A

Practical File

On

Data ware house and data mining

Paper Code: ITD08



Bachelor of Engineering

(2016-2020)

Computer Science Engineering

**Submitted By:**

Kavita Maurya

Batch-3

Group-1

2016UCO1579

**DIVISION OF COMPUTER SCIENCE ENGINEERING**

**NETAJI SUBHAS UNIVERSITY OF TECHNOLOGY, DWARKA**

**1.**

**AIM:**

Write a procedure for Clustering Weather data using EM Algorithm.

**DESCRIPTION:**

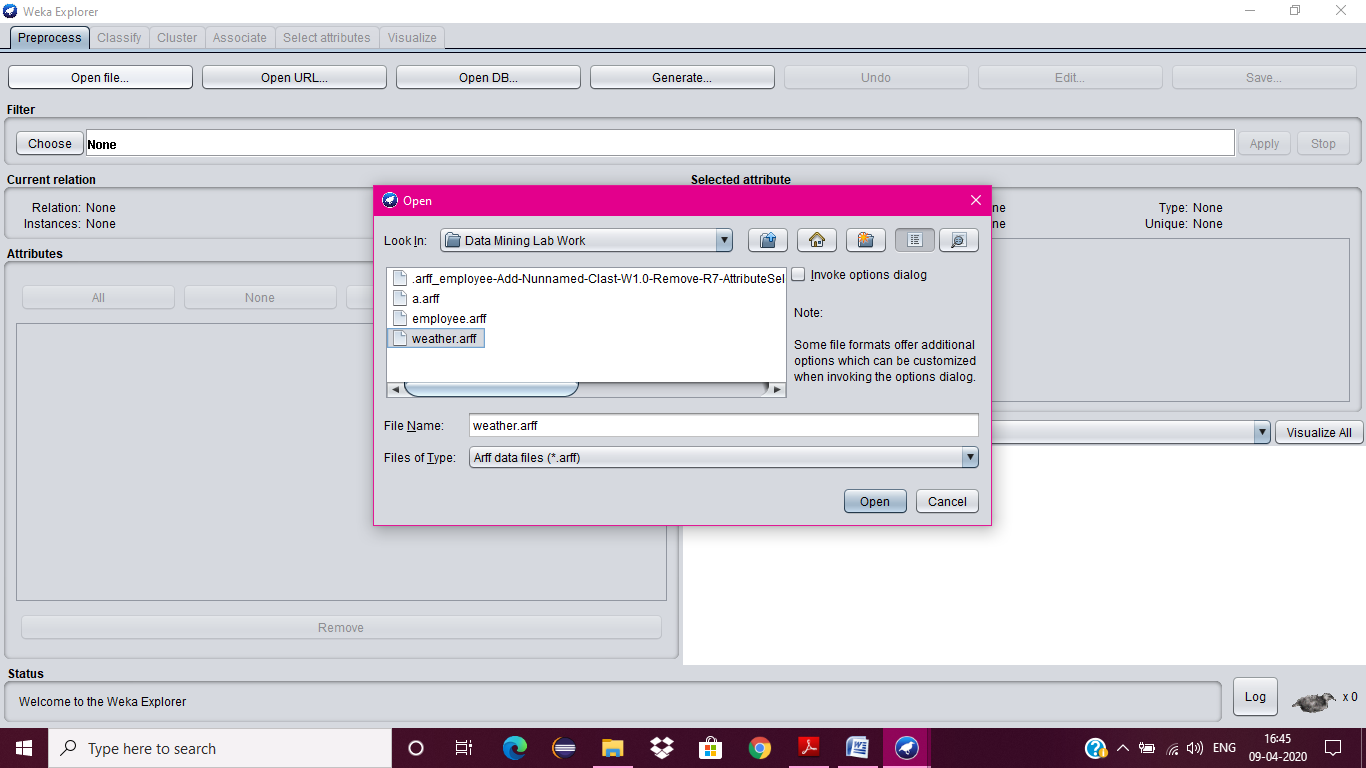
Cluster analysis or clustering is the task of assigning a set of objects into groups (called clusters) so that the objects in the same cluster are more similar (in some sense or another) to each other than to those in other clusters. Clustering is a main task of explorative data mining, and a common technique for statistical data analysis used in many fields, including machine learning, pattern recognition, image analysis, information retrieval, and bioinformatics.

**PROCEDURE:**

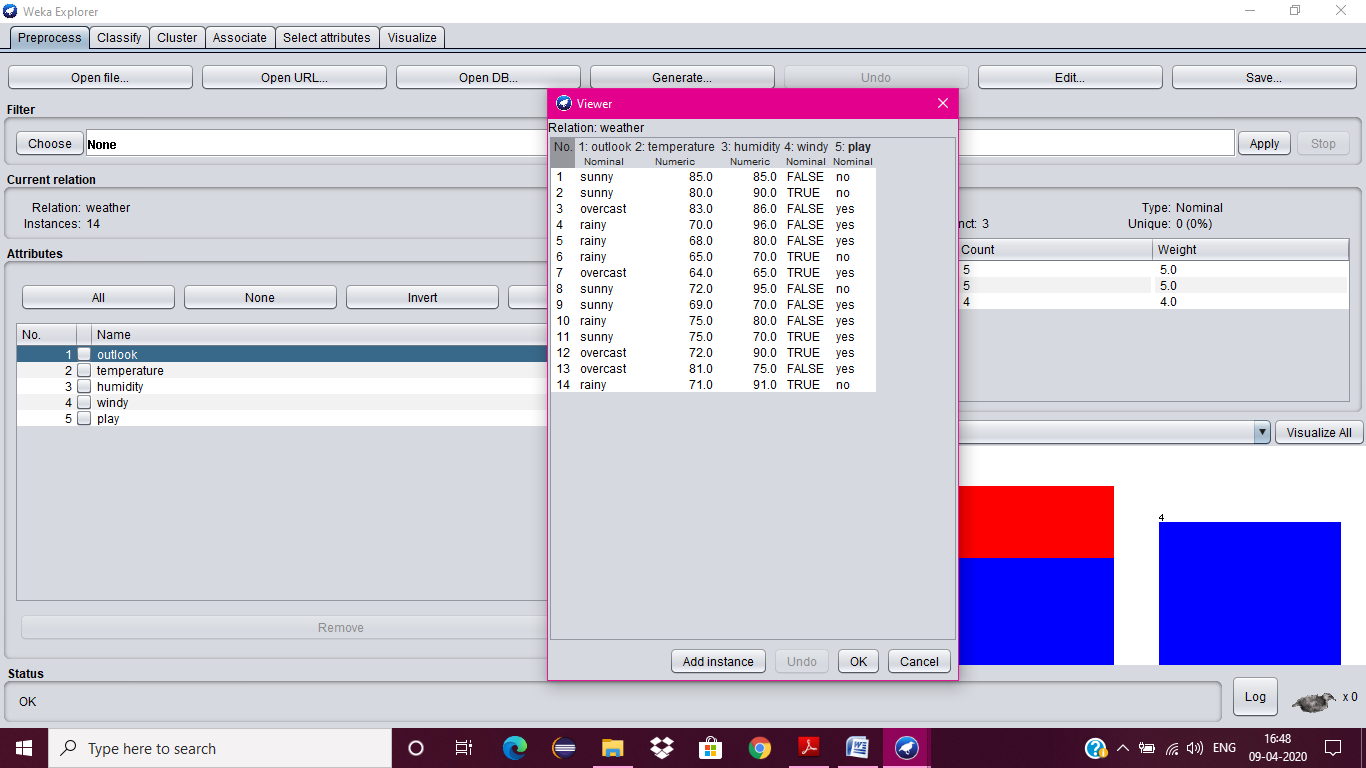
1. Create or download **weather.arff** file having weather data.
2. Click on weka and then click on explorer.



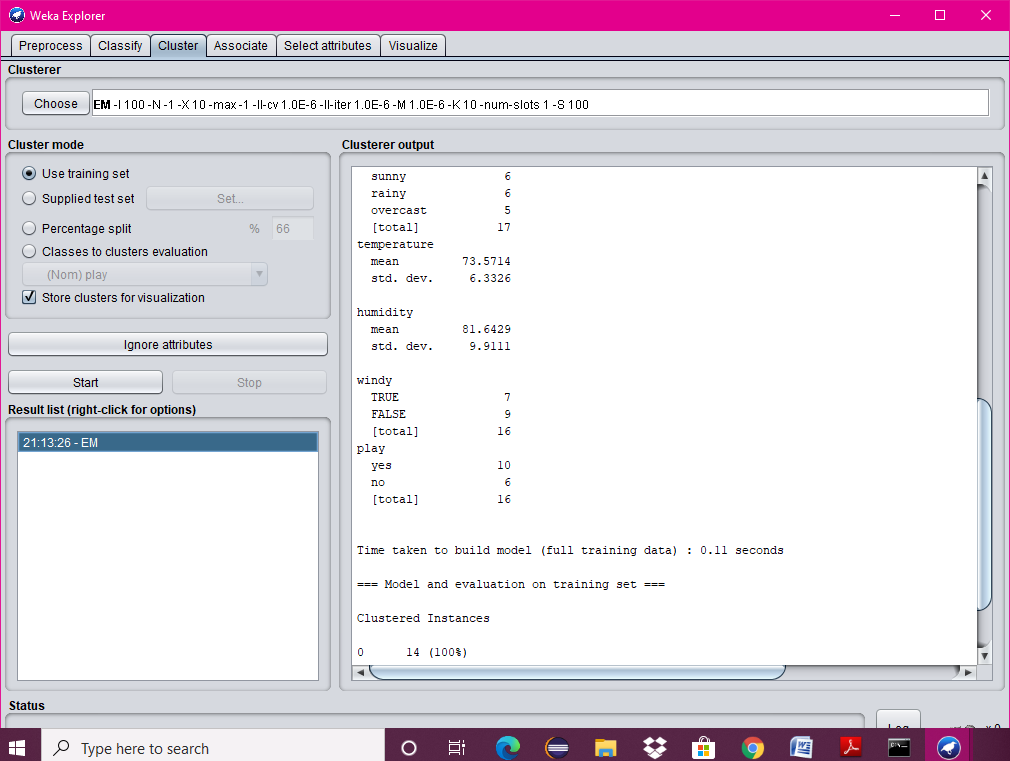
1. Explorer shows many options. In that click on ‘open file’ and select the arff file.



1. Click on edit button which shows weather table on weka.



1. Close the file.
2. Click on Cluster menu. In this there are different algorithms are there.
3. Click on Choose button and then select EM algorithm.
4. Click on Start button and then output will be displayed on the screen.



**2.**

**AIM:**

Write a procedure for Clustering Banking data using the farthest first Algorithm.

**DESCRIPTION:**

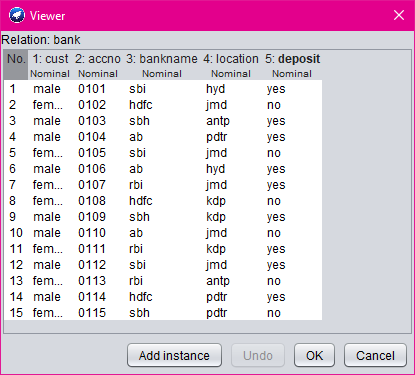
Cluster analysis or clustering is the task of assigning a set of objects into groups (called clusters) so that the objects in the same cluster are more similar (in some sense or another) to each other than to those in other clusters. Clustering is a main task of explorative data mining, and a common technique for statistical data analysis used in many fields, including machine learning, pattern recognition, image analysis, information retrieval, and bioinformatics.

**PROCEDURE:**

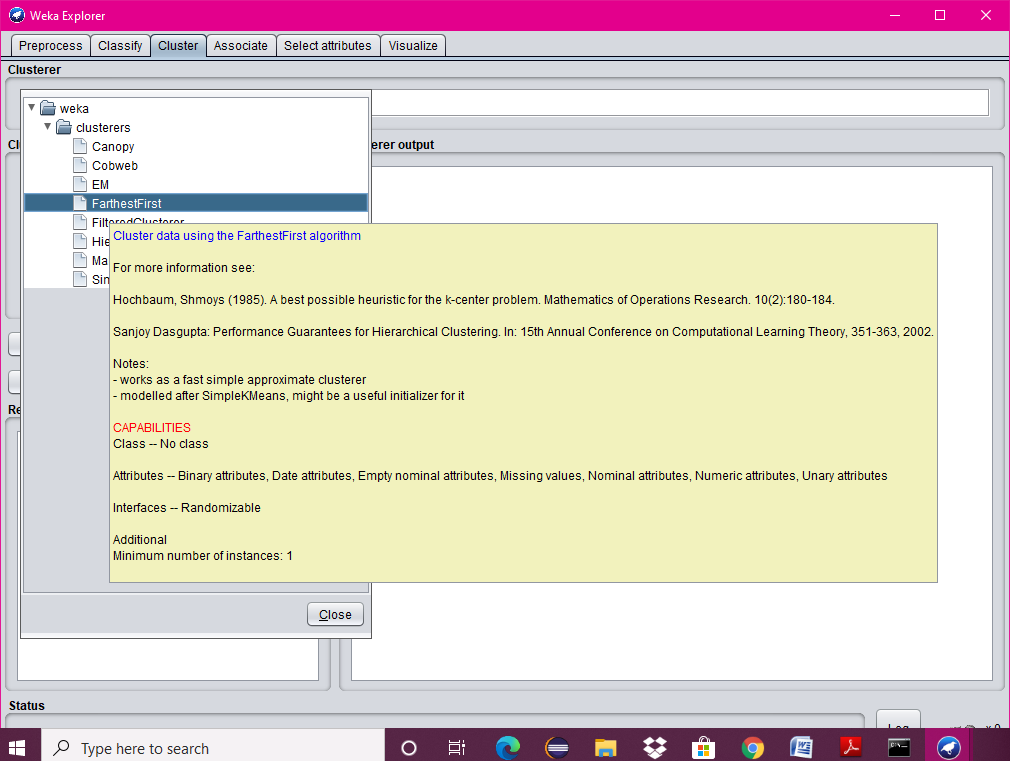
1. Create or download **banking.arff** file having banking data.
2. Click on weka and then click on explorer.



1. Explorer shows many options. In that click on ‘open file’ and select the arff file.
2. Click on edit button which shows banking table on weka.



1. Close the file.
2. Click on Cluster menu. In this there are different algorithms are there.
3. Click on Choose button and then select **FarthestFirst** algorithm.



1. Click on Start button and then output will be displayed on the screen.

