- owns(john,X).
X = car(chevy).

?- owns(Who,car(bmw)).
Who = jack.

?- owns(jane,X),truck(X).
X = car(chevy).

?- owns(olivia,_).
true.
```

e) pet relationships
%facts

cat(fubby).
black_spots(fubby).
dog(figaro).
white_spots(figaro).

%rules
owns(mary,pet):- cat(pet),black_spots(pet).
loves(who,what):- owns(who,what).

```
?-
% c:/Users/User38/Desktop/ShraddhaD/pet relations.pl compiled 0.00 sec, 0 clauses
?- owns(mary,_).
true.

?- listing(cat).
cat(fubby).

true.

?- listing(owns).
owns(mary, Pet) :-
    cat(Pet),
    black_spots(Pet).

true.
```

Signature: ____________

Practical 5 Write a prolog program for maximum and minimum

```
max(X,Y):-
(
X=Y->
write('both are equal')
;
X>Y->
(
Z is X,
write(Z)
)
;
(
Z is Y,
write(Z)
)
).
```

```
:-
% c:/Users/User38/Desktop/ShraddhaD/max(X,Y).pl compiled 0.00 sec, 1 clauses
?- max(5,7).
7
true.

?- max(9,9).
both are equal
true.

?- max(99,7).
99
true.
```

```
min(X,Y):-
(
X=Y->
write('both are equal')
;
X<Y->
(
Z is Y,
write(Z)
)
;
(
Z is X,
write(Z)
)
).
```

```
?- min(5,7).
7
true.

?- min(6,6).
both are equal
true.

?- min(99,7).
99
true.
```

b) prolog program for arithmetic Expressions

```
?- X is 19+9.
X = 28.

?- X is 361/9.
X = 40.111111111111114.

?- X is -(16,6).
X = 10.

?- X is -(16,-6).
X = 22.

?- X is 5 mod 3.
X = 2.

?- X is 25 mod 4.
X = 1.

?- X is 4 mod 25.
X = 4.

?- X is -(-(5,3),6).
X = -4.
```

Signature: _____________

Practical 6

a)
sum(X,Y,Z):- Z is X+Y.

```
?- sum(9,6,Z).
Z = 15.

?- sum(74,16,Z).
Z = 90.
```

b)
max(X,Y,M):-X>Y,M is X.
max(X,Y,M):-Y>=X,M is Y.

```
?- max(3,6,M).
M = 6.

?- max(23,6,M).
M = 23 .

?- max(16,16,M).
M = 16.
```

c)
multi(N1,N2,R):-R is N1*N2.

```
?- multi(2,6,R).
R = 12.

?- multi(200,16,R).
R = 3200.
```

d)Power function
power(Num,Pow,Ans):-Ans is Num^Pow.

```
?- power(5,2,Ans).
Ans = 25.

?- power(39,6,Ans).
Ans = 3518743761.
```

Signature: ____________

Practical 7 prolog program to calculate user input

a)Program in prolog to find cube of a number

```
cube:-
write('Write a number:'),
read(Number),
process(Number).

process(stop):-!.
process(Number):-
C is Number*Number*Number,
write('Cube of'),write(Number),write(': '),write(C),nl,cube.
```

```
% c:/Users/User38/Desktop/ShraddhaD/to find cube of number.pl compiled 0.00 sec, 3 clauses
?- cube.
Write a number:9.
Cube of9: 729
Write a number:|: 16.
Cube of16: 4096
Write a number:|: 3.
Cube of3: 27
Write a number:|: 19.
Cube of19: 6859
Write a number:|: 678.
Cube of678: 311665752
Write a number:|:
```

b)prolog program to implement addition multiplication by taking 2 numbers from user

```
multiply:-
write('Write a number X:'),read(X),nl,
write('Write a number Y:'),read(Y),nl,
addmul(X,Y).

addmul(X,Y):-
S is X+Y,
M is X*Y,
write('Addition is: '),write(S),nl,
write('Multiplication is: '),write(M).
```

```
?- start.
Write a number X:9.

Write a number Y:|: 3.

Addition is: 12
Multiplication is: 27
true.
```

Signature: ____________

Practical 8 Family Relationship

```
female(pam).
female(liz).
female(pat).
female(ann).
male(jim).
male(bob).
male(tom).
male(peter).
parent(pam,bob).
parent(tom,bob).
parent(tom,liz).
parent(bob,ann).
parent(bob,pat).
parent(pat,jim).
parent(bob,peter).
parent(peter,jim).
mother(X,Y):-parent(X,Y),female(X).
father(X,Y):-parent(X,Y).male(X).
haschild(X):-parent(X,_).
sister(X,Y):-parent(Z,X),parent(Z,Y),female(X),X\==Y.
brother(X,Y):-parent(Z,X),parent(Z,Y),male(X),X\==Y.
```

```
?- parent(X,jim).
X = pat ;
X = peter.

?- mother(X,Y).
X = pam,
Y = bob ;
X = pat,
Y = jim ;
false.

?- sister(X,Y).
X = liz,
Y = bob ;
X = ann,
Y = pat ;
X = ann,
Y = peter ;
X = pat,
Y = ann ;
X = pat,
Y = peter ;
false.

?- haschild(X).
X = pam ;
X = tom ;
X = tom ;
X = bob ;
X = bob ;
X = pat ;
X = bob ;
X = peter.

?- brother(X,Y).
X = bob,
Y = liz ;
X = peter,
Y = ann ;
X = peter,
Y = pat
```

Signature: ____________

Practical 9 Tower of Hanoi

Move top disk from source to target
Move top disk from auxiliary to source
Move top disk from auxiliary to target
Move top disk from source to target

```
move(1,X,Y,_):-
write('Move top disk from '),write(X),write('to'),write('Y'),nl.
move(N,X,Y,Z):-
N>1,
M is N-1,
move(M,X,Y,Z),
move(1,X,Y,_),
move(M,Z,Y,X).
```

```
% c:/Users/User38/Desktop/ShraddhaD/Move top disk from source to target.pl
?- move(4,source,target,auxiliary).
Move top disk from sourcetoY
Move top disk from sourcetoY
Move top disk from auxiliarytoY
Move top disk from sourcetoY
Move top disk from auxiliarytoY
Move top disk from auxiliarytoY
Move top disk from sourcetoY
Move top disk from sourcetoY
Move top disk from auxiliarytoY
Move top disk from auxiliarytoY
Move top disk from sourcetoY
Move top disk from auxiliarytoY
Move top disk from sourcetoY
Move top disk from sourcetoY
Move top disk from auxiliarytoY
true
```

Signature: ____________