**Practical No: 8**

**Clustering algorithm**

**AIM: Write an application to implement clustering algorithm.**

**Code: Hierarchical clustering**

import matplotlib.pyplot as plt

import pandas as pd

import numpy as np

import scipy.cluster.hierarchy as shc

from sklearn.cluster import AgglomerativeClustering

# Read the customer data from a CSV file

customer\_data = pd.read\_csv('/content/mall\_customers.csv')

# Display the shape and the first few rows of the data

print(customer\_data.shape)

customer\_data.head()

# Extract the relevant columns from the data

data = customer\_data.iloc[:, 3:5].values

# Create a dendrogram plot

plt.figure(figsize=(10, 7))

plt.title("Customer Dendrograms")

dend = shc.dendrogram(shc.linkage(data, method='ward'))

# Perform hierarchical clustering

cluster = AgglomerativeClustering(n\_clusters=5, affinity='euclidean', linkage='ward')

cluster\_labels = cluster.fit\_predict(data)

# Create a scatter plot to visualize the clusters

plt.figure(figsize=(10, 7))

plt.scatter(data[:, 0], data[:, 1], c=cluster\_labels, cmap='rainbow')

plt.show()

OUTPUT:

A screenshot of a screen shot

Description automatically generated

A diagram of a city

Description automatically generated

A chart of different colored dots

Description automatically generated

**b) Clustering**

**Code:**

from numpy import where

from sklearn.datasets import make\_classification

from matplotlib import pyplot

x,y = make\_classification(n\_samples=1000,n\_features=2,n\_informative=2,n\_redundant=0,n\_clusters\_per\_class=1,random\_state=4)

for class\_value in range(2):

row\_ix=where(y==class\_value)

pyplot.scatter(x[row\_ix,0],x[row\_ix,1])

pyplot.show()

OUTPUT:

A blue and orange dotted line

Description automatically generated