# Q1:

## a)



Naive Bayes — A — RNN
Naive Bayes — D — Decision Tree
RNN — C — Decision Tree
Decision Tree — B — Linear Regr.

## b)

Project Beta: Unsupervised

Project Gamma: Reinforcement

# Q2:

## a)

Age Group

$\boxed{4+ \mid 4-}$

Young / Middle | Senior

$\boxed{2+ \mid 2-}$  $\boxed{1+ \mid 1-}$  $\boxed{1+ \mid 1-}$
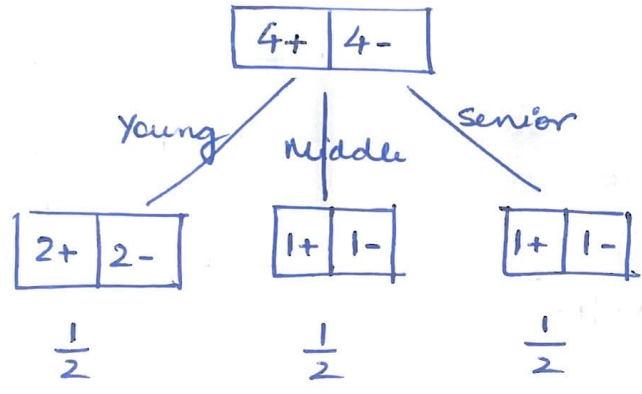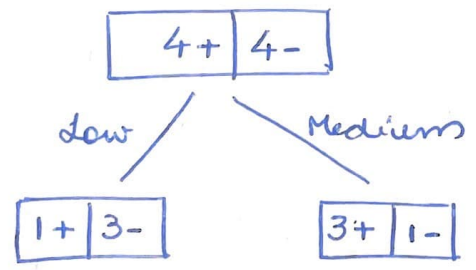
$\frac{1}{2}$     $\frac{1}{2}$     $\frac{1}{2}$

weighted =

$\frac{4}{8} \times \frac{1}{2} + \frac{2}{8} \times \frac{1}{2} + \frac{2}{8} \times \frac{1}{2}$
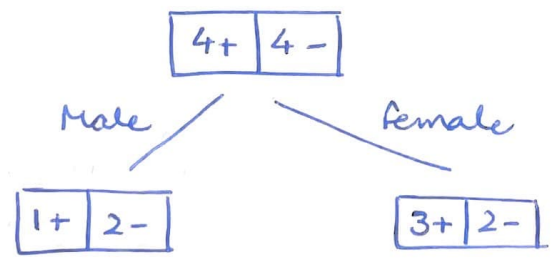
$= \frac{1}{2} = 0.5$

Income Range

$\boxed{4+ \mid 4-}$

Low /   Medium

$\boxed{1+ \mid 3-}$    $\boxed{3+ \mid 1-}$

$1 - \left(\frac{1}{4}\right)^2 - \left(\frac{3}{4}\right)^2 = \frac{3}{8}$    $\frac{3}{8}$

weighted $= \frac{4}{8} \times \frac{3}{8} + \frac{4}{8} \times \frac{3}{8}$

$= \frac{3}{8} = 0.375$

Gender

$\boxed{4+ \mid 4-}$

Male /   Female

$\boxed{1+ \mid 2-}$    $\boxed{3+ \mid 2-}$

$1 - \left(\frac{1}{3}\right)^2 - \left(\frac{2}{3}\right)^2 = \frac{4}{9}$    $1 - \left(\frac{3}{5}\right)^2 - \left(\frac{2}{5}\right)^2 = \frac{12}{25}$
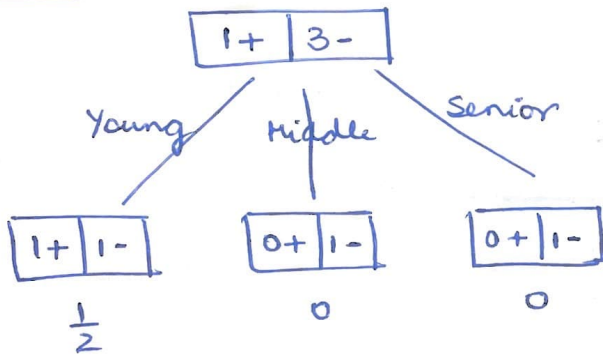
weighted $= \frac{3}{8} \times \frac{4}{9} + \frac{5}{8} \times \frac{12}{25}$

$= \frac{28}{60} = 0.467$

Select income for splitting

Income:
$$\boxed{4+ \mid 4-}$$
   Low / ＼ Medium
$$\boxed{1+ \mid 3-}$$        $$\boxed{3+ \mid 1-}$$
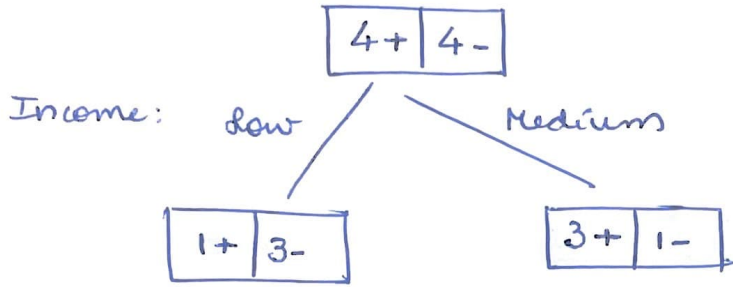
| Customer | AG | Gender | PD |
|---|---|---|---|
| 1 | Young | M | N |
| 3 | Middle | M | N |
| 5 | Senior | F | N |
| 7 | Young | M | Y |

| Customer | AG | Gender | PD |
|---|---|---|---|
| 2 | Young | F | Yes |
| 4 | Middle | F | Yes |
| 6 | Senior | F | Yes |
| 8 | Young | F | No |

**AG:**

$$\boxed{1+ \mid 3-}$$
Young / Middle ＼ Senior
$$\boxed{1+ \mid 1-}$$   $$\boxed{0+ \mid 1-}$$   $$\boxed{0+ \mid 1-}$$
$\frac{1}{2}$            0            0

weighted $= \frac{1}{2} \times \frac{1}{2} = 0.25$

**Gender:**

$$\boxed{1+ \mid 3-}$$
Male / ＼ female
$$\boxed{1+ \mid 2-}$$        $$\boxed{0+ \mid 1-}$$
$1 - \frac{1}{9} - \frac{4}{9} = \frac{4}{9}$        0
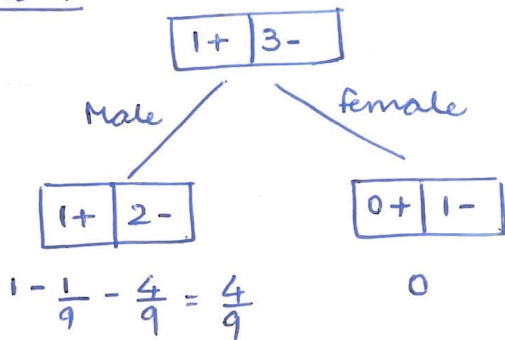
weighted $= \frac{4}{9} \times \frac{3}{4} = \frac{1}{3} = 0.33$

↓

AG will be Selected

**AG:**

$$\boxed{3+ \mid 1-}$$
Young / Middle ＼ Senior
$$\boxed{1+ \mid 1-}$$   $$\boxed{1+ \mid 0-}$$   $$\boxed{1+ \mid 0-}$$
$\frac{1}{2}$            0            0
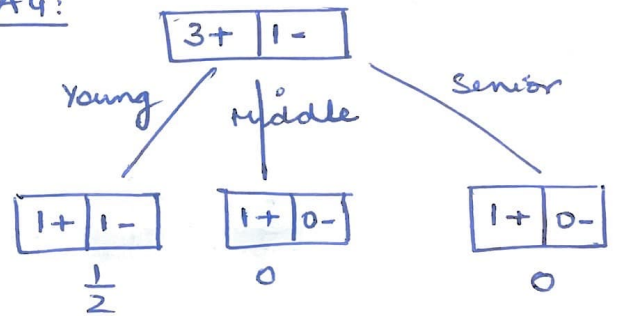
weighted $= \frac{1}{2} \times \frac{1}{2} = \frac{1}{4} = 0.25$

**Gender:**

$$\boxed{3+ \mid 1-}$$
Male / ＼ Female
$$\boxed{0+ \mid 0-}$$        $$\boxed{3+ \mid 1-}$$
1        $1 - \frac{9}{16} - \frac{1}{16} = \frac{6}{16}$
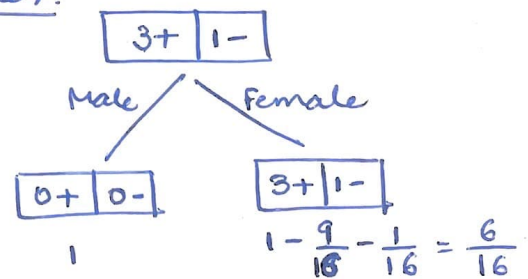
weighted $= \frac{0}{4} \times 1 + \frac{4}{4} \times \frac{6}{16} = 0.375$

↓

AG will be selected.

```
                         ┌─────┬─────┐
                         │ 4+  │ 4-  │
                         └─────┴─────┘
      Income:        low  ╱         ╲  Medium
                   ┌─────┬─────┐   ┌─────┬─────┐
                   │ 1+  │ 3-  │   │ 3+  │ 1-  │
                   └─────┴─────┘   └─────┴─────┘
```

Age Group:

Left subtree (low income):  Young / Middle | Senior ╲

- Young → ┌─────┬─────┐  │ 1+ │ 1- │ → **Yes** → (Tie Break)
- Middle → ┌─────┬─────┐  │ 0+ │ 1- │ → **NO**
- Senior → ┌─────┬─────┐  │ 0+ │ 1- │ → **NO**

Right subtree (Medium income): Young / Middle | Senior ╲

- Young → ┌─────┬─────┐  │ 1+ │ 1- │ → **Yes** → (Tie Break)
- Middle → ┌─────┬─────┐  │ 1+ │ 0- │ → **Yes**
- Senior → ┌─────┬─────┐  │ 1+ │ 0- │ → **Yes**

Tie Break table (left):

| Customer | Gender | AD |
|----------|--------|----|
| 1        | M      | N  |
| 7        | M      | Y  |

Tie Break table (right):
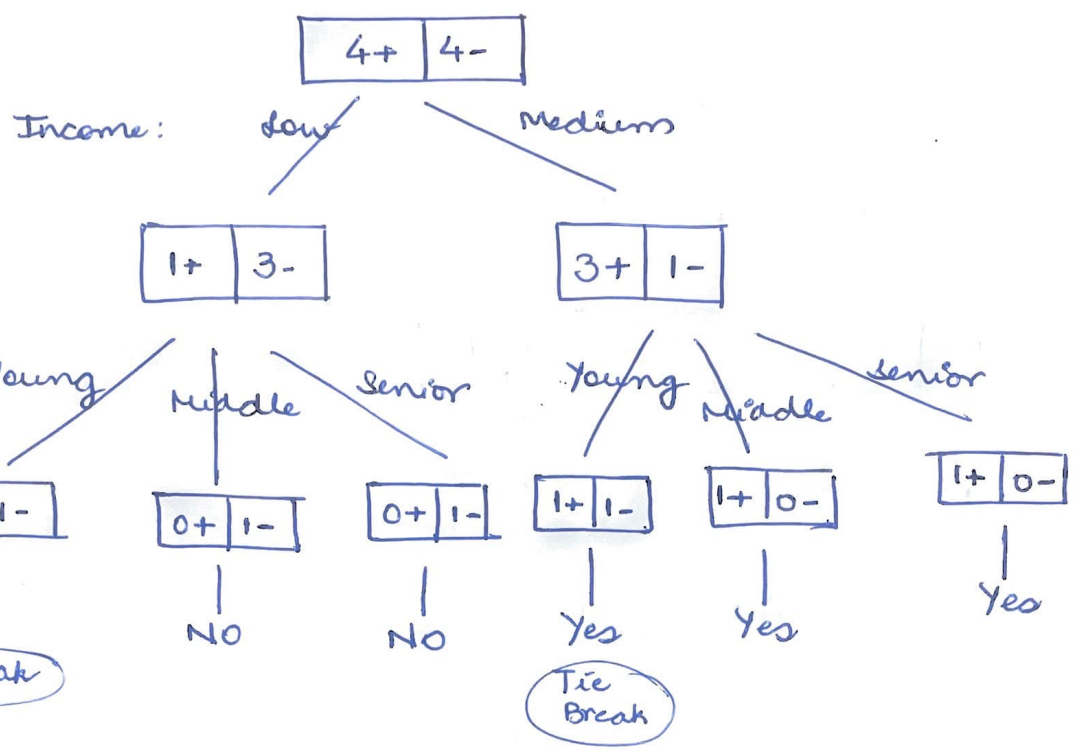
| Customer | Gender | PD |
|----------|--------|-----|
| 2        | F      | Yes |
| 8        | F      | NO  |

AG: Middle Aged
Income: Medium    } → Yes.
Gender: Male

**b)**

| Customer | Actual | Predicted |   |
|----------|--------|-----------|---|
| 1        | NO     | Yes       | ✗ |
| 2        | Yes    | Yes       | ✓ |
| 3        | NO     | No        | ✓ |
| 4        | Yes    | Yes       | ✓ |
| 5        | NO     | NO        | ✓ |
| 6        | Yes    | Yes       | ✓ |
| 7        | Yes    | Yes       | ✓ |
| 8        | NO     | Yes       | ✗ |

$$\text{Accuracy} = \frac{6}{8} = 0.75$$

c) Prune the decision tree:

$$\boxed{4+ \mid 4-}$$

Income:     Low       Medium

$$\boxed{1+ \mid 3-} \qquad\qquad \boxed{3+ \mid 1-}$$
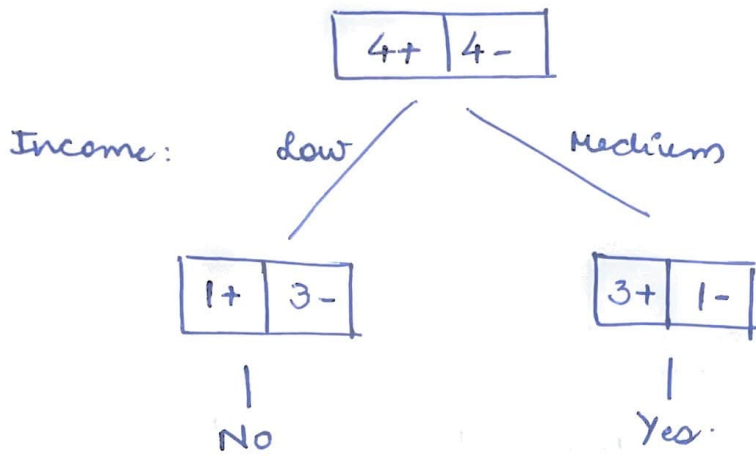
|       | | |
|---|---|---|
| No | | Yes |

| Customer | Actual | Predicted | |
|---|---|---|---|
| 1 | No | No | ✓ |
| 2 | Yes | Yes | ✓ |
| 3 | NO | NO | ✓ |
| 4 | Yes | Yes | ✓ |
| 5 | No | NO | ✓ |
| 6 | Yes | Yes | ✓ |
| 7 | Yes | No | ✗ |
| 8 | No | Yes | ✗ |

$$\text{Accuracy} = \frac{6}{8} = 0.75$$

no change in accuracy.

## Q3:

a)   $P(NO) = P(Yes) = \frac{1}{2}$

**Age Group**

| | Yes | NO |
|---|---|---|
| Young | 2/4 | 2/4 |
| (Middle) | 1/4 | 1/4 |
| Senior | 1/4 | 1/4 |

**Income**

| | Yes | NO |
|---|---|---|
| Low | 1/4 | 3/4 |
| (Medium) | 3/4 | 1/4 |

**Gender**

| | Yes | NO |
|---|---|---|
| (Male) | 1/4 | 2/4 |
| Female | 3/4 | 2/4 |

$$P(NO|x) = \frac{1}{2} \times \frac{1}{4} \times \frac{1}{4} \times \frac{2}{4} = \frac{1}{64} = 0.0156$$

$$P(Yes|x) = \frac{1}{2} \times \frac{1}{4} \times \frac{3}{4} \times \frac{1}{4} = \frac{3}{128} = 0.023 \checkmark$$

$$P(Yes|x) > P(NO|x)$$

4) Naive Bayes - make independent choices

Decision Tree - considers only one feature at a time

Prediction by both classifiers is same.

## Q4:

a) Recall is important.

TP = 120

FN = 10

FP $\cancel{TN}$ = 30

Precision = $\dfrac{TP}{TP + FP}$ = $\dfrac{120}{120 + 30}$ = $\dfrac{120}{150}$ = $\dfrac{4}{5}$ = 0.8

Recall = $\dfrac{TP}{TP + FN}$ = $\dfrac{120}{120 + 10}$ = $\dfrac{120}{130}$ = $\dfrac{12}{13}$ = 0.923

b) Precision is important

TP = 50

FP = 5

FN = 20

Precision = $\dfrac{TP}{TP + FP}$ = $\dfrac{50}{55}$ = 0.909

Recall = $\dfrac{TP}{TP + FN}$ = $\dfrac{50}{70}$ = 0.714