

OUTPUT

- 1) Loading the data we can load dataset into weka by clicking an open button in preprocessing interface and selecting the appropriate file.
- 2) Once the data is loaded, weka will recognize the attributes and during the scan of the data weka will complete some basic strategies on each attribute.
- 3) Clicking on an attribute in the left panel will show the basic statistics on the attributes we can obtain min, max, mean, standard deviation and deviation etc.
- 4) The Visualization in the right button panel in the form of Cross tabulation across two attributes.
- 5) Remove: Selecting or filtering attributes
 Removing: Choose Tab → weka → filters → unsupervised → attribute → remove
 Replacing Missing Values: Choose Tab → weka → filters → unsupervised → attribute →
 • Nominal to Binary → replacing missing values
 Same for Nominal to Binary
 Discretize: Choose Tab → weka → filters → unsupervised → attribute → Discretize.

Preprocess | Classify | Cluster | Associated | Selected attributes | Visualize

Open file... | Open URL | Open DB | Generate | Undo | Edit | Save

Filter

Current Relation

Relation: student-weka-filter Attributes: 6
Instance: 14 Sum of weights: 14

Selected Attribute

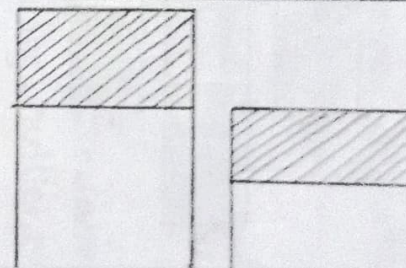
Name: age Distinct: 3 Type: Nominal
Missing: 0 (0%) Unique: 0 (0%)

Attributes

No.		Name
1	<input type="checkbox"/>	Age
2	<input type="checkbox"/>	Price
3	<input type="checkbox"/>	Income
4	<input type="checkbox"/>	Student
5	<input type="checkbox"/>	Credit - rating
6	<input type="checkbox"/>	buyspc

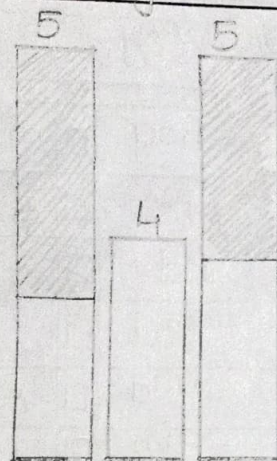
No.	Label	Count	Weight
1	<30	5	5.0
2	30-40	4	4.0
3	>40	5	5.0

Class: busype (Nom) ☒ Visualize all

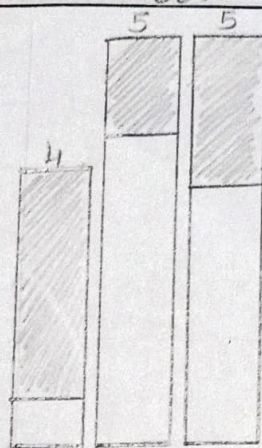


6

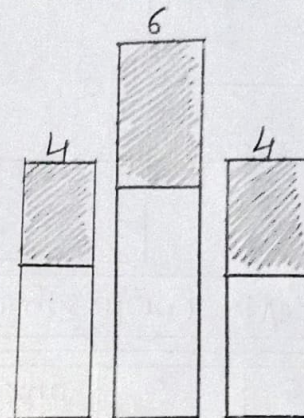
Age



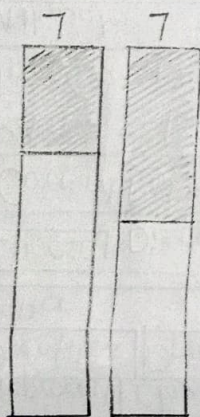
Prices



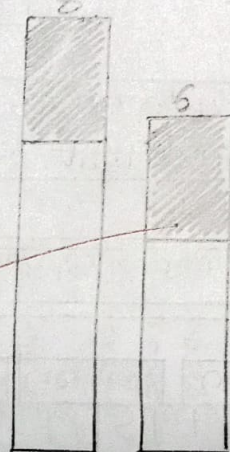
Income



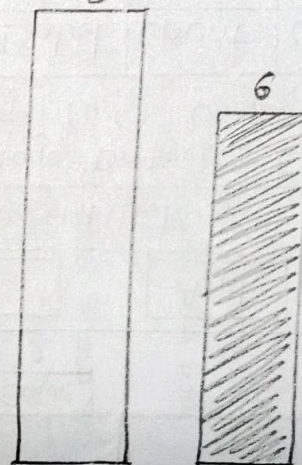
Student



Credit-rating



buyspc



VIVA QUESTIONS

1. What is the main goal of Data Mining?

Ans. To find patterns that can lead to inferences or predictions from large and unstructured data sets.

2. What are the Two Types of Data Mining Tasks?

Ans. Two types of Data Mining Tasks are descriptive tasks and predictive tasks.

3. What are the Data Mining Functionalities?

Ans. Data characterization, Data Discrimination, Association analysis, classification, prediction, clustering, outlier, evolution.

4. What are the Features of WEKA?

Ans. quickly solve data, solve storage problem of infrastructure engineers, it's good if it also has a lot of read traffic.

5. Navigate the options available in the WEKA.

Ans. Explorer, Experimenter, Knowledge flow, Console, preprocessor, classify, cluster, associate, select attributes, visualize.

6. What is ARFF file format?

Ans. Header Section: '@relation' - data set name '@attribute' - definitions, '@data' to mark start of the data

data Section: actual data instances with values separated by comma.

Ans
22/08/2024

Week - 1

Aim:-

This experiment illustrates some of the basic data processing operations that can be performed using WEKA - Explorer. The simple data available in arff format.

Step 1:-

loading the data. We can load the dataset into weka by clicking on open button in preprocessing interface and selecting the appropriate file.

Step 2:-

Once the data is loaded, weka will recognize the attributes and during the scan of the data weka will complete some basic strategies on each attribute. The left panel in the above figure shows the list of recognized attributes while the top panel indicates the names of the base relation or table and the current working relation (which are same initially)

Step 3:-

Clicking on an attribute in the left panel will show the basic statistics on the attributes for the categorical attributes the frequency of each attribute value is shown while for continuous attributes we can obtain min, max, mean, standard deviation and deviation etc.

Step 4:-

The visualization in the right button panel in the form of cross-tabulation across two attributes.

Step 5:-

Remove: Selecting or filtering attributes Removing an

chooseTab → weka → filters → unsupervised
→ attribute → remove.

Replace Missing Values:-

chooseTab → weka → filters → unsupervised
→ attribute → Replace missing values

Nominal to Binary:-

chooseTab → weka → filters → unsupervised
→ attribute → Nominal to Binary

Discretize:-

chooseTab → weka → filters → unsupervised
→ attribute → Discretize.

Preprocess	Clarity	Cluster	Associated	Selected attribute	Visualize	
Open file..	Open URL	Open DB	Generate	undo	edit	Save

filter

choose	Discardize (before selecting diagram)	Apply
--------	---------------------------------------	-------

Current Relation

Relation : student-weka.filter Attributes : 6
 Instance : 14 Sum of weights : 14

Attributes

All	None	Invert	Pattern
No.			Name
1	<input type="checkbox"/>		Age
2	<input type="checkbox"/>		Price
3	<input type="checkbox"/>		InCome
4	<input type="checkbox"/>		Student
5	<input type="checkbox"/>		Credit - rating
6	<input type="checkbox"/>		buyspc

Remove

Selected Attribute

Name: age			Type: Nominal
Missing: 0(0%)			Distinct: 3 unique: 0(0%)
No.	Label	Count	weight
1	< 30	5	5.0
2	30 - 40	4	4.0
3	> 40	5	5.0

Class: buyspc (NOM) V/Visualize ALL

