Problem 1:

Step1:

Gender	M	F
C0	8	4
C1	2	6

$Gini(M) = 1 - (\frac{8}{10})^2 - (\frac{2}{10})^2 = 0.32$

$Gini(F) = 1 - (\frac{4}{10})^2 - (\frac{6}{10})^2 = 0.48$

$Gini(Gender) = \frac{10}{20} \times 0.32 + \frac{10}{20} \times 0.48 = 0.4$

Car type	{Family}	{Sports, Luxury}
C0	7	5
C1	0	8

$Gini(\{Family\}) = 1 - (\frac{7}{7})^2 - (\frac{0}{7})^2 = 0$

$Gini(\{Sports, Luxury\}) = 1 - (\frac{5}{13})^2 - (\frac{8}{13})^2 = 0.473$

$Gini(Car type) = \frac{7}{20} \times 0 + \frac{13}{20} \times 0.473 = 0.307$

Car type	{Sports}	{Family, Luxury}
C0	4	8
C1	2	6

$Gini(\{Sports\}) = 1 - (\frac{4}{6})^2 - (\frac{2}{6})^2 = 0.444$

$Gini(\{Family, Luxury\}) = 1 - (\frac{8}{14})^2 - (\frac{6}{14})^2 = 0.490$

$Gini(Car type) = \frac{6}{20} \times 0.444 + \frac{14}{20} \times 0.490 = 0.476$

Car type	{Luxury}	{Family, Sports}
C0	1	11
C1	6	2

$Gini(\{Luxury\}) = 1 - (\frac{1}{7})^2 - (\frac{6}{7})^2 = 0.245$

$Gini(\{Family, Sports\}) = 1 - (\frac{11}{13})^2 - (\frac{2}{13})^2 = 0.260$

$Gini(Car type) = \frac{7}{20} \times 0.245 + \frac{13}{20} \times 0.260$
$= 0.255$

Shirt Size	{Small}	{Medium, Large, Extra Large}
C0	3	9
C1	2	6

$Gini(\{Small\}) = 1 - (\frac{3}{5})^2 - (\frac{2}{5})^2 = 0.48$

$Gini(\{Medium, Large, Extra Large\})$
$= 1 - (\frac{9}{15})^2 - (\frac{6}{15})^2 = 0.48$

$Gini(Shirt Size) = \frac{5}{20} \times 0.48 + \frac{15}{20} \times 0.48$
$= 0.48$

Shirt Size	{Medium}	{Small, Large, Extra Large}
C0	3	9
C1	4	4

$Gini(\{Medium\}) = 1 - (\frac{3}{7})^2 - (\frac{4}{7})^2 = 0.490$

$Gini(\{Small, Large, Extra Large\})$
$= 1 - (\frac{9}{13})^2 - (\frac{4}{13})^2 = 0.426$

$Gini(Shirt Size) = \frac{7}{20} \times 0.490 + \frac{13}{20} \times 0.426$
$= 0.448$

Shirt Size	{Large}	{Small, Medium, Extra Large}
C0	3	9
C1	1	7

$Gini(\{Large\}) = 1 - (\frac{3}{4})^2 - (\frac{1}{4})^2 = 0.375$

$Gini(\{Small, Medium, Extra Large\})$
$= 1 - (\frac{9}{16})^2 - (\frac{7}{16})^2 = 0.492$

$Gini(Shirt Size) = \frac{4}{20} \times 0.375 + \frac{11}{20} \times 0.492$
$= 0.469$

Shirt Size	{Extra Large}	{Small, Medium, Large}
C0	2	10
C1	1	7

$Gini(\{Extra Large\}) = 1 - (\frac{2}{3})^2 - (\frac{1}{3})^2 = 0.444$

$Gini(\{Small, Medium, Large\}) = 1 - (\frac{10}{17})^2 - (\frac{7}{17})^2$
$= 0.484$

$Gini(Shirt Size) = \frac{3}{20} \times 0.444 + \frac{17}{20} \times 0.484$
$= 0.478$

Shirt Size	{Small, Medium}	{Large, Extra Large}
C0	6	6
C1	6	2

$Gini(\{Small, Medium\})$
$= 1 - (\frac{6}{12})^2 - (\frac{6}{12})^2 = 0.5$

$Gini(\{Large, Extra Large\})$
$= 1 - (\frac{6}{8})^2 - (\frac{2}{8})^2 = 0.375$

$Gini(Shirt Size)$
$= \frac{12}{20} \times 0.5 + \frac{8}{20} \times 0.375$
$= 0.45$

Shirt Size	{Small, Large}	{Medium, Extra Large}
C0	6	6
C1	3	5

$Gini(\{Small, Large\})$
$= 1 - (\frac{6}{9})^2 - (\frac{3}{9})^2 = 0.444$

$Gini(\{Medium, Extra Large\})$
$= 1 - (\frac{6}{11})^2 - (\frac{5}{11})^2 = 0.496$

$Gini(Shirt Size) = \frac{9}{20} \times 0.444$
$+ \frac{11}{20} \times 0.496$
$= 0.473$

Shirt Size	{Small, Extra Large}	{Medium, Large}
C0	6	6
C1	3	5

$Gini(\{Small, Extra Large\})$
$= 1 - (\frac{6}{9})^2 - (\frac{3}{9})^2 = 0.444$

$Gini(\{Medium, Large\})$
$= 1 - (\frac{6}{11})^2 - (\frac{5}{11})^2 = 0.496$

$Gini(Shirt Size) = \frac{9}{20} \times 0.444$
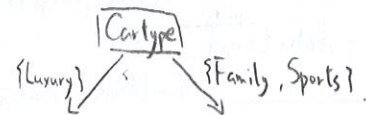$+ \frac{11}{20} \times 0.496$
$= 0.473$

∴ pick Car type {Luxury}, {Family, Sports} for step 1.

```
        |Car type|
  {Luxury}    {Family, Sports}
```

Step 2:

(1) for {Luxury} part:

Customer ID	Gender	Shirt Size	Class
10	F	Large	C0
14	M	Extra Large	C1
15	F	Small	C1
16	F	Small	C1
18	F	Medium	C1
19	F	Medium	C1
20	F	Large	C1

Gender	M	F
C0	0	1
C1	1	5

$Gini(M) = 1 - (\frac{0}{1})^2 - (\frac{1}{1})^2 = 0$

$Gini(F) = 1 - (\frac{1}{6})^2 - (\frac{5}{6})^2 = 0.278$

$Gini(Gender) = \frac{1}{7} \times 0 + \frac{6}{7} \times 0.278 = 0.238$

Shirt Size	{Small}	{Medium, Large, Extra Large}
C0	0	1
C1	2	4

$Gini(\{Small\}) = 1 - (\frac{0}{2})^2 - (\frac{2}{2})^2 = 0$

$Gini(\{Medium, Large, Extra Large\})$
$= 1 - (\frac{1}{5})^2 - (\frac{4}{5})^2 = 0.32$

$Gini(Shirt Size) = \frac{2}{7} \times 0 + \frac{5}{7} \times 0.32$
$= 0.229$

Left column

Shirt Size

	{Medium}	{Small, Large, Extra Large}
Co	0	1
C1	2	4

$Gini(\{Medium\}) = 1-(\frac{0}{2})^2-(\frac{2}{2})^2 = 0$

$Gini(\{Small, Large, Extra\ Large\})$
$= 1-(\frac{1}{5})^2-(\frac{4}{5})^2 = 0.32$

$Gini(Shirt\ Size) = \frac{2}{7}\times 0 + \frac{5}{7}\times 0.32$
$= 0.229$

Shirt Size

	{Large}	{Small, Medium, Extra Large}
Co	1	0
C1	1	5

$Gini(\{Large\}) = 1-(\frac{1}{2})^2-(\frac{1}{2})^2 = 0.5$

$Gini(\{Small, Medium, Extra\ Large\})$
$= 1-(\frac{0}{5})^2-(\frac{5}{5})^2 = 0$

$Gini(Shirt\ Size) = \frac{2}{7}\times 0.5 + \frac{5}{7}\times 0$
$= 0.143$

Shirt Size

	{Extra Large}	{Small, Medium, Large}
Co	0	1
C1	1	5

$Gini(\{Extra\ Large\}) = 1-(\frac{0}{1})^2-(\frac{1}{1})^2 = 0$

$Gini(\{Small, Medium, Large\})$
$= 1-(\frac{1}{6})^2-(\frac{5}{6})^2 = 0.278$

$Gini(Shirt\ Size)$
$= \frac{1}{7}\times 0 + \frac{6}{7}\times 0.278$
$= 0.238$

Shirt Size

	{Small, Medium}	{Large, Extra Large}
Co	0	1
C1	4	2

$Gini(\{Small, Medium\}) = 1-0^2-(\frac{4}{4})^2 = 0$

$Gini(\{Large, Extra\ Large\}) = 1-(\frac{1}{3})^2-(\frac{2}{3})^2$
$= 0.444$

$Gini(Shirt\ Size) = \frac{4}{7}\times 0 + \frac{3}{7}\times 0.444$
$= 0.190$

Shirt Size

	{Small, Large}	{Medium, Extra Large}
Co	1	0
C1	3	3

$Gini(\{Small, Large\}) = 1-(\frac{1}{4})^2-(\frac{3}{4})^2$
$= 0.375$

$Gini(\{Medium, Extra\ Large\}) = 1-0^2-(\frac{3}{3})^2 = 0$

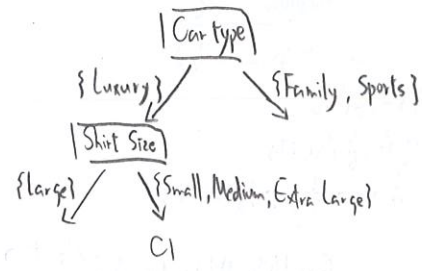$Gini(Shirt\ Size) = \frac{4}{7}\times 0.375 + \frac{3}{7}\times 0$
$= 0.214$

Shirt Size

	{Small, Extra Large}	{Medium, Large}
Co	0	1
C1	3	3

$Gini(\{Small, Extra\ Large\}) = 1-0-(\frac{3}{3})^2 = 0$

$Gini(\{Medium, Large\}) = 1-(\frac{1}{4})^2-(\frac{3}{4})^2 = 0.375$

$Gini(Shirt\ Size) = \frac{3}{7}\times 0 + \frac{4}{7}\times 0.375$
$= 0.214$

∴ pick Shirt Size {Large} {Small, Medium, Extra Large}
for step 2 {luxury} part

```
            |Car type|
  {Luxury}/         \{Family, Sports}
  |Shirt Size|
{Large}/   \{Small, Medium, Extra Large}
        C1
```

Right column

2) for {Family, Sports} part:

Customer ID	Gender	Shirt Size	Class
1	M	Small	Co
2	M	Medium	Co
3	M	Medium	Co
4	M	Large	Co
5	M	Extra Large	Co
6	M	Extra Large	Co
7	F	Small	Co
8	F	Small	Co
9	F	Medium	Co
11	M	Large	Co
12	M	Extra Large	Co
13	M	Medium	C1
17	F	Medium	C1

Gender

	M	F
Co	8	3
C1	1	1

$Gini(M) = 1-(\frac{8}{9})^2-(\frac{1}{9})^2 = 0.198$

$Gini(F) = 1-(\frac{3}{4})^2-(\frac{1}{4})^2 = 0.375$

$Gini(Gender) = \frac{9}{13}\times 0.198 + \frac{4}{13}\times 0.375 = 0.252$

Shirt Size

	{Small}	{Medium, Large, Extra Large}
Co	3	8
C1	0	2

$Gini(\{Small\}) = 1-(\frac{3}{3})^2-(\frac{0}{3})^2 = 0$

$Gini(\{Medium, Large, Extra\ Large\})$
$= 1-(\frac{8}{10})^2-(\frac{2}{10})^2 = 0.32$

$Gini(Shirt\ Size) = \frac{3}{13}\times 0 + \frac{10}{13}\times 0.32$
$= 0.246$

Shirt Size

	{Medium}	{Small, Large, Extra Large}
Co	3	8
C1	2	0

$Gini(\{Medium\}) = 1-(\frac{3}{5})^2-(\frac{2}{5})^2 = 0.48$

$Gini(\{Small, Large, Extra\ Large\})$
$= 1-(\frac{8}{8})^2-(\frac{0}{8})^2 = 0$

$Gini(Shirt\ Size) = \frac{5}{13}\times 0.48 + \frac{8}{13}\times 0$
$= 0.185$

Shirt Size

	{Large}	{Small, Medium, Extra Large}
Co	2	9
C1	0	2

$Gini(\{Large\}) = 1-(\frac{2}{2})^2-(\frac{0}{2})^2 = 0$

$Gini(\{Small, Medium, Extra\ Large\})$
$= 1-(\frac{9}{11})^2-(\frac{2}{11})^2 = 0.298$

$Gini(Shirt\ Size) = \frac{2}{13}\times 0 + \frac{11}{13}\times 0.298$
$= 0.252$

Shirt Size

	{Extra Large}	{Small, Medium, Large}
Co	3	8
C1	0	2

$Gini(\{Extra\ Large\}) = 1-(\frac{3}{3})^2-(\frac{0}{3})^2 = 0$

$Gini(\{Small, Medium, Large\})$
$= 1-(\frac{8}{10})^2-(\frac{2}{10})^2 = 0.32$

$Gini(Shirt\ Size) = \frac{3}{13}\times 0 + \frac{10}{13}\times 0.32$
$= 0.246$

Shirt Size

	{Small, Medium}	{Large, Extra Large}
Co	6	5
C1	2	0

$Gini(\{Small, Medium\}) = 1-(\frac{6}{8})^2-(\frac{2}{8})^2$
$= 0.375$

$Gini(\{Large, Extra\ Large\}) = 1-(\frac{5}{5})^2-(\frac{0}{5})^2$
$= 0$

$Gini(Shirt\ Size) = \frac{8}{13}\times 0.375 + \frac{5}{13}\times 0$
$= 0.231$

Shirt Size

	{Small, Large}	{Medium, Extra large}
Co	5	6
C1	0	2

$Gini(\{Small, Large\}) = 1 - (\frac{5}{5})^2 - (\frac{0}{5})^2 = 0$

$Gini(\{Medium, Extra\ large\}) = 1 - (\frac{6}{8})^2 - (\frac{2}{8})^2$
$= 0.375$

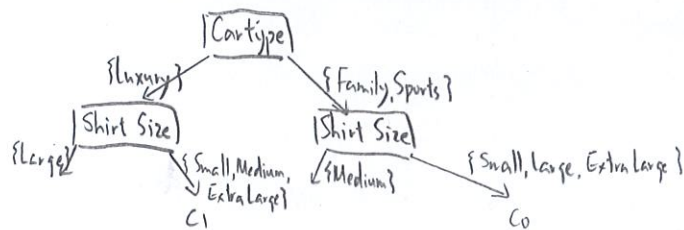$Gini(Shirt\ Size) = \frac{5}{13} \times 0 + \frac{8}{13} \times 0.375$
$= 0.231$

Shirt Size

	{Small, Extra Large}	{Medium, Large}
Co	6	5
C1	0	2

$Gini(\{Small, Extra\ Large\}) = 1 - (\frac{6}{6})^2 - (\frac{0}{6})^2 = 0$

$Gini(\{Medium, Large\}) = 1 - (\frac{5}{7})^2 - (\frac{2}{7})^2 = 0.408$

$Gini(Shirt\ Size) = \frac{6}{13} \times 0 + \frac{7}{13} \times 0.408$
$= 0.220$

∴ pick Shirt Size {Medium}, {Small, Large, Extra Large}
for step 2 {Family, Sports} part

[Tree diagram: Cartype → {Luxury}: Shirt Size → {Large}, {Small, Medium, Extra Large}: C1; {Family, Sports}: Shirt Size → {Medium}, {Small, Large, Extra Large}: Co]

Step 3:
(1) for {Luxury}, {Large} part

Customer ID	Gender	Class
10	F	Co
20	F	C1

Can't split!
∵ Same features, different class.

(2) for {Family, Sports}, {Medium} part

Customer ID	Gender	Cartype	Class
2	M	Sports	Co
3	M	Family	Co
9	F	Family	Co
13	M	Sports	C1
17	F	Sports	C1

Gender

	M	F
Co	2	1
C1	1	1

$Gini(M) = 1 - (\frac{2}{3})^2 - (\frac{1}{3})^2 = 0.444$
$Gini(F) = 1 - (\frac{1}{2})^2 - (\frac{1}{2})^2 = 0.5$
$Gini(Gender) = \frac{3}{5} \times 0.444 + \frac{2}{5} \times 0.5 = 0.466$

Cartype

	Sports	Family
Co	1	2
C1	2	0

$Gini(Sports) = 1 - (\frac{1}{3})^2 - (\frac{2}{3})^2 = 0.444$
$Gini(Family) = 1 - (\frac{2}{2})^2 - (\frac{0}{2})^2 = 0$
$Gini(Cartype) = \frac{3}{5} \times 0.444 + \frac{2}{5} \times 0 = 0.267$

∴ pick Cartype {Family}, {Sports}
for step 3 {Family, Sports}, {Medium} part:

[Tree diagram: Cartype → {Luxury}: Shirt Size → {Large}, {Small, Medium, Extra Large}: C1; {Family, Sports}: Shirt Size → {Medium}: Cartype → {Family}, {Sports}: Co; {Small, Large, Extra Large}: Co]
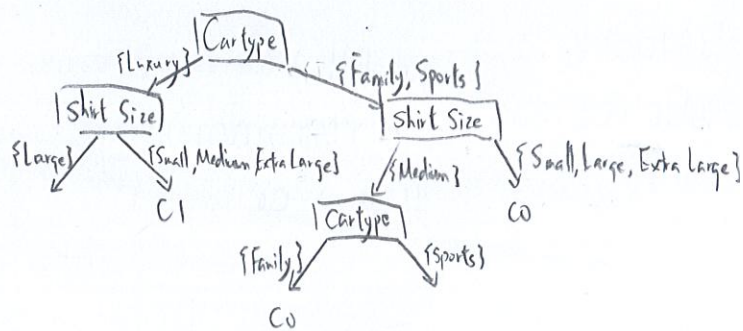
Step 4: for {Family, Sports} {Medium} {Sports} part:

Customer ID	Gender	Class
2	M	Co
13	M	C1
17	F	C1

Gender

	M	F
Co	1	0
C1	1	1

$Gini(M) = 1 - (\frac{1}{2})^2 - (\frac{1}{2})^2 = 0.5$
$Gini(F) = 1 - (\frac{1}{1})^2 - (\frac{0}{1})^2 = 0$
$Gini(Gender) = \frac{2}{3} \times 0.5 + \frac{1}{3} \times 0 = 0.333$

∴ pick Gender for step 4:

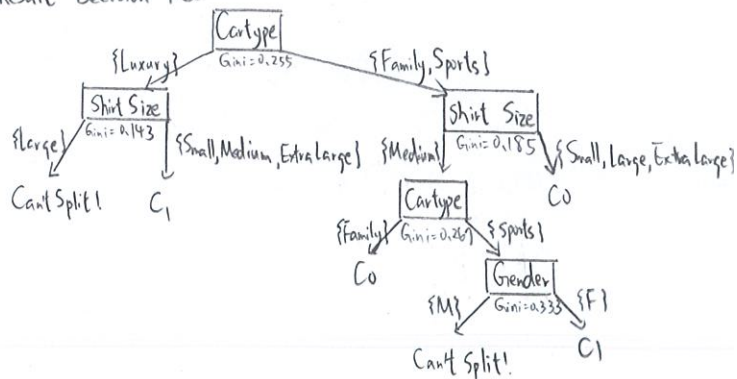[Tree diagram: Cartype → {Luxury}: Shirt Size → {Large}: C1, {Small, Medium, Extra Large}; {Family, Sports}: Shirt Size → {Medium}: Cartype → {Family}: Co, {Sports}: Gender → {M}, {F}: C1; {Small, Large, Extra Large}: Co]
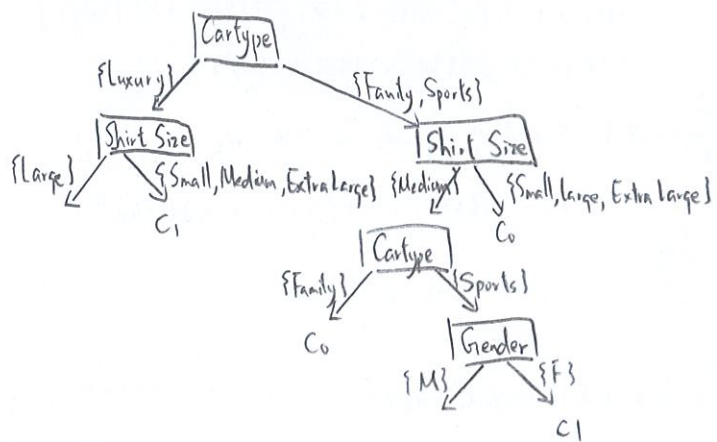
Step 5: for {Family, Sports}, {Medium}, {Sports}, {M} part:

Customer ID	Class
2	Co
13	C1

Can't Split!
∵ Same features, different class.

∴ Result Decision Tree:

[Tree diagram: Cartype (Gini=0.255) → {Luxury}: Shirt Size (Gini=0.143) → {Large}: Can't Split!, {Small, Medium, Extra Large}: C1; {Family, Sports}: Shirt Size (Gini=0.185) → {Medium}: Cartype (Gini=0.267) → {Family}: Co, {Sports}: Gender (Gini=0.333) → {M}: Can't Split!, {F}: C1; {Small, Large, Extra Large}: Co]

Problem 2:

Gender	Car type	Shirt Size	
M	Sports	Medium	A: Attributes.

$P(A|C_0) = \frac{8}{12} \times \frac{4}{12} \times \frac{3}{12} = 0.056$ $P(A|C_0) \cdot P(C_0) = 0.056 \times \frac{12}{20} = 0.0336$

$P(A|C_1) = \frac{2}{8} \times \frac{2}{8} \times \frac{4}{8} = 0.03125$ $P(A|C_1) \times P(C_1) = 0.03125 \times \frac{8}{20} = 0.0125$

$\therefore P(A|C_0) P(C_0) > P(A|C_1) P(C_1) \quad \therefore C_0$

Problem 3

1. objectives: $\min L(w) = \frac{\|\vec{w}\|^2}{2}$

   Constraints: $f(\vec{x}_i) = \begin{cases} 1 & \text{if } \vec{w} \cdot \vec{x}_0 + b \geq 1 \\ -1 & \text{if } \vec{w} \cdot \vec{x}_0 + b \leq -1 \end{cases}$

2. pick support vectors: $\begin{cases} y: 1 \ (4,1) \\ y: -1 \ (2,3) \end{cases}$

3. $\phi(\vec{\lambda}) = \sum_{i=1}^{2} \alpha_i - \sum_{i=1, j=1}^{2} \alpha_i \alpha_j y_i y_j (\vec{x}_i \cdot \vec{x}_j)$

   $= (\alpha_1 + \alpha_3) - \frac{1}{2} [\alpha_i^2 y_i^2 (\vec{x}_1 \cdot \vec{x}_1) + 2\alpha_1 \alpha_2 y_1 y_2 (\vec{x}_1 \cdot \vec{x}_2) + \alpha_2^2 y_2^2 (\vec{x}_2 \cdot \vec{x}_2)]$

   $= (\alpha_1 + \alpha_2) - \frac{1}{2} [\alpha_1^2 (16+1) + 2 \times \alpha_1 \alpha_2 (-1)(8+3) + \alpha_2^2 (4+9)]$

   $= (\alpha_1 + \alpha_2) - \frac{1}{2} (17\alpha_1^2 - 22\alpha_1 \alpha_2 + 13\alpha_2^2)$

   $\sum_{i=1}^{2} (\alpha_i y_i) = \alpha_1 y_1 + \alpha_2 y_2 = \alpha_1 - \alpha_2 = 0 \Rightarrow \alpha_1 = \alpha_2$

   $\Rightarrow 2\alpha_1 - \frac{1}{2}(17\alpha_1^2 - 22\alpha_1^2 + 13\alpha_1^2) = 2\alpha_1 - 4\alpha_1^2$

   $\frac{d\phi}{d\alpha_1} = 2 - 8\alpha_1 = 0 \Rightarrow \alpha_1 = \frac{1}{4}, \ \alpha_2 = \frac{1}{4}$

4. $\vec{w} = \sum_{i=1}^{N} \alpha_i y_i \vec{x}_i = \alpha_1 y_1 \vec{x}_1 + \alpha_2 y_2 \vec{x}_2 = \frac{1}{4}(4,1) + \frac{1}{4}(-1)(2,3) = [1, 0.25] - [0.5, 0.75]$

   $= [0.5, -0.5]$

   $b = \frac{1}{2} \left(\min_{i:y_i=1} (\vec{w} \cdot \vec{x}_i) + \max_{i:y_i=-1} (\vec{w} \cdot \vec{x}_i) \right)$

   $= \frac{1}{2}((\vec{w} \cdot \vec{x}_1) + (\vec{w} \cdot \vec{x}_2)) = \frac{1}{2}(0.5 \times 4 - 0.5 \times 1 + 0.5 \times 2 - 0.5 \times 3) = \frac{1}{2}$

   $\therefore$ hyperplane: $y = \begin{bmatrix} 0.5 \\ -0.5 \end{bmatrix} x + \frac{1}{2}$.