**Name = Nihar**

**Sap ID = 500091867**

**Roll No = R2142210525**

**Experiment No -1**

**Title** - Introduction to Weka Tool

Q) **Short Note On Weka Tool**

**Ans Weka Tool is named after flightless New Zealand bird. Weka is a set of machine learning algorithm that can be applied to a data set directly. It contain tools for data pre-processing, classification, regression, clustering, association rules and visualization. It is data mining software that uses a collection of machine learning algorithm. These algorithms can be applied directly to the data or called from the java code.**

**Weka is a collection of tools for:**

**1. Regression**

**2. Clustering**

**3. Association**

**4. Data pre-processing**

**5. Classification**

**6. Visualization**

Q) **Explore the tools with different options like :-**

* **Preprocess:- According to Weka preprocessing refers to the process of cleaning, transforming, and preparing data for analysis. This can include tasks such as removing missing values, converting data types, normalizing numerical values, and encoding categorical variables. Weka provides a number of tools for pre-processing data, including the Data Cleaner, Attribute Filter, and Attribute Selection filters. You can access these tools by opening the "Pre-process" tab in the Weka Explorer, and then selecting the appropriate filter from the list. Once you have applied the necessary filters, you can save the pre-processed data for further analysis or modelling.**
* **Classify:- Classification is the process of using a trained model to predict the class label of new instances. Weka provides a number of algorithms for classification, including decision trees, naive Bayes, k-nearest neighbour’s, and support vector machines.**
* **Cluster:- Clustering refers to the process of grouping similar instances together based on their attributes. Weka provides several algorithms for clustering, including k-means, hierarchical clustering, and density-based clustering, which can be found under the "Cluster" tab in the Weka Explorer.**
* **Associate:- Association rule mining is the process of discovering relationships between items in large datasets. Weka provides an implementation of the Apriorism algorithm for mining association rules, which can be found under the "Associate" tab in the Weka Explorer.**
* **Select Attributes:- Attribute selection refers to the process of selecting a subset of attributes from a dataset that are most relevant to a specific task, such as classification or clustering. Weka provides several attribute selection algorithms, including Correlation-based Feature Selection (CFS), Information Gain, and Relief.**
* **Visualize:- Visualizing data refers to the process of representing data in a graphical format to help understand and interpret the data better. Weka provides several visualization tools that can be used to visualize data, including scatter plots, histograms, and box plots. These visualization tools can be found under the "Visualize" tab in the Weka Explorer.**