**5. Write a program to implement dbscan algorithm**

**OUTPUT:**

=== Run information ===

Scheme: weka.clusterers.MakeDensityBasedClusterer -M 1.0E-6 -W weka.clusterers.SimpleKMeans -- -init 0 -max-candidates 100 -periodic-pruning 10000 -min-density 2.0 -t1 -1.25 -t2 -1.0 -N 2 -A "weka.core.EuclideanDistance -R first-last" -I 500 -num-slots 1 -S 10

Relation: breast-cancer

Instances: 286

Attributes: 10

age

menopause

tumor-size

inv-nodes

node-caps

deg-malig

breast

breast-quad

irradiat

Class

Test mode: evaluate on training data

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

MakeDensityBasedClusterer:

Wrapped clusterer:

kMeans

======

Number of iterations: 3

Within cluster sum of squared errors: 1177.0

Initial starting points (random):

Cluster 0: 50-59,premeno,10-14,0-2,no,2,right,left\_up,no,no-recurrence-events

Cluster 1: 40-49,premeno,15-19,0-2,yes,3,right,left\_up,no,recurrence-events

Missing values globally replaced with mean/mode

Final cluster centroids:

Cluster#

Attribute Full Data 0 1

(286.0) (225.0) (61.0)

======================================================================================

age 50-59 50-59 40-49

menopause premeno premeno premeno

tumor-size 30-34 25-29 30-34

inv-nodes 0-2 0-2 0-2

node-caps no no yes

deg-malig 2 2 3

breast left left left

breast-quad left\_low left\_low left\_low

irradiat no no no

Class no-recurrence-events no-recurrence-events recurrence-events

Fitted estimators (with ML estimates of variance):

Cluster: 0 Prior probability: 0.7847

Attribute: age

Discrete Estimator. Counts = 1 2 27 65 84 46 7 1 1 (Total = 234)

Attribute: menopause

Discrete Estimator. Counts = 8 106 114 (Total = 228)

Attribute: tumor-size

Discrete Estimator. Counts = 9 5 28 28 43 48 34 12 18 3 8 1 (Total = 237)

Attribute: inv-nodes

Discrete Estimator. Counts = 190 19 8 8 2 4 1 1 1 1 1 1 1 (Total = 238)

Attribute: node-caps

Discrete Estimator. Counts = 21 206 (Total = 227)

Attribute: deg-malig

Discrete Estimator. Counts = 68 122 38 (Total = 228)

Attribute: breast

Discrete Estimator. Counts = 118 109 (Total = 227)

Attribute: breast-quad

Discrete Estimator. Counts = 80 87 24 20 19 (Total = 230)

Attribute: irradiat

Discrete Estimator. Counts = 44 183 (Total = 227)

Attribute: Class

Discrete Estimator. Counts = 194 33 (Total = 227)

Cluster: 1 Prior probability: 0.2153

Attribute: age

Discrete Estimator. Counts = 1 1 11 27 14 13 1 1 1 (Total = 70)

Attribute: menopause

Discrete Estimator. Counts = 1 25 38 (Total = 64)

Attribute: tumor-size

Discrete Estimator. Counts = 1 1 2 4 9 8 28 9 6 2 2 1 (Total = 73)

Attribute: inv-nodes

Discrete Estimator. Counts = 25 19 11 4 3 4 1 1 2 1 1 1 1 (Total = 74)

Attribute: node-caps

Discrete Estimator. Counts = 37 26 (Total = 63)

Attribute: deg-malig

Discrete Estimator. Counts = 5 10 49 (Total = 64)

Attribute: breast

Discrete Estimator. Counts = 36 27 (Total = 63)

Attribute: breast-quad

Discrete Estimator. Counts = 19 26 11 6 4 (Total = 66)

Attribute: irradiat

Discrete Estimator. Counts = 26 37 (Total = 63)

Attribute: Class

Discrete Estimator. Counts = 9 54 (Total = 63)

Time taken to build model (full training data) : 0.01 seconds

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

Clustered Instances

0 226 ( 79%)

1 60 ( 21%)

Log likelihood: -9.79929