

**Discrete Structures**

Project Review

Submitted by:

Abhishek Singh (2K19/CO/024) Abhishek Kumar Singh (2K19/CO/021)

**AIM**

To partition a given set of objects into subsets (or clusters) in such a way that any two objects from the same subset are close (or similar) to each other, while any two objects from different subsets are far apart.

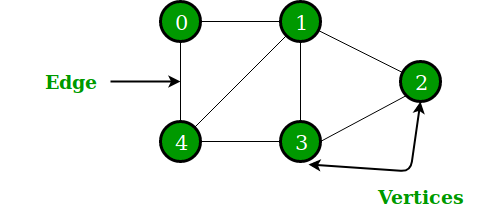
**MOTIVATION**

Clustering algorithms are used for disease classification in medical science and also for customer classification in marketing research and for environmental health risk assessment in environmental engineering. It is a main task of exploratory data mining, and a common technique for statistical data analysis, used in many fields, including pattern recognition, image analysis, information retrieval, bioinformatics, data compression, computer graphics and machine learning.

**PREREQUISITE**

1. **Graph**

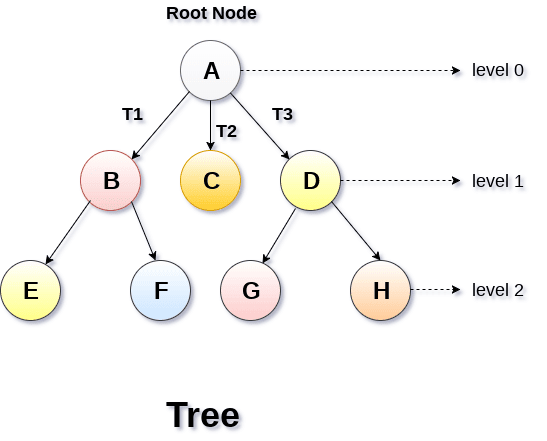
A Graph is a non-linear data structure consisting of nodes and edges. The nodes are sometimes also referred to as vertices and the edges are lines or arcs that connect any two nodes in the graph.

**

1. **Trees**

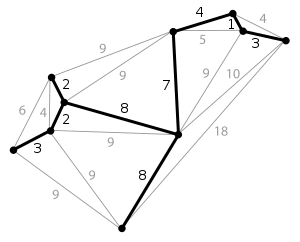
A Tree is a recursive data structure containing the set of one or more data nodes where one node is designated as the root of the tree while the remaining nodes are called as the children of the root. Nodes of a tree either maintain a parent-child relationship between them or they are sister nodes.

The following image shows a tree, where the node A is the root node of the tree while the other nodes can be seen as the children of A.



1. **Minimum Spanning Tree**

A minimum spanning tree (MST) or minimum weight spanning tree is a subset of the edges of a connected, edge-weighted undirected graph that connects all the vertices together, without any cycles and with the minimum possible total edge weight. That is, it is a spanning tree whose sum of edge weights is as small as possible. More generally, any edge-weighted undirected graph (not necessarily connected) has a minimum spanning forest, which is a union of the minimum spanning trees for its connected components.

.

1. **C++**

C++ is a general-purpose programming language created by Bjarne Stroustrup as an extension of the C programming language, or "C with Classes. The language has expanded significantly over time, and modern C++ now has object-oriented, generic, and functional features in addition to facilities for low-level memory manipulation.

**APPROACH**

We will devise an algorithm to compute the largest possible value of 𝑑 such that the given points can be partitioned into 𝑘 non-empty subsets in such a way that the distance between any two points from different subsets is at least 𝑑.