Experiment 1

Shashwat Shah
60004220126
Div B C2-2

Aim: Perform data preprocessing task using weka data mining tool.

Theory: weka is an opensource software that performs tools for pre-processing implementation of several machine learning algorithms & visual tools.

We can use the weka tool to perform

1) Pre-processing
2) Classification
3) Clustering
4) Association Rule
5) Visualization.

Pre-processing - It involves cleaning & transforming raw data into a format suitable for analysis. The goal is to enhance the data & ensure that it is well suited for a specific requirement of data mining task.

It involves steps such as:-

1. Data cleaning - Addressing errors & inconsistencies in data
2. Data integration - Combine data from different sources.
3. Data transformation - changing the format or structure of the data.
4. Data reduction - Reducing the volume but producing the same result.
5. Data Discretization - Transforming continuous data into discrete categories

(6) Handling noisy data - Make sure of the data.
(7) Normalization - scaling numerical features to the range
(8) Handling the imbalanced data.
(9) Data Imputation. - Filling the missing data.
(10) Feature Engineering

Conclusion: Thus the weka mining tool was explored using the weka tool tasks like pre-processing.