

# 1.K MEANS ALGORITHM:

DATA SET: diabetes

OUTPUT:

The first screenshot shows the Weka Explorer interface with the SimpleKMeans algorithm selected. The cluster mode is set to 'Percentage split' with a percentage of 66. The clusterer output shows the initial starting points (random) and the first iteration of the algorithm. The output includes the number of iterations (4), the within cluster sum of squared errors (149.5177664581119), and the initial starting points (random). The output also shows the final cluster centroids for two clusters: Cluster 0 (1,126,56,29,152,28.7,0.801,21,tested\_negative) and Cluster 1 (8,95,72,0,0,36.8,0.485,57,tested\_negative).

The second screenshot shows the Weka Explorer interface with the SimpleKMeans algorithm selected. The cluster mode is set to 'Percentage split' with a percentage of 66. The clusterer output shows the final cluster centroids and the evaluation metrics. The output includes the number of iterations (2), the within cluster sum of squared errors (115.21146310981545), and the final cluster centroids for two clusters: Cluster 0 (10,148,84,48,237,37.6,1.001,51,tested\_positive) and Cluster 1 (6,154,74,32,193,29.3,0.839,39,tested\_negative). The output also shows the missing values globally replaced with mean/mode and the final cluster centroids.

Weka Explorer

PreprocessClassifyClusterAssociateSelect attributesVisualize

Clusterer

ChooseSimpleKMeans-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

Cluster mode

Use training set

Supplied test set

Percentage split

%66

Classes to clusters evaluation

(Nom) class

Store clusters for visualization

Ignore attributes

Start

Stop

Result list (right-click for options)

193450 - SimpleKMeans

193542 - SimpleKMeans

Clusterer output

Number of iterations: 2

Within cluster sum of squared errors: 115.21146310581545

Initial starting points (random):

Cluster 0: 10,148,84,48,237,37.6,1.001,51,tested\_positive

Cluster 1: 6,154,74,32,193,29.3,0.839,39,tested\_negative

Missing values globally replaced with mean/mode

Final cluster centroids:

Attribute	Full Data	Cluster#	0	1
	(506.0)		(184.0)	(322.0)
preg	3.919	4.7026	3.4255	
plas	121.164	139.3967	110.7453	
pres	69.1779	70.8587	68.2174	
skin	19.9486	21.962	18.7981	
insu	78.585	90.0217	72.0497	
mass	32.0275	35.2766	30.1708	
pedi	0.4798	0.5516	0.4388	
age	33.7925	37.4674	31.6925	
class		tested_negative	tested_positive	tested_negative

Time taken to build model (percentage split) : 0.01 seconds

Clustered Instances

0	84 ( 32%)
1	178 ( 68%)

Status

OK

OneDrive - Personal Processing changes

Log

x 0

27°C

Partly cloudy

Search

ENG IN

19:36

12-02-2023