3.Write a program to implement k-means clustering algorithm

OUTPUT:

=== Run information ===

Scheme: 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: hypothyroid

Instances: 3772

Attributes: 30

age

sex

on thyroxine

query on thyroxine

on antithyroid medication

sick

pregnant

thyroid surgery

I131 treatment

query hypothyroid

query hyperthyroid

lithium

goitre

tumor

hypopituitary

psych

TSH measured

TSH

T3 measured

T3

TT4 measured

TT4

T4U measured

T4U

FTI measured

FTI

TBG measured

TBG

referral source

Class

Test mode: evaluate on training data

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

kMeans

======

Number of iterations: 14

Within cluster sum of squared errors: 6902.201759447764

Initial starting points (random):

Cluster 0: 69,F,t,f,f,f,f,f,f,f,f,f,f,f,f,f,t,1.5,t,1.8,t,136,t,0.92,t,149,f,0,other,negative

Cluster 1: 27,F,f,f,f,f,f,f,f,f,f,f,f,f,f,f,t,0.15,t,1.6,t,101,t,1.09,t,92,f,0,SVHC,negative

Missing values globally replaced with mean/mode

Final cluster centroids:

Cluster#

Attribute Full Data 0 1

(3772.0) (892.0) (2880.0)

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

age 51.7359 48.8587 52.627

sex F F F

on thyroxine f f f

query on thyroxine f f f

on antithyroid medication f f f

sick f f f

pregnant f f f

thyroid surgery f f f

I131 treatment f f f

query hypothyroid f f f

query hyperthyroid f f f

lithium f f f

goitre f f f

tumor f f f

hypopituitary f f f

psych f f f

TSH measured t t t

TSH 5.0868 1.7273 6.1273

T3 measured t t t

T3 2.0135 2.5923 1.8342

TT4 measured t t t

TT4 108.3193 151.9865 94.7946

T4U measured t t t

T4U 0.995 1.1534 0.9459

FTI measured t t t

FTI 110.4696 136.134 102.5208

TBG measured f f f

TBG 0 0 0

referral source other other other

Class negative negative negative

Time taken to build model (full training data) : 0.14 seconds

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

Clustered Instances

0 892 ( 24%)

1 2880 ( 76%)