Aim: To Perform BI using Weka

Lab Outcome:

- To Explore Weka's interface for pre-processing and classification tasks.
- To apply clustering and association mining with result visualizations.

Theory:

Q1. What is Weka?

Weka (Waikato Environment for Knowledge Analysis) is a free, open-source machine learning and data mining software suite written in Java.

It Developed at the University of Waikato, New Zealand, it is a cornerstone tool for both education and research in data science.

It provides a graphical user interface (GUI), command-line interface, and a comprehensive Java API, enabling users to apply a vast library of algorithms to datasets for tasks such as classification, regression, clustering, association rule mining, and data visualization.

Weka excels in data pre-processing, feature selection, and experimental evaluation, making it an invaluable platform for prototyping models, analysing data patterns, and building foundational knowledge in business intelligence and predictive analytics without requiring extensive programming expertise.

Q2. Explain the features of Weka.

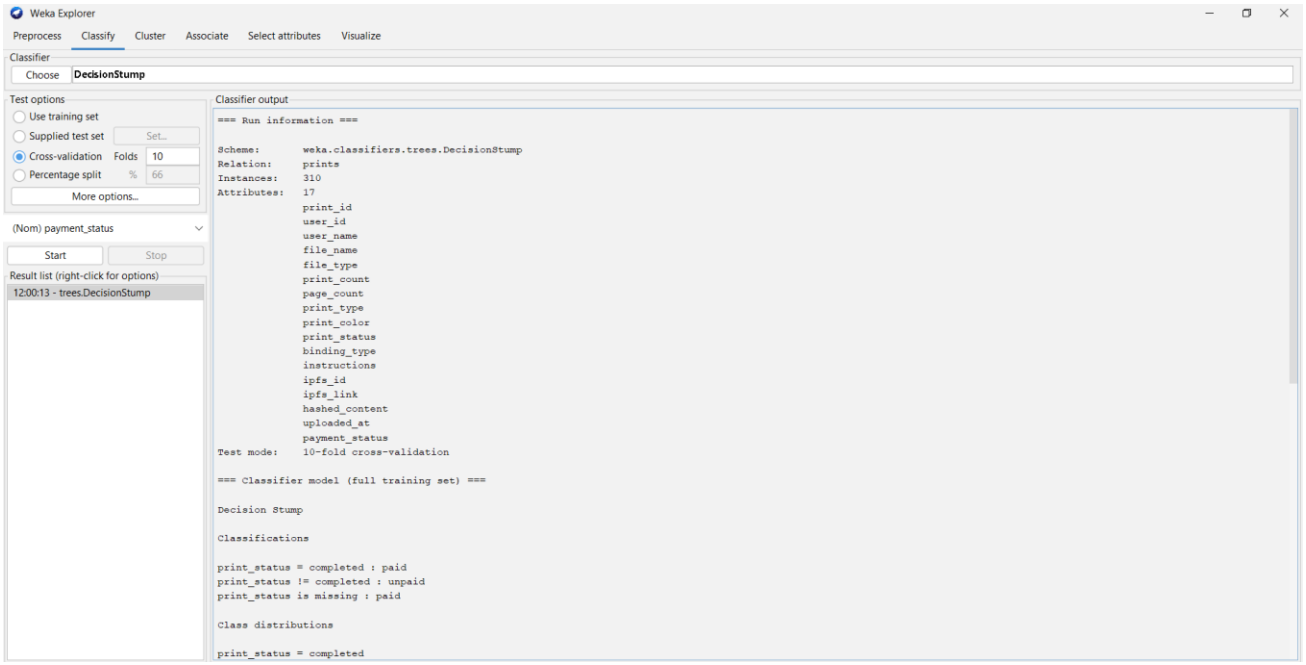
a) Pre-process: Import, clean, transform, and filter datasets.
b) Classify: Apply supervised learning algorithms (e.g., decision trees, naive Bayes).
c) Cluster: Apply unsupervised clustering algorithms (e.g., k-means, hierarchical).
d) Associate: Discover association rules (e.g., Apriori).
e) Select Attributes: Identify the most relevant features or attributes.
f) Visualize: Plot data and results in 2D/3D graphs.

Q3. Give Installation steps for Weka.

1) Download the installer from the official Weka website (stable version for your OS).
2) Run the installer and follow the setup instructions (ensure Java is installed).
3) Launch Weka GUI Chooser.
4) Select "Explorer" to start data mining.

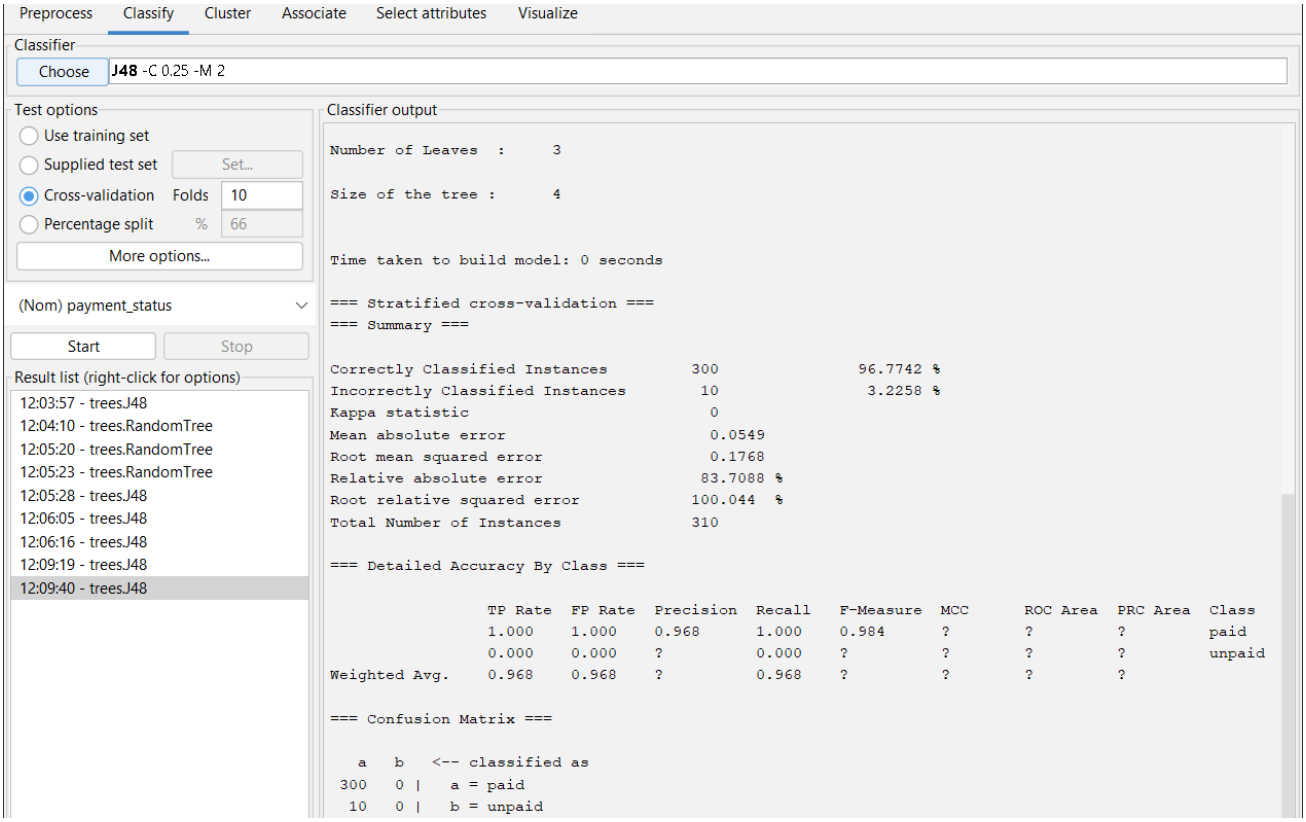
Q4. How to a Load Dataset in Weka?

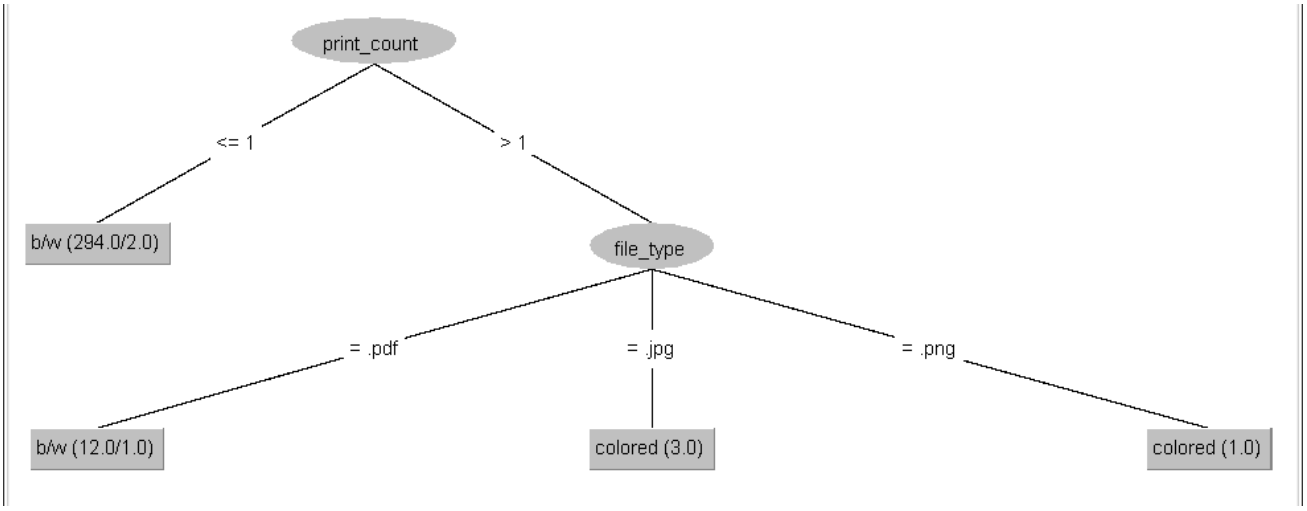
1) Download or create a dataset.
2) Open Weka Explorer.
3) Click "Open File" under the 'Pre-process' tab.
4) Choose a dataset file (e.g., .arff, .csv).
5) Click "Open" – data attributes and statistics will appear.



Q5. Classify (J48 Decision Tree) with Visualization in Weka.

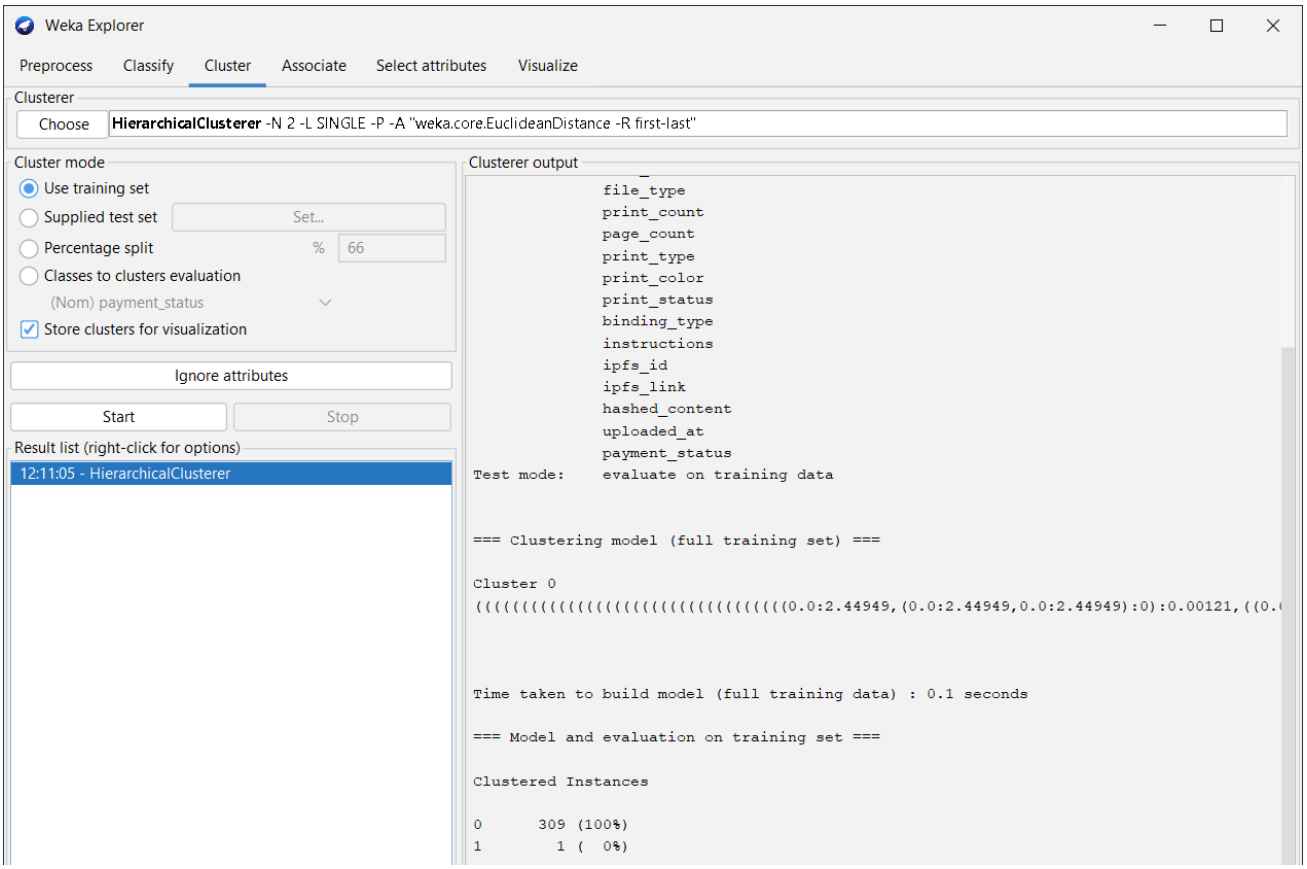
1) Go to “Classify” tab.
2) Click “Choose” → trees → J48.
3) Click “Start”.
4) Right-click the result in the list → “Visualize tree” to view the tree.

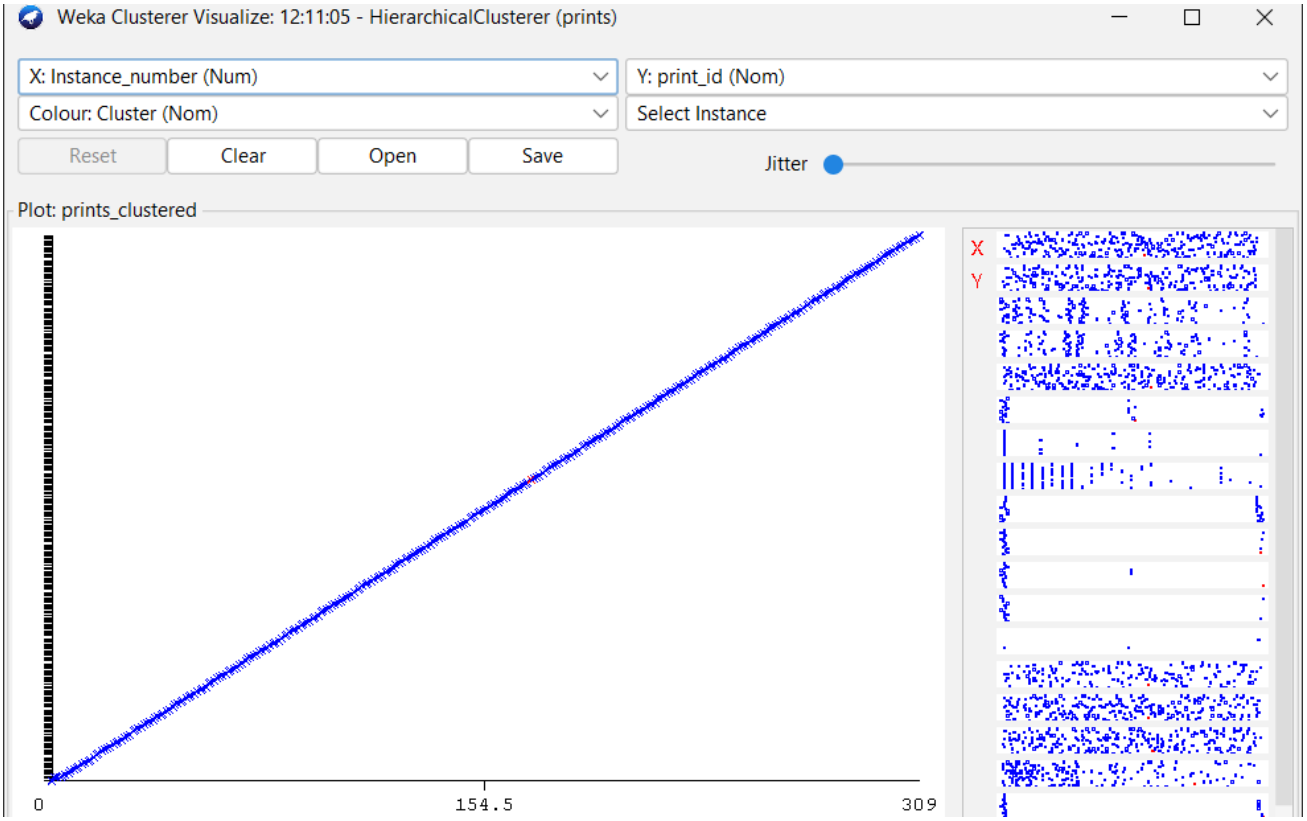


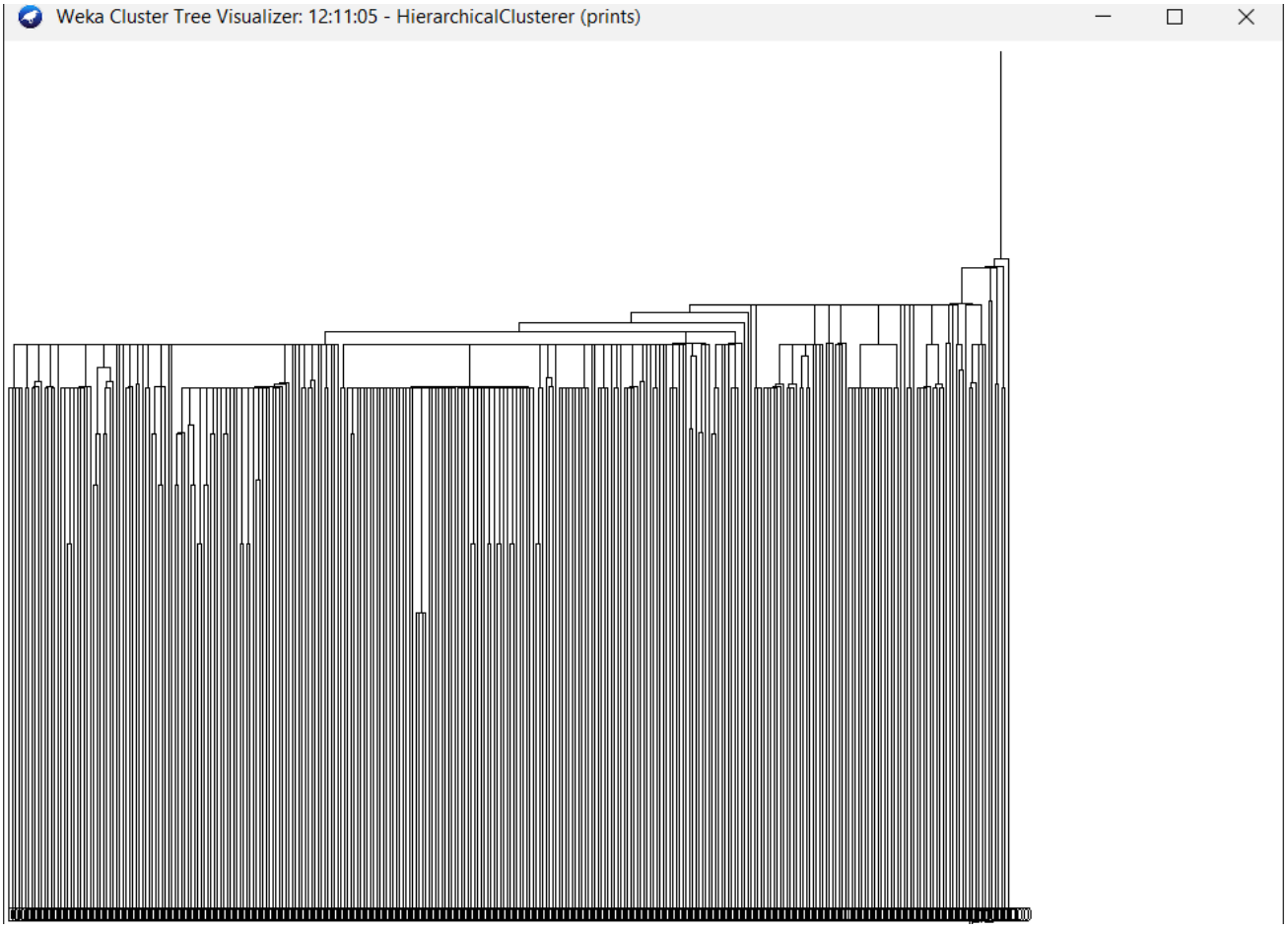


Q6. Clustering (Hierarchical Cluster) with Visualization in Weka.

1) Go to “Cluster” tab.
2) Click “Choose” → hierarchical → HierarchicalClusterer.
3) Set parameters (e.g., number of clusters, linkage type).
4) Click “Start”.
5) Right-click result → “Visualize cluster assignments”.

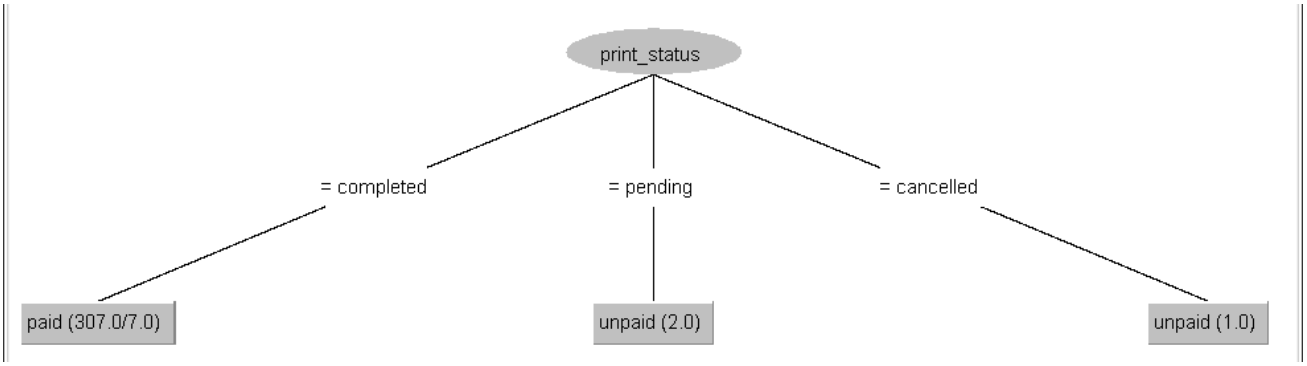



Weka Clusterer Visualize: 12:11:05 - HierarchicalClusterer (prints)
X: Instance_number (Num)
Y: print_id (Nom)
Colour: Cluster (Nom)
Select Instance
Reset
Clear
Open
Save
Jitter
Plot: prints_clustered
X
Y
0
154.5
309


Weka Cluster Tree Visualizer: 12:11:05 - HierarchicalClusterer (prints)

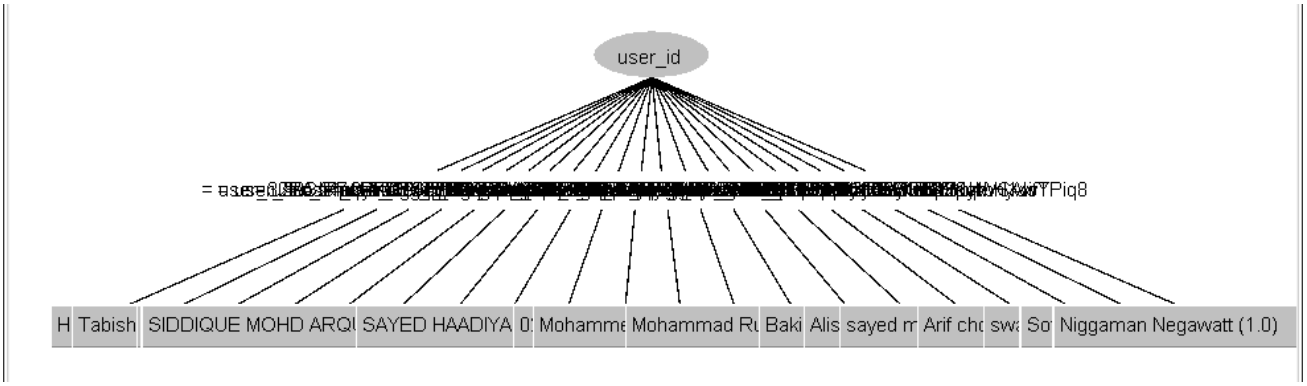
Q7. Association (Filtered Associator) in Weka.

1) Go to “Associate” tab.
2) Click “Choose” → “FilteredAssociator”.
3) Configure the base associator (e.g., Apriori) and any filters.
4) Click “Start” – rules will appear in the output.



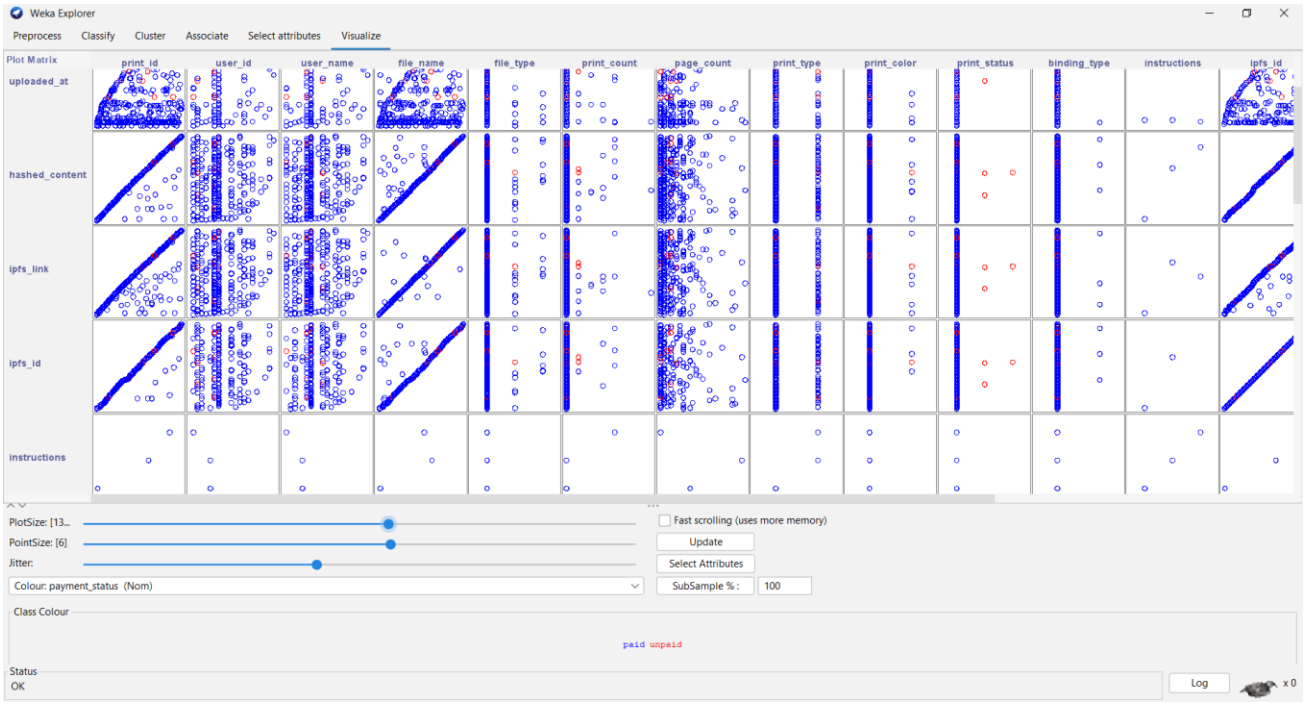
Q8. Select Attributes with Visualization in Weka.

1) Go to “Select attributes” tab.
2) Choose an attribute evaluator (e.g., InfoGain) and search method.
3) Click “Start”.
4) The output shows ranked attributes.
5) Use “Visualize” tab to plot selected attributes.



Q9. Visualize (Graphs & Plots) in Weka.

1) Go to “Visualize” tab in Explorer.
2) A scatter plot matrix of all attributes appears.
3) Click any small plot to open a large interactive 2D plot.
4) Adjust colours, point size, and axes using the controls.



Conclusion:

Successfully demonstrated Weka's effectiveness for core BI tasks: data pre-processing, classification, clustering, association mining, and visualization, providing actionable insights from raw datasets through an accessible graphical interface.