NAME:- SHIVPRASAD . C. PREMARAJAN.

ROLL NO:- 9696
SUB : DWM

Assignment 1

Q.1] A bank wants to develop a DW for effective decision making about their loan schemes. Loans are provided for various purpose (Home, Car, Education, etc). The whole country is categorized into regions namely (East, west, South, North). Loan is distributed to customers at the interest rate that changes from time to time. Different loans have different RDI. clearly design the star schema & snowflake schema.

⇒ Information Package Diagram.

Business Process: Loan Analysis

Time	Customer	Region	Loan.
Time_key	Customer_Key	Region_key	loanKey
Day	Account_number	Region_name	Loan_type
day_of_week	Account_type		ROI
month	Customer_name		Payment.
Quarter			
year.			

Fact: Loan_amount, No. of customer.

Manual 3/10/23

DWM Assignment no-2

Q.1] The following Data (in increasing order) for the attribute age: 13, 15, 16, 16, 19, 20, 20, 21, 22, 22, 25, 25, 25, 25, 30, 33, 33, 35, 35, 35, 35, 36, 40, 45, 46, 52, 70.

⇒

Partition into (equi-depth) bins.

Bin1 :- 13, 15, 16, 16, 19, 20, 20, 21, 22.

Bin2 :- 22, 25, 25, 25, 25, 30, 33, 33, 35

Bin3 :- 35, 35, 35, 36, 40, 45, 46, 52, 70

a) Smoothing by bin mean :-

Bin1 :- 18, 18, 18, 18, 18, 18, 18, 18, 18.

Bin2 :- 28, 28, 28, 28, 28, 28, 28, 28, 28

Bin3 :- 44, 44, 44, 44, 44, 44, 44, 44, 44.

b) Smoothing by bin median :-

Bin1 :- 19, 19, 19, 19, 19, 19, 19, 19, 19.

Bin2 :- 25, 25, 25, 25, 25, 25, 25, 25, 25

Bin3 :- 40, 40, 40, 40, 40, 40, 40, 40, 40.

Name:- Chavaratil Shivprasad Premsagar
Roll no:- 9696
FR. CONCEICAO RODRIGUES COLLEGE OF ENGINEERING

Marvellous

Assignment 3.

Q.1] Apply Naive Bayesian Classification Algorithm on following data

tuple (Homeowner = Yes, Status = Employed, Income = Avg)

⇒

p(Defaulted = 'yes') = 3/10

p(Defaulted = 'No') = 7/10

To calculate probabity

$$P(C_i/x) = \frac{p(x|C_i)\, p(C_i)}{p(x)}$$

$\because p(x)$ is constant for all classes

$\therefore p(C_i/x) = p(x|C_i)\, p(C_i)$

$p(x|C_i) = p(x|$Defaulted = 'yes'$)$

$= p($Homeowner = 'yes' $|$ Defaulted = 'yes'$) \times$

$= p($Status = 'Employed' $|$ Defaulted = 'yes'$) \times$

$= p($Income = 'Average' $|$ Defaulted = 'yes'$)$

$= 0 \times \frac{2}{3} \times 0$

$= 0$

$p($Defaulted = 'yes' $|\, x) =$

$\therefore p(x|$ Defaulted = 'yes'$) \times p($Defaulted = 'yes'$)$

$= 0 \times 3/10 = 0.$