

## CS 539: Homework 1

$$e(s) = 0.667 \times \left(\frac{6}{9}\right) \log_2\left(\frac{6}{9}\right) - \left(\frac{3}{9}\right) \log_2\left(\frac{3}{9}\right) = 0.9182$$

## Debt

low	med.	high
(1, -3)	(1, -1)	(-2, 0)
$e = 0.8112$	$e = 0.9182$	$e = 0.9182$
$ce = 0.8888$		

$$IG(\text{Debt}) = 0.667 \times 0.9182 - 0.8888 \\ = 0.0294$$

## Income

low	med.	high
(3, -2)	(0, -2)	(0, -2)
$e = 0.9709$	$e = 0$	$e = 0$
$ce = 0.5393$		

$$IG(\text{income}) = 0.3788$$

## Married

Yes	No
(2, -3)	(1, -3)
$e = 0.9709$	$0.8112$
$ce = \left(0.9709 \times \frac{5}{9}\right) + \left(0.8112 \times \frac{4}{9}\right)$	
$ce = 0.8999$	

$$IG(\text{married}) = 0.018$$

### Own's Property

Yes $(0, -3)$ $e = 0$ $ce = \left( \frac{3}{9} \times 0 \right) + \left( \frac{6}{9} \times \cancel{-3} \right)$ $ce = \cancel{0} - 2/3$	No $(-3, +3)$ $e = \cancel{-3}$ $ce = \cancel{0} + 1$
--	--

$$IG(\text{Own's property}) = \cancel{0.5848} \approx 0.2515$$

### Gender

Male $(-4, +2)$ $e = 0.9182$ $ce = \left( 0.9182 \times \frac{6}{9} \right) + \left( 0.9182 \times \frac{3}{9} \right)$ $ce = 0.9182$	Female $(-2, +1)$ $e = 0.9182$
---	--------------------------------------

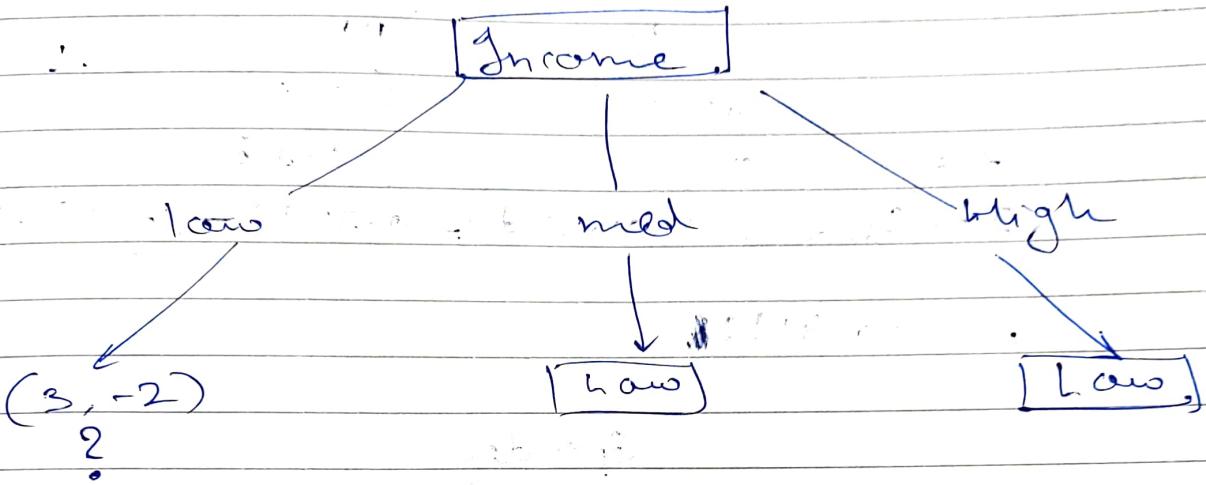
$$IG(\text{Gender}) = 0$$

### IG:

Debt	-	0.294 ✓
Income	-	0.3788 ✓ ← Highest
Married	-	0.018 ✓
Own's property	-	<del>0.5848</del> 0.2515 ✓
Gender	-	0 ✓

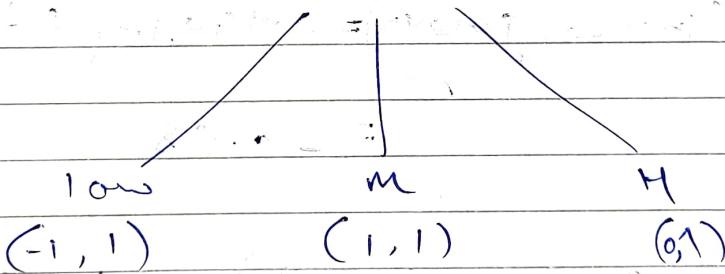
~~∴ IG of "Own's property" is highest node = Own's property~~

~~Splitting~~ Since the IG of Income is the highest, Income is the root node.



$$\therefore S_{\text{income}} = \frac{1}{5} (3, -2) + e(S_{\text{income}}) = \underline{0.9709}$$

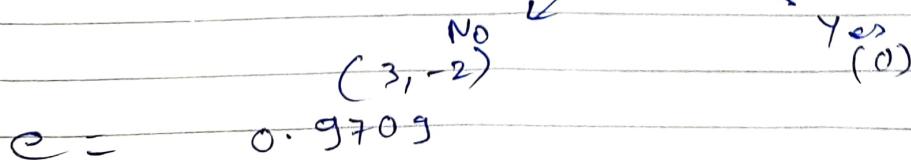
Debt



$$e = 1$$

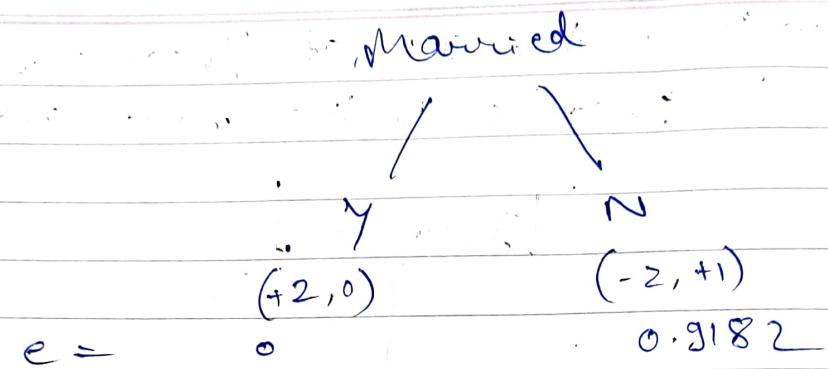
$$\text{IG}(S_{\text{income}}, \text{Debt}) = 0.9709 - \frac{2}{5} - \frac{2}{5} = 0.1709$$

Owns Property



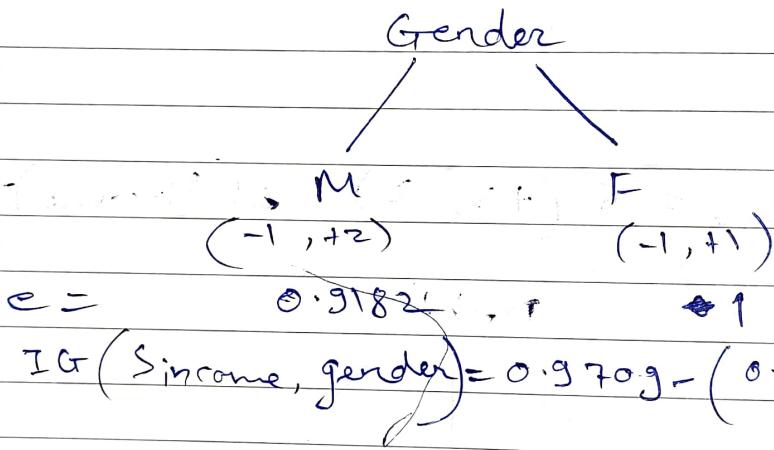
$$e = 0.9709$$

$$\begin{aligned} -\text{IG}(S_{\text{income}}, \text{Own's prop.}) &= 0.9709 - \left(\frac{5}{5} \times 0.9709\right) \\ &= 0. \end{aligned}$$

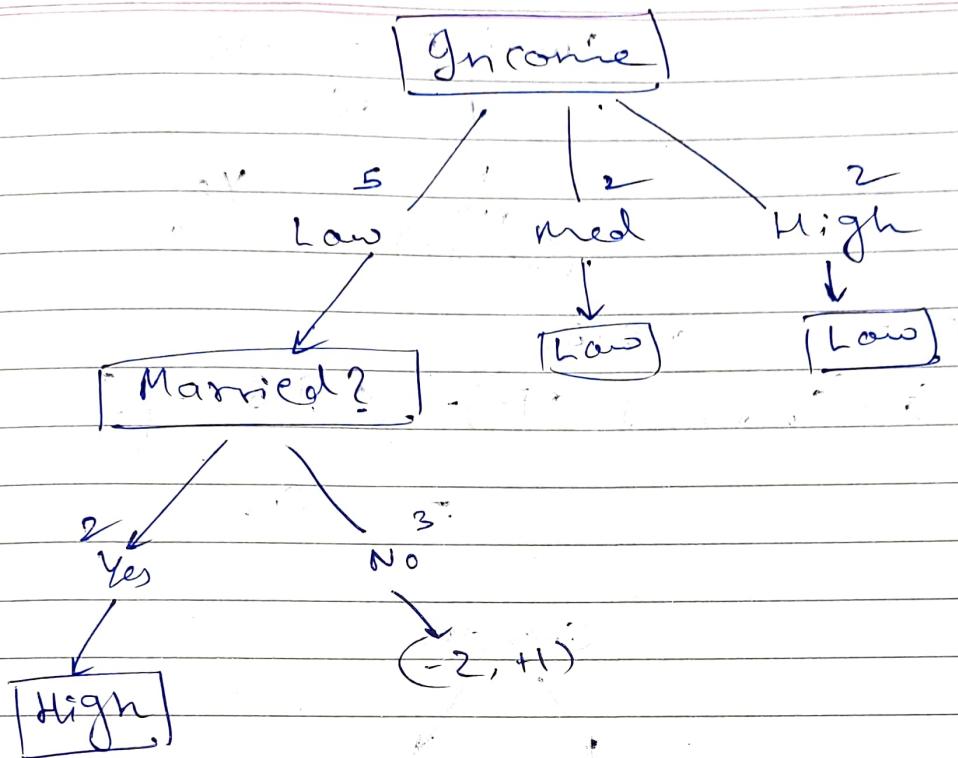


$$IG(S_{\text{income}}, \text{married}) = 0.9709 - \left( \frac{3}{5} \times 0.9182 \right)$$

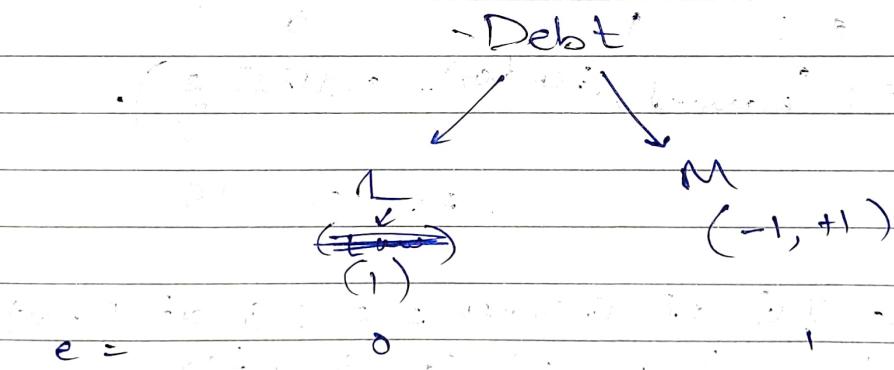
$$= 0.4199$$



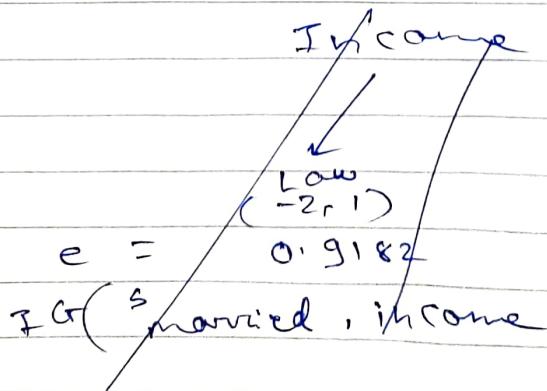
$\therefore IG(S_{\text{income}}, \text{married})$  is the highest.  
ie selected married.



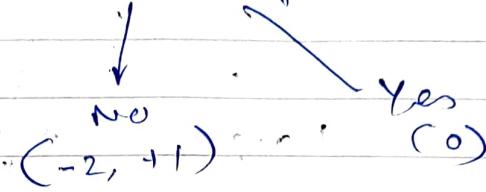
$$S_{\text{married}} = 3 \cdot (-2, +1) \quad e = 0.9182$$



$$\text{IG}(\text{S}_{\text{married}}, \text{Debt}) = 0.9182 - \left( \frac{2}{3} \times 1 \right) = 0.2515$$



Owns Property



$$e = 0.9182$$

$$\text{IG}(\text{Smashed}, \text{OP}) = 0.9182 - \left( \frac{2}{3} \times 0.9182 \right)$$

$$= 0.$$

Gender

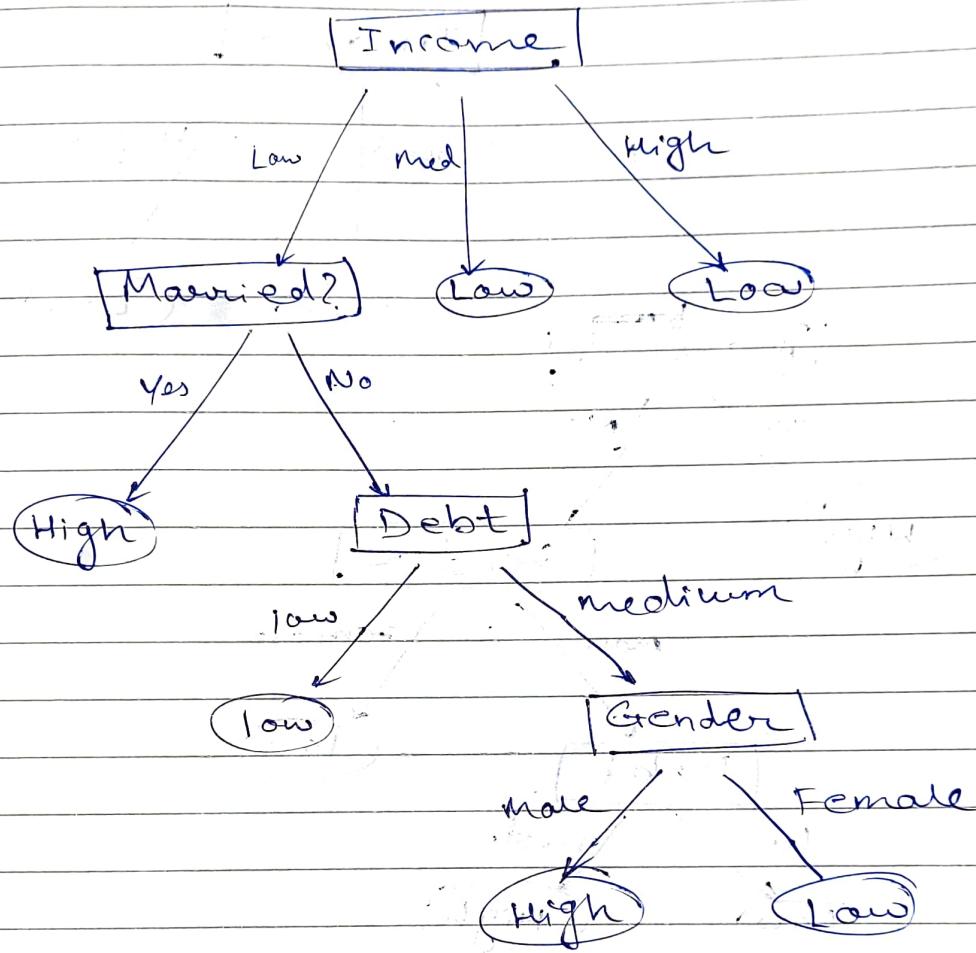


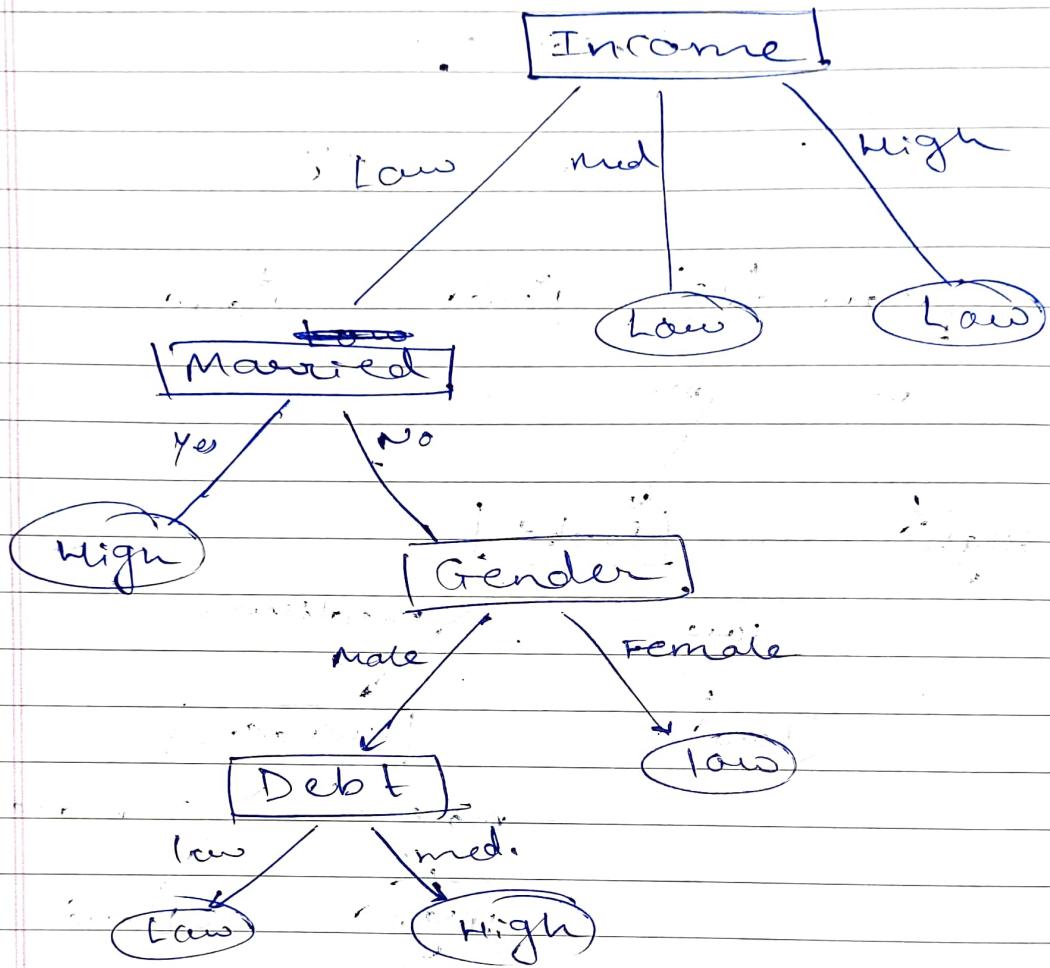
$$e = 1 - 0 = 0$$

$$\text{IG}(\text{Smashed}, \text{Gender}) = 0.9182 \left( \frac{2}{3} \right).$$

$$= 0.2515$$

$\therefore$  the IG of Gender == IG of Debt,  
there are 2 possible solutions.

Solution 1

Solution 2

## Credit Risk Prediction:

1) Tom

Soln. 1 : Credit Risk = Low. (without using Gender & Own's prop.)

Soln. 2 : Credit Risk = Low (using Gender but not Own's property)

2) Ana

Soln 1: Low

∴ Ana's income is medium, we directly reach the leaf node / target label

Sol 2: Low

### \* Observation:

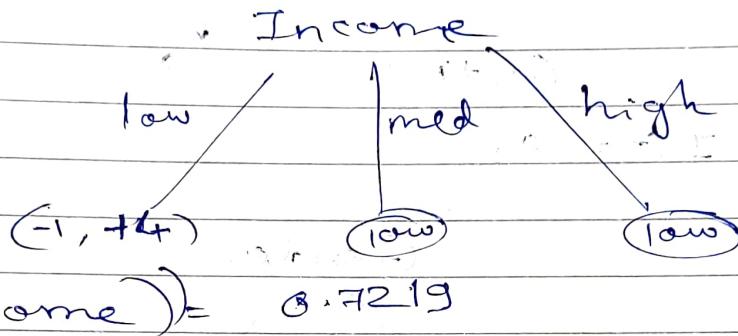
- we got the same prediction using both the solutions; however, the depth of the tree increased by one level in soln. 2
- ~~Since~~, soln. 1 is the more optimal tree as its depth is less than the tree presented in soln. 2.

### \* Task 2-2

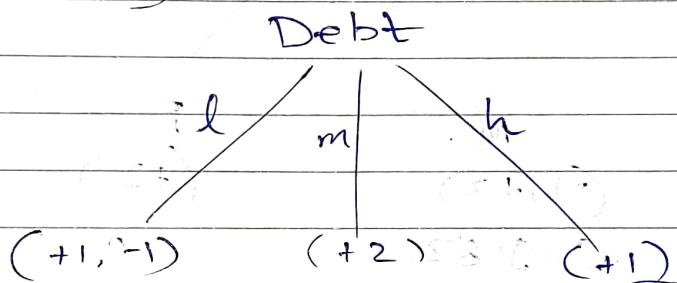
Feature not playing a role: "Owns\_Property"

CLASSMATE  
Date \_\_\_\_\_  
Page \_\_\_\_\_

Task 2-2: If Sofia's credit risk is high  
Root Node = 'Income' as it has the highest information gain.



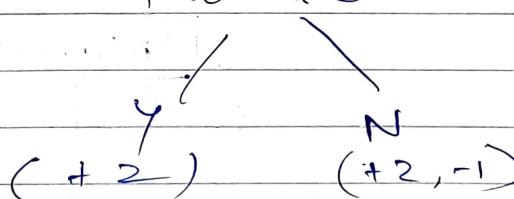
$$e(S(\text{Income})) = 0.7219$$



$$IG(\text{Debt}) = 0.7219 - \left( \frac{2}{5} \times 1 \right) = 0.3219.$$

IG: Information Gain  
= 0.3219

Married

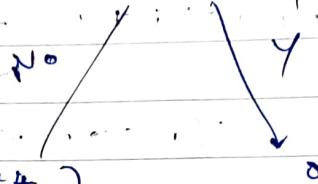


$$e = 0 \quad 0.9182$$

$$IG(\text{Married}) = 0.7219 - \left( \frac{3}{5} \times 0.9182 \right)$$

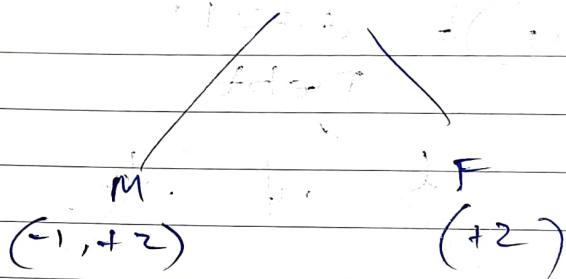
$$= 0.1709$$

Owes Property



$$IG = 0.$$

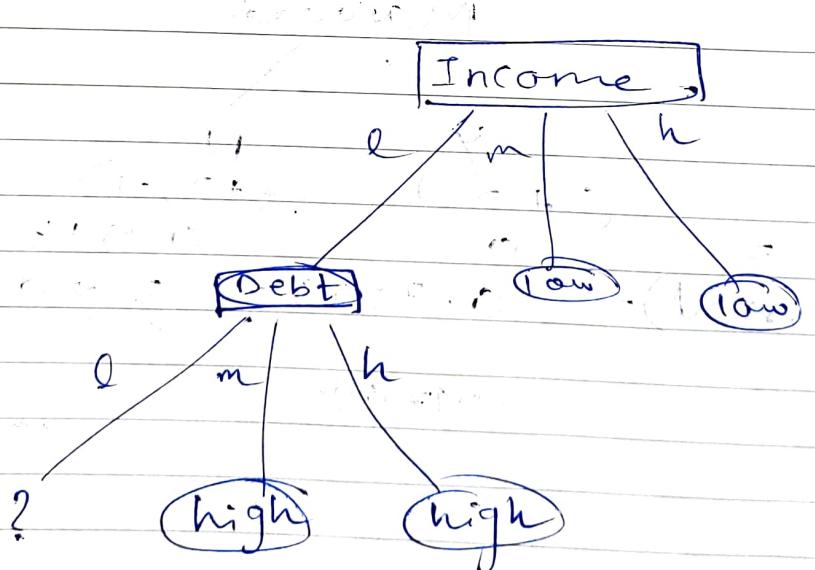
Gender



$$C = 0.9182 \quad (0.1709)$$

$$IG = 0.1709.$$

Since Debt has the highest IG,  
next node is Debt.



Only possible node here is 'Married'

2. Solution :-

