

<i>Likes Popcorn</i>	<i>Age</i>	<i>Favorite Color</i>	<i>Loves Troll 2</i>
<i>Yes</i>	<i>12</i>	<i>Blue</i>	<i>Yes</i>
<i>Yes</i>	<i>87</i>	<i>Green</i>	<i>Yes</i>
<i>No</i>	<i>44</i>	<i>Blue</i>	<i>No</i>
<i>Yes</i>	<i>19</i>	<i>Red</i>	<i>No</i>
<i>No</i>	<i>32</i>	<i>Green</i>	<i>Yes</i>
<i>No</i>	<i>14</i>	<i>Blue</i>	<i>Yes</i>

Convert Categorical Data  
To Numerics

<i>Likes Popcorn</i>	<i>Age</i>	<i>Favorite Color</i>	<i>Loves Troll 2</i>
1	12	1	1
1	87	2	1
0	44	1	0
1	19	3	0
0	32	2	1
0	14	1	1

Likes Popcorn	Age	Favorite Color	Loves Troll 2
1	12	1	1
1	87	2	1
0	44	1	0
1	19	3	0
0	32	2	1
0	14	1	1

درخت یس بونی اولیه

0.5

← تفسیر: احتمال Yes بودن Label

Likes Popcorn	Age	Favorite Color	Loves Troll 2
1	12	1	1
1	87	2	1
0	44	1	0
1	19	3	0
0	32	2	1
0	14	1	1

Residual  
-0.5  
-0.5  
+0.5  
+0.5  
-0.5  
-0.5

درخت یس بی اولیه

0.5

حال Residual ها را برای هر داده  
محاسبه می کنیم  $(\hat{y} - y)$

Likes Popcorn	Age	Favorite Color	Loves Troll 2
1	12	1	1
1	87	2	1
0	44	1	0
1	19	3	0
0	32	2	1
0	14	1	1

Residual  
 -0.5  
 -0.5  
 +0.5  
 +0.5  
 -0.5  
 -0.5

حالت مقادیر Residual را در یک بزرگ مقدار

می دهیم و Similarity Score را حساب می کنیم

$$\left( \sum_{i=1}^6 \text{Residual}_i \right)$$

$$\sum_{i=1}^6 [\text{Prev. Prob.}_i \times (1 - \text{Prev. Prob.}_i)]$$

درخت بیس بینی اولیه

0.5

-0.5, -0.5, +0.5, +0.5, -0.5, -0.5

sim. score

$$\frac{(-0.5 - 0.5 + 0.5 + 0.5 - 0.5 - 0.5)^2}{6 \times 0.5 \times (1 - 0.5)} = 0.66$$



Likes Popcorn	Age	Favorite Color	Loves Troll 2
1	12	1	1
1	87	2	1
0	44	1	0
1	19	3	0
0	32	2	1
0	14	1	1

Residual  
-0.5  
-0.5  
+0.5  
+0.5  
-0.5  
-0.5

حال باید Split های مختلف را بررسی کرده و به سمت بهترین gain برویم

0.5 ← Like Popcorn

65.5, 38, 25.5, 16.5 و 13 ← Age

2.5 و 1.5 ← Fav. Color

Split های ممکن

Sim. Score  
0.66

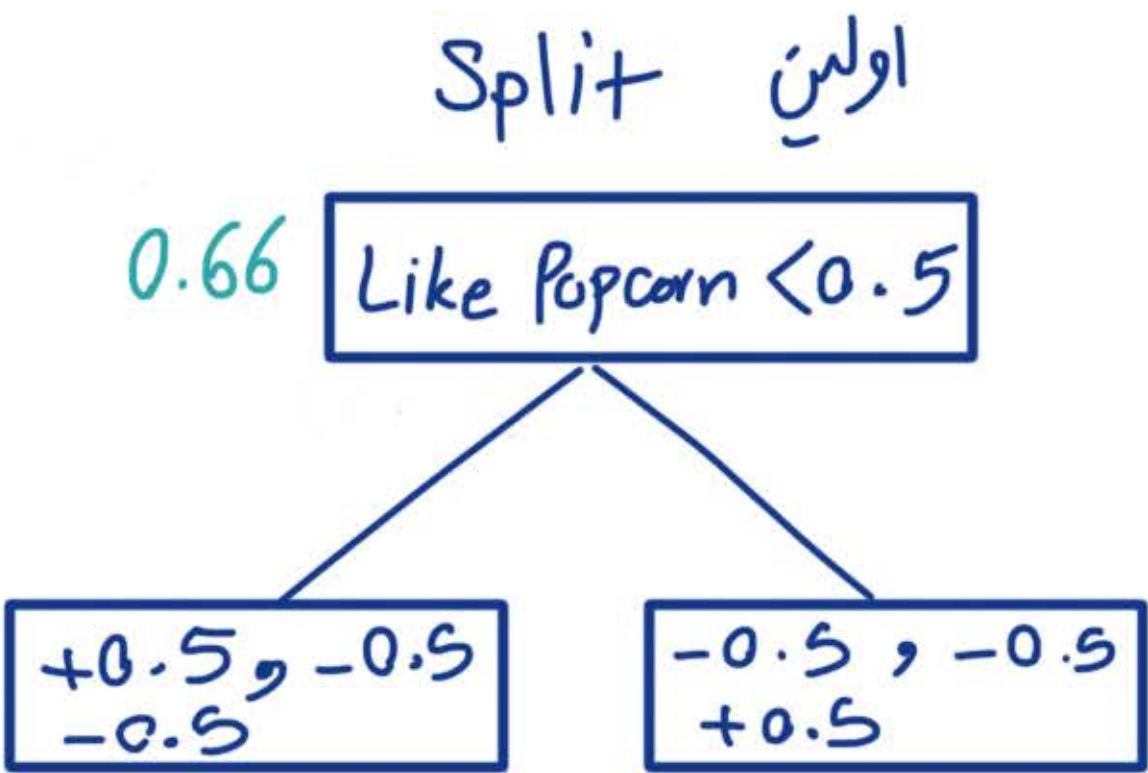
-0.5, -0.5, +0.5, +0.5, -0.5, -0.5

درخت یس بی اولیه

0.5

Likes Popcorn	Age	Favorite Color	Loves Troll 2
1	12	1	1
1	87	2	1
0	44	1	0
1	19	3	0
0	32	2	1
0	14	1	1

Residual  
-0.5  
-0.5  
+0.5  
+0.5  
-0.5  
-0.5



درخت یس بانی اولی

0.5

Likes Popcorn	Age	Favorite Color	Loves Troll 2
1	12	1	1
1	87	2	1
0	44	1	0
1	19	3	0
0	32	2	1
0	14	1	1

Residual  
 -0.5  
 -0.5  
 +0.5  
 +0.5  
 -0.5  
 -0.5

اولین Split

0.66

Like Popcorn < 0.5

+0.5 , -0.5  
 -0.5

-0.5 , -0.5  
 +0.5

درخت یسین بنی اولیه

0.5

similarity score =

$$\frac{(\cancel{0.6} - \cancel{0.6} - 0.5)^2}{3 \times 0.25} = 0.33$$

Likes Popcorn	Age	Favorite Color	Loves Troll 2
1	12	1	1
1	87	2	1
0	44	1	0
1	19	3	0
0	32	2	1
0	14	1	1

Residual  
 -0.5  
 -0.5  
 +0.5  
 +0.5  
 -0.5  
 -0.5

اولین Split

0.66

Like Popcorn < 0.5

+0.5, -0.5  
 -0.5

0.33

-0.5, -0.5  
 +0.5

درخت یسین بنی اولیه

0.5

similarity score =

$$\frac{(\cancel{0.6} - \cancel{0.6} - 0.5)^2}{3 \times 0.25} = 0.33$$

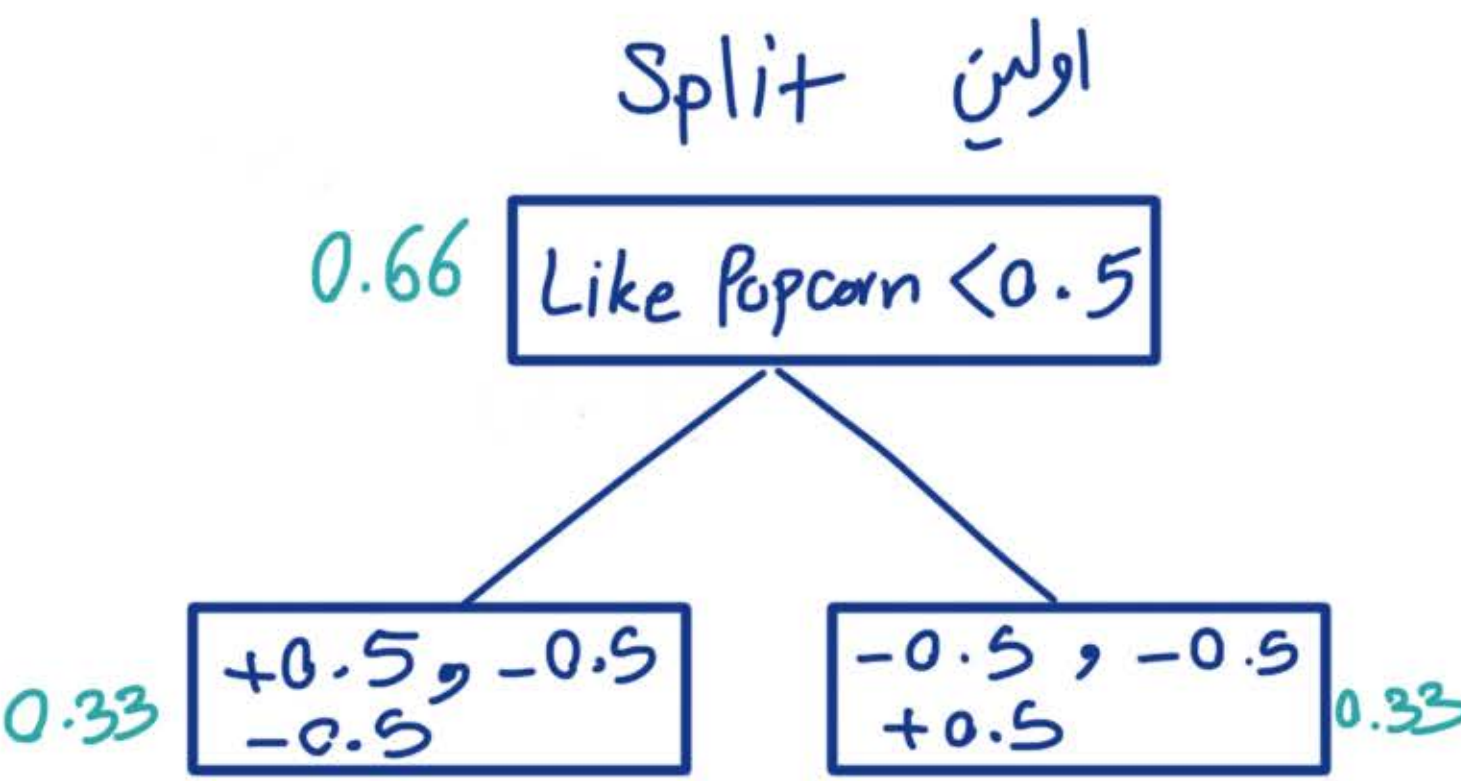


Likes Popcorn	Age	Favorite Color	Loves Troll 2
1	12	1	1
1	87	2	1
0	44	1	0
1	19	3	0
0	32	2	1
0	14	1	1

Residual  
-0.5  
-0.5  
+0.5  
+0.5  
-0.5  
-0.5

درخت یسین بنی اولیه

0.5



Gain:  $\text{Similarity}_{\text{Left}} + \text{Similarity}_{\text{Right}} - \text{Similarity}_{\text{Root}}$

= 0.33 + 0.33 - 0.66 = 0

Likes Popcorn	Age	Favorite Color	Loves Troll 2
1	12	1	1
1	87	2	1
0	44	1	0
1	19	3	0
0	32	2	1
0	14	1	1

Residual  
-0.5  
-0.5  
+0.5  
+0.5  
-0.5  
-0.5

دومین Split

0.66

Age < 13

-0.5

-0.5, +0.5,  
+0.5, -0.5, -0.5

درخت یسین بنی اولیه

0.5

$$\text{similarity score} = \frac{(-0.5)^2}{0.25} = 1$$

Likes Popcorn	Age	Favorite Color	Loves Troll 2
1	12	1	1
1	87	2	1
0	44	1	0
1	19	3	0
0	32	2	1
0	14	1	1

Residual  
-0.5  
-0.5  
+0.5  
+0.5  
-0.5  
-0.5

دومین Split

0.66

Age < 13

-0.5

-0.5, +0.5,  
+0.5, -0.5, -0.5

درخت یسین بنی اولیه

0.5

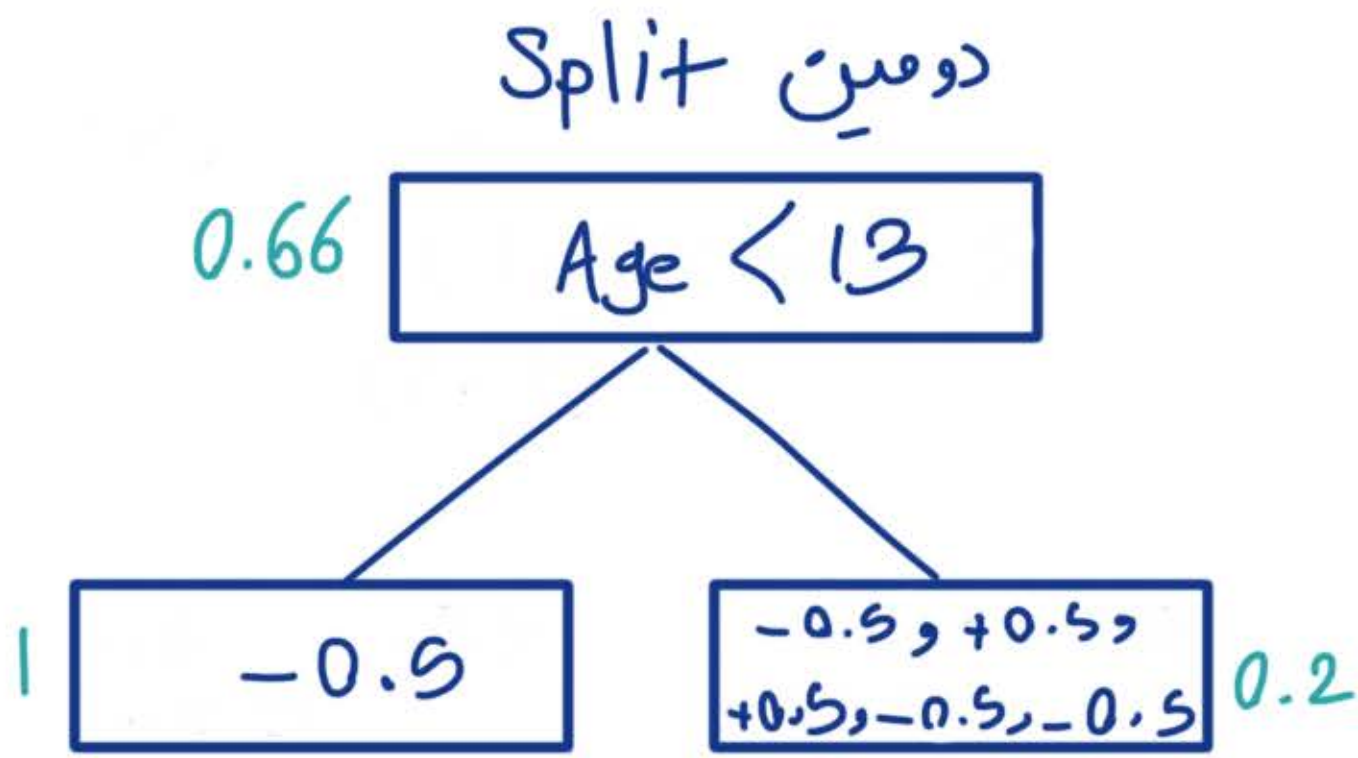
$$\text{similarity score} = \frac{(-0.5 + \cancel{+0.5} + \cancel{+0.5} - \cancel{0.5} - \cancel{0.5})^2}{5 \times 0.25} = 0.2$$

Likes Popcorn	Age	Favorite Color	Loves Troll 2
1	12	1	1
1	87	2	1
0	44	1	0
1	19	3	0
0	32	2	1
0	14	1	1

Residual  
-0.5  
-0.5  
+0.5  
+0.5  
-0.5  
-0.5

درخت یس بی اولیه

0.5



$$\text{Gain} = 1 + 0.2 - 0.66$$

$$= 0.54$$



Likes Popcorn	Age	Favorite Color	Loves Troll 2
1	12	1	1
1	87	2	1
0	44	1	0
1	19	3	0
0	32	2	1
0	14	1	1

Residual  
-0.5  
-0.5  
+0.5  
+0.5  
-0.5  
-0.5

Split سوچین

0.66

Age < 16.5

-0.5 و -0.5

-0.5 و +0.5 و  
+0.5 و -0.5

درخت یس بنی اولیہ

0.5

$$\text{similarity score} = \frac{(-0.5 - 0.5)^2}{2 \times 0.25} = 2$$

Likes Popcorn	Age	Favorite Color	Loves Troll 2
1	12	1	1
1	87	2	1
0	44	1	0
1	19	3	0
0	32	2	1
0	14	1	1

Residual  
-0.5  
-0.5  
+0.5  
+0.5  
-0.5  
-0.5

Split سوچین

0.66

Age < 16.5

2

-0.5, -0.5

-0.5, +0.5, +0.5, -0.5

درخت یس بنی اولیہ

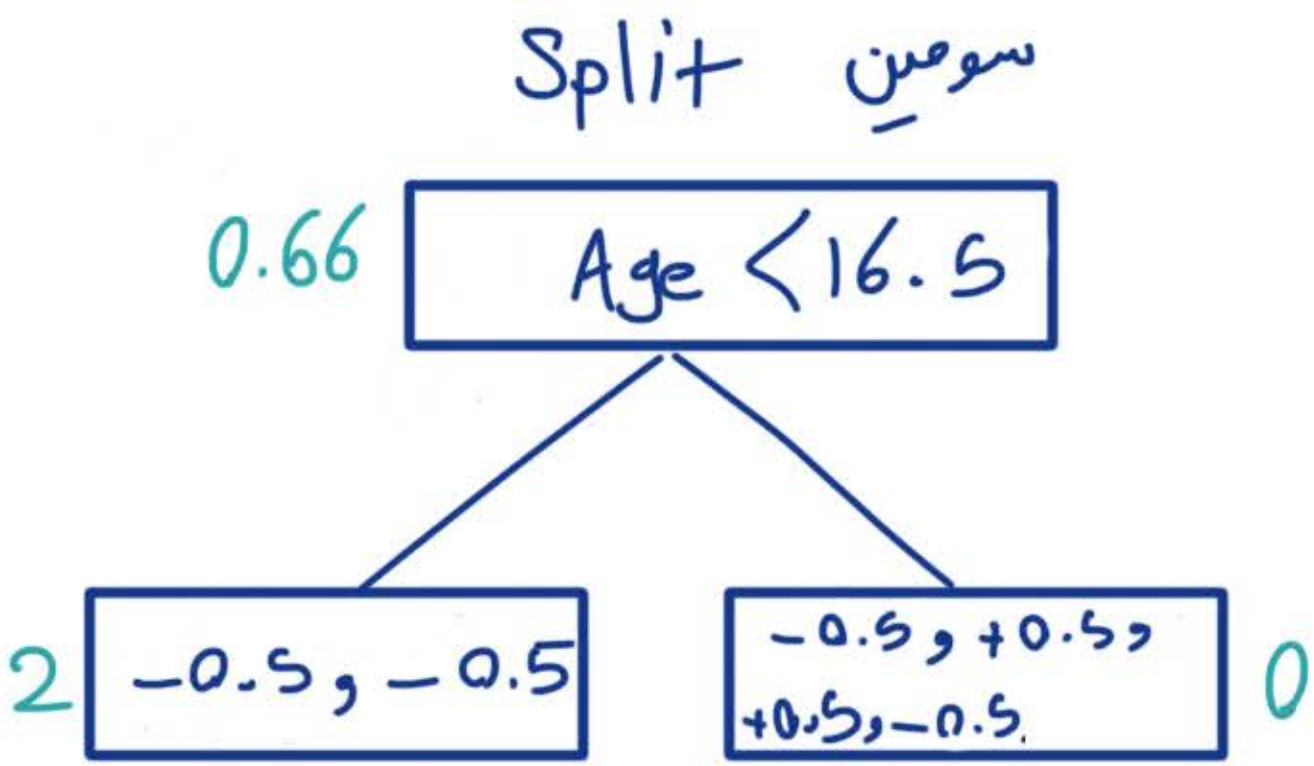
0.5

$$\text{similarity score} = \frac{(-0.5 + 0.5 + 0.5 - 0.5)^2}{4 \times 0.25} = 0$$

Likes Popcorn	Age	Favorite Color	Loves Troll 2
1	12	1	1
1	87	2	1
0	44	1	0
1	19	3	0
0	32	2	1
0	14	1	1

Residual  
-0.5  
-0.5  
+0.5  
+0.5  
-0.5  
-0.5

درخت یس بی اولیه  
0.5



Gain = 2 + 0 - 0.66 = 1.34

Likes Popcorn	Age	Favorite Color	Loves Troll 2
1	12	1	1
1	87	2	1
0	44	1	0
1	19	3	0
0	32	2	1
0	14	1	1

Residual  
-0.5  
-0.5  
+0.5  
+0.5  
-0.5  
-0.5

Split چھار میں

0.66

Age < 25.5

-0.5, +0.5, -0.5

-0.5, +0.5, -0.5

درخت پیس بنی اولیہ

0.5



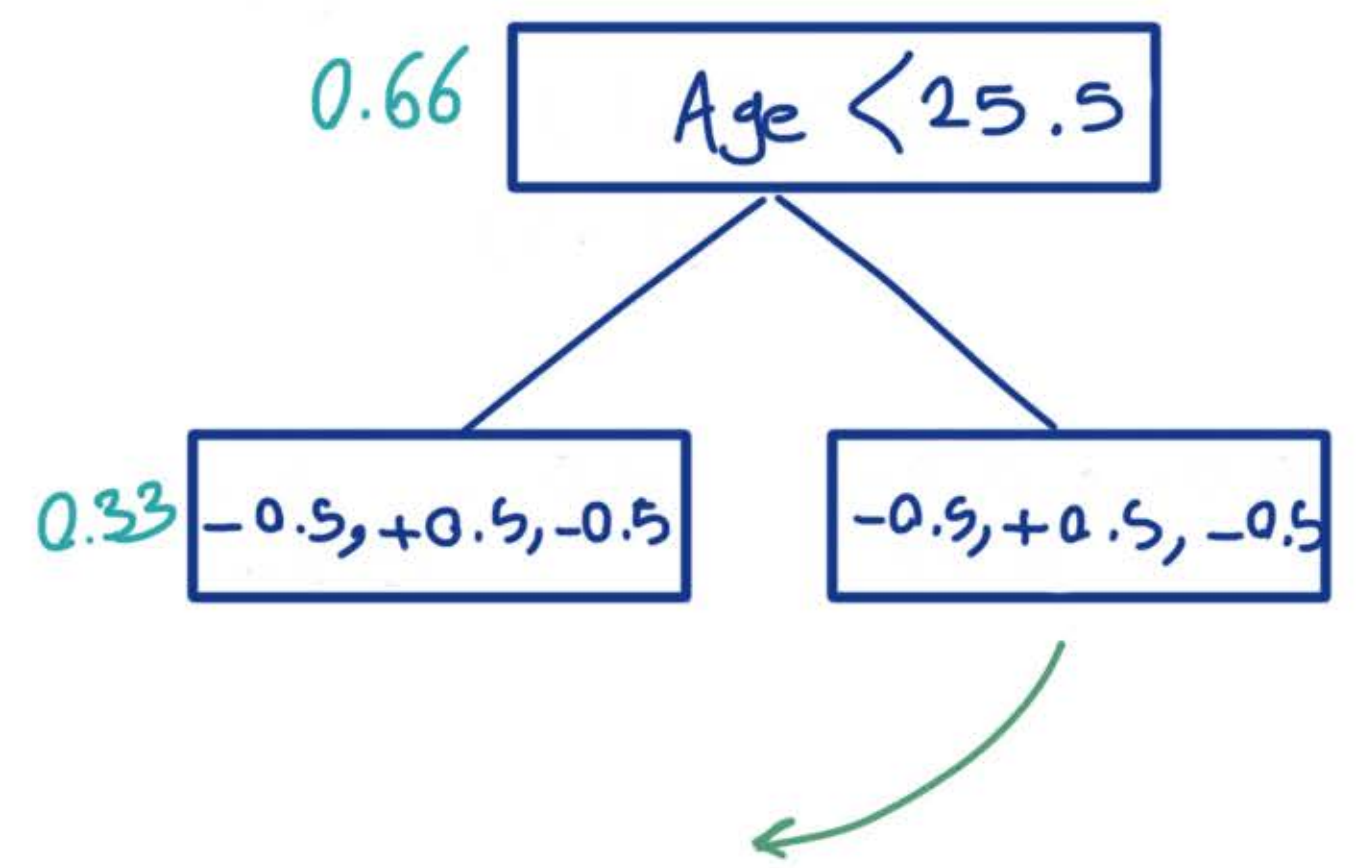
$$\text{Sim score} = \frac{(-0.5 + 0.5 - 0.5)^2}{3 \times 0.25} = 0.33$$



Likes Popcorn	Age	Favorite Color	Loves Troll 2
1	12	1	1
1	87	2	1
0	44	1	0
1	19	3	0
0	32	2	1
0	14	1	1

Residual  
 -0.5  
 -0.5  
 +0.5  
 +0.5  
 -0.5  
 -0.5

Split چھار میں



درخت پیس بینی اولیہ

0.5

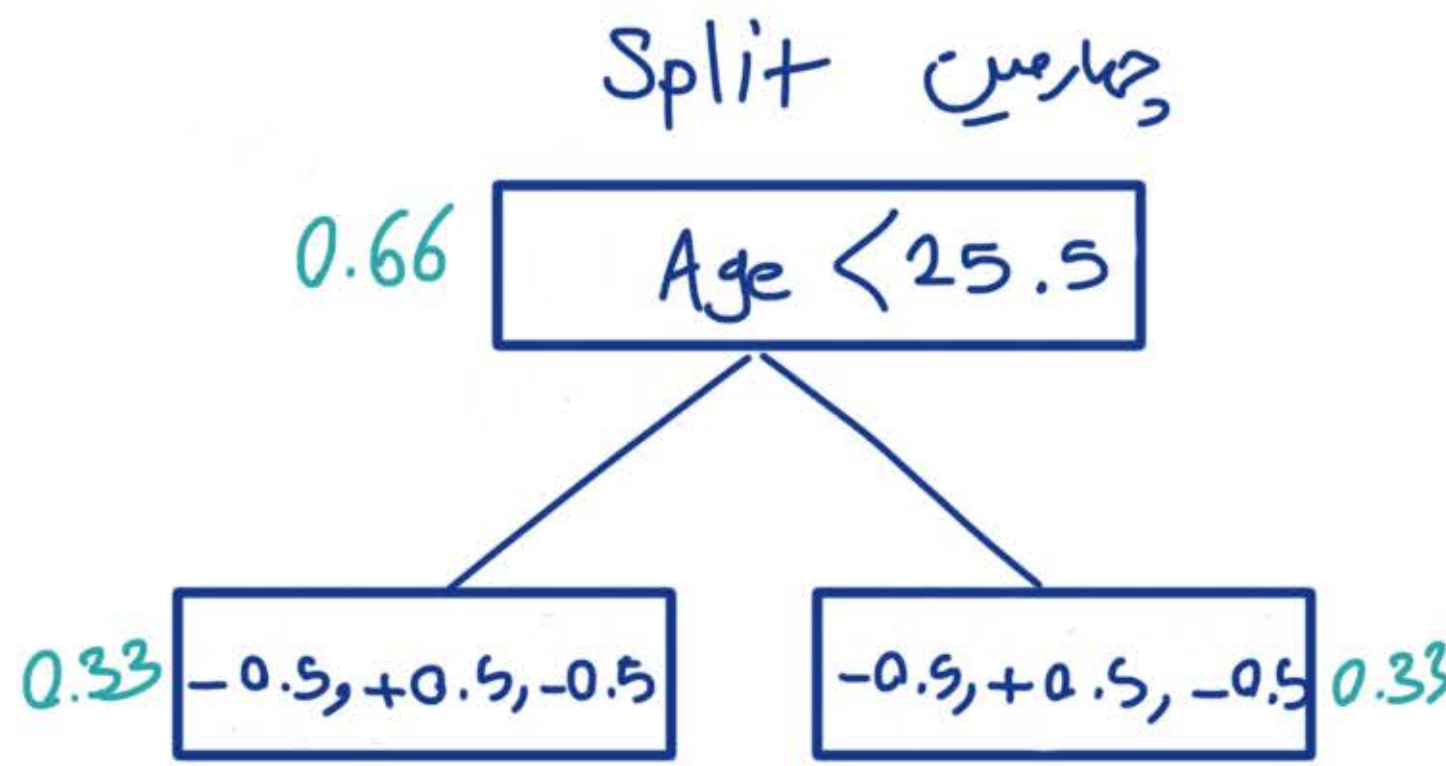
$$\text{sim score} = \frac{(-0.5 + 0.5 - 0.5)^2}{3 \times 0.25} = 0.33$$

Likes Popcorn	Age	Favorite Color	Loves Troll 2
1	12	1	1
1	87	2	1
0	44	1	0
1	19	3	0
0	32	2	1
0	14	1	1

Residual  
-0.5  
-0.5  
+0.5  
+0.5  
-0.5  
-0.5

درخت یس بی اولیه

0.5



Gain = 0.33 + 0.33 - 0.66 = 0

Likes Popcorn	Age	Favorite Color	Loves Troll 2
1	12	1	1
1	87	2	1
0	44	1	0
1	19	3	0
0	32	2	1
0	14	1	1

Residual  
-0.5  
-0.5  
+0.5  
+0.5  
-0.5  
-0.5

Split بچمین

0.66

Age < 38

-0.5, +0.5  
-0.5, -0.5

-0.5, +0.5

درخت یسین بینی اولیه

0.5

$$\text{Sim score} = \frac{(-0.5 + \cancel{0.5} - \cancel{0.5} - 0.5)^2}{4 \times 0.25} = 1$$

Likes Popcorn	Age	Favorite Color	Loves Troll 2
1	12	1	1
1	87	2	1
0	44	1	0
1	19	3	0
0	32	2	1
0	14	1	1

Residual  
-0.5  
-0.5  
+0.5  
+0.5  
-0.5  
-0.5

Split بچمین

0.66

Age < 38

1  
-0.5, +0.5  
-0.5, -0.5

-0.5, +0.5

درخت یس بینی اولیه

0.5

$$\text{Sim score} = \frac{(-0.5 + 0.5)^2}{2 \times 0.25} = 0$$

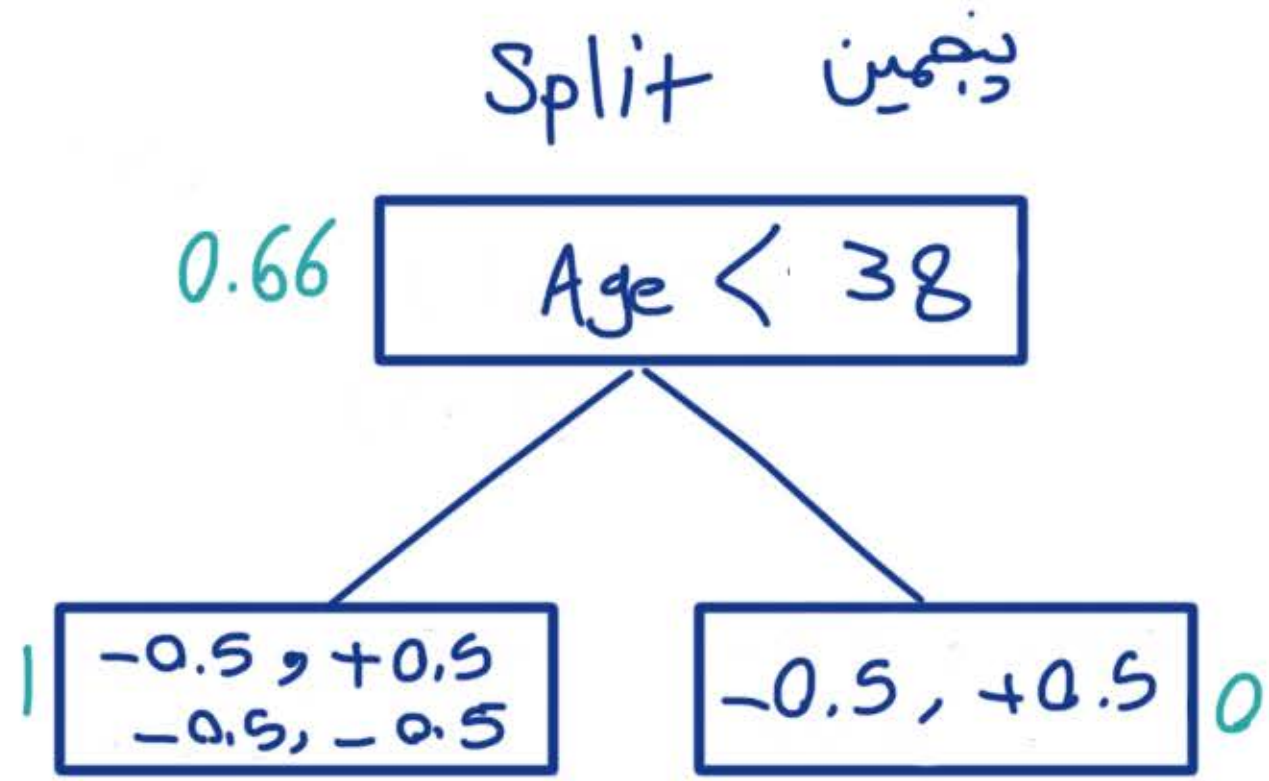


Likes Popcorn	Age	Favorite Color	Loves Troll 2
1	12	1	1
1	87	2	1
0	44	1	0
1	19	3	0
0	32	2	1
0	14	1	1

Residual  
-0.5  
-0.5  
+0.5  
+0.5  
-0.5  
-0.5

درخت یس بی اولیه

0.5



Gain = 1 + 0 - 0.66 = 0.33

Likes Popcorn	Age	Favorite Color	Loves Troll 2
1	12	1	1
1	87	2	1
0	44	1	0
1	19	3	0
0	32	2	1
0	14	1	1

Residual  
 -0.5  
 -0.5  
 +0.5  
 +0.5  
 -0.5  
 -0.5

Split سٹپ

0.66

Age < 65.5

-0.5, +0.5, +0.5  
 , -0.5, -0.5

-0.5

درخت پیس بینی اولیہ

0.5

$$\text{Sim. Score} = \frac{(-0.5 + \cancel{0.5} + \cancel{0.5} - \cancel{0.5} - \cancel{0.5})^2}{5 \times 0.25} = 0.2$$

Likes Popcorn	Age	Favorite Color	Loves Troll 2
1	12	1	1
1	87	2	1
0	44	1	0
1	19	3	0
0	32	2	1
0	14	1	1

Residual  
-0.5  
-0.5  
+0.5  
+0.5  
-0.5  
-0.5

Split سٹپ

0.66

Age < 65.5

-0.5, +0.5, +0.5  
, -0.5, -0.5

0.2

-0.5

درخت پیس بینی اولیہ

0.5

$$\text{Sim. Score} = \frac{(-0.5)^2}{0.25} = 1$$

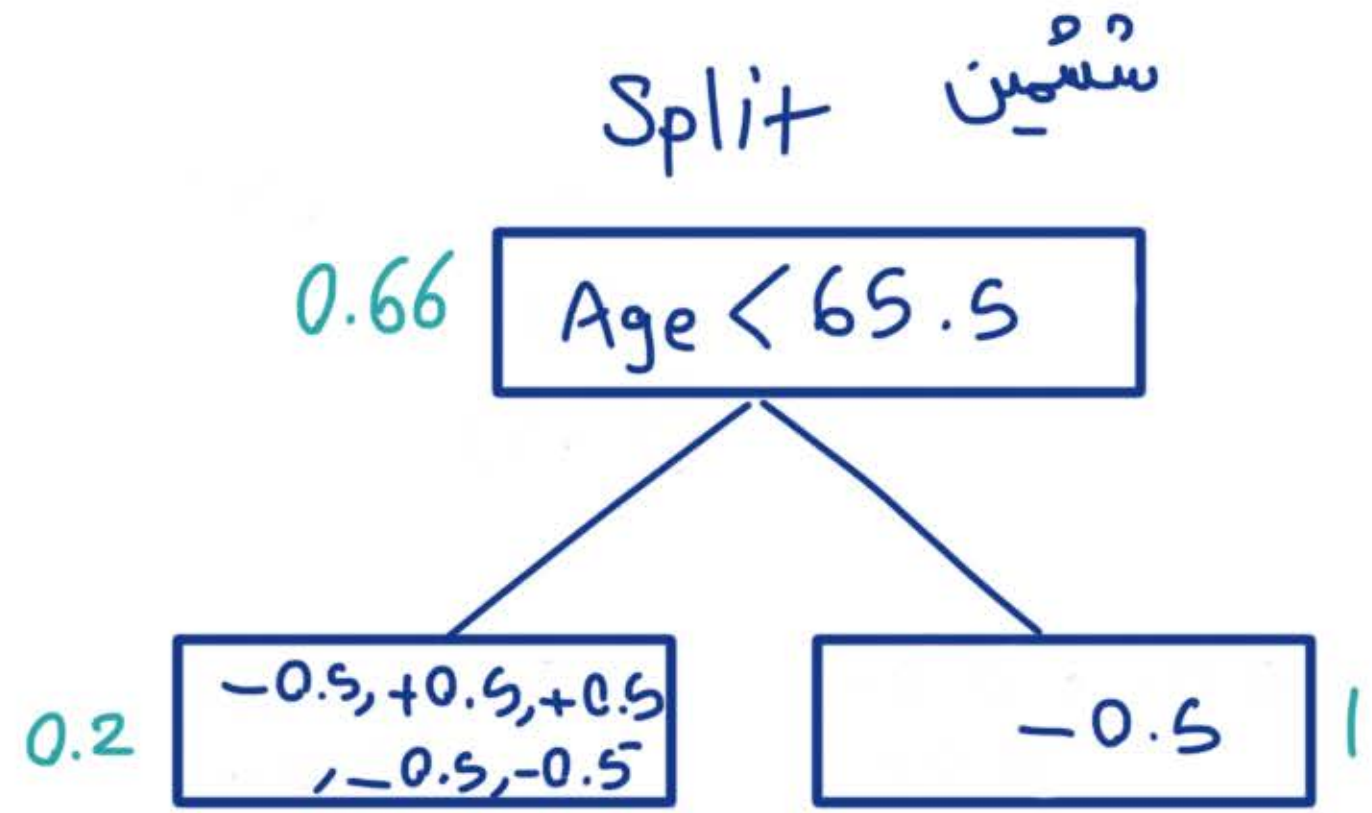


Likes Popcorn	Age	Favorite Color	Loves Troll 2
1	12	1	1
1	87	2	1
0	44	1	0
1	19	3	0
0	32	2	1
0	14	1	1

Residual  
-0.5  
-0.5  
+0.5  
+0.5  
-0.5  
-0.5

درخت یس بی اولیه

0.5



Gain = 0.2 + 1 - 0.66 = 0.54



Likes Popcorn	Age	Favorite Color	Loves Troll 2
1	12	1	1
1	87	2	1
0	44	1	0
1	19	3	0
0	32	2	1
0	14	1	1

Residual  
-0.5  
-0.5  
+0.5  
+0.5  
-0.5  
-0.5

Split هفتمین

0.66

Far. Color < 1.5

-0.5, +0.5, -0.5

-0.5, +0.5, -0.5

درخت ییس بینی اولیه

0.5

$$\text{sim score} = \frac{(-0.5 + \cancel{0.5} - \cancel{0.5})^2}{3 \times 0.25} = 0.33$$

Likes Popcorn	Age	Favorite Color	Loves Troll 2
1	12	1	1
1	87	2	1
0	44	1	0
1	19	3	0
0	32	2	1
0	14	1	1

Residual  
-0.5  
-0.5  
+0.5  
+0.5  
-0.5  
-0.5

Split هفتمین

0.66

Far. Color < 1.5

-0.5, +0.5, -0.5

-0.5, +0.5, -0.5

0.33



$$\text{sim score} = \frac{(-0.5 + \cancel{0.5} - \cancel{0.5})^2}{3 \times 0.25} = 0.33$$

درخت ییس بینی اولیه

0.5

Likes Popcorn	Age	Favorite Color	Loves Troll 2
1	12	1	1
1	87	2	1
0	44	1	0
1	19	3	0
0	32	2	1
0	14	1	1

Residual  
-0.5  
-0.5  
+0.5  
+0.5  
-0.5  
-0.5

Split هفتمین

0.66

Far. Color < 1.5

-0.5, +0.5, -0.5

-0.5, +0.5, -0.5

0.33

0.33

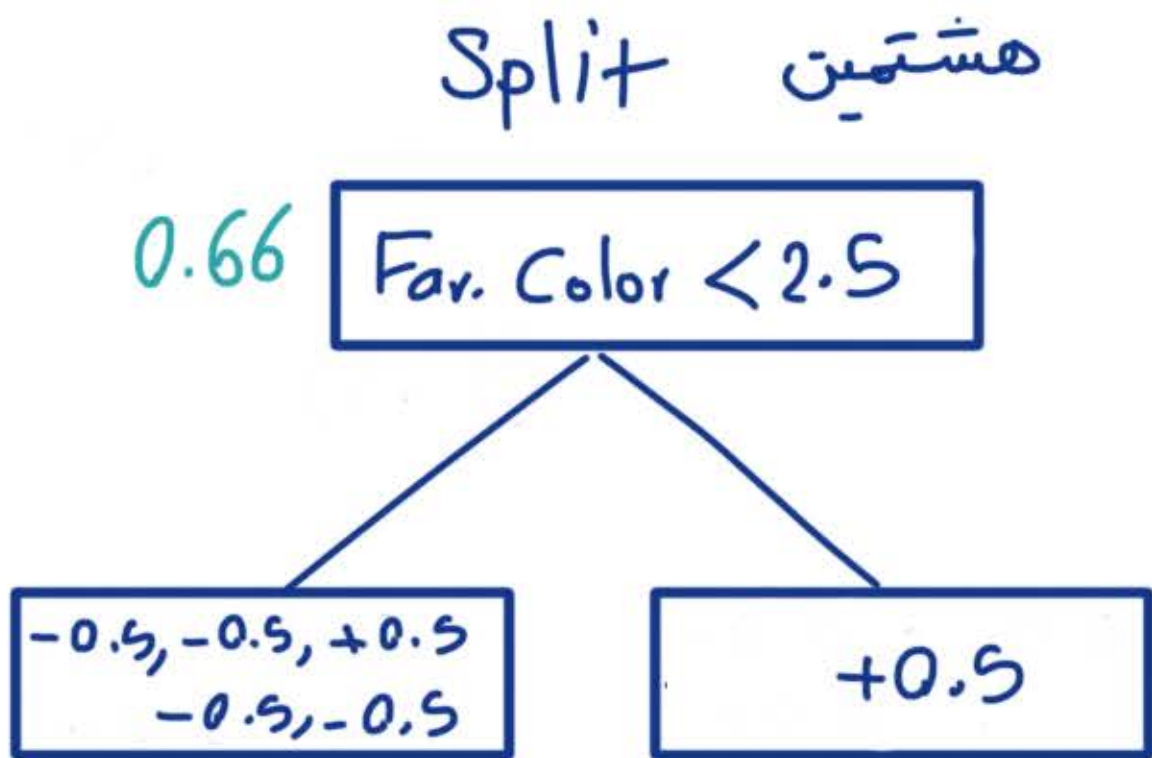
درخت ییس بینی اولیه

0.5

Gain = 0.33 + 0.33 - 0.66 = 0

Likes Popcorn	Age	Favorite Color	Loves Troll 2
1	12	1	1
1	87	2	1
0	44	1	0
1	19	3	0
0	32	2	1
0	14	1	1

Residual  
 -0.5  
 -0.5  
 +0.5  
 +0.5  
 -0.5  
 -0.5



درخت ییس بینی اولیه

0.5

$$\text{Sim score} = \frac{(-0.5 - 0.5 + 0.5 - 0.5 - 0.5)^2}{5 \times 0.25} = 1.8$$

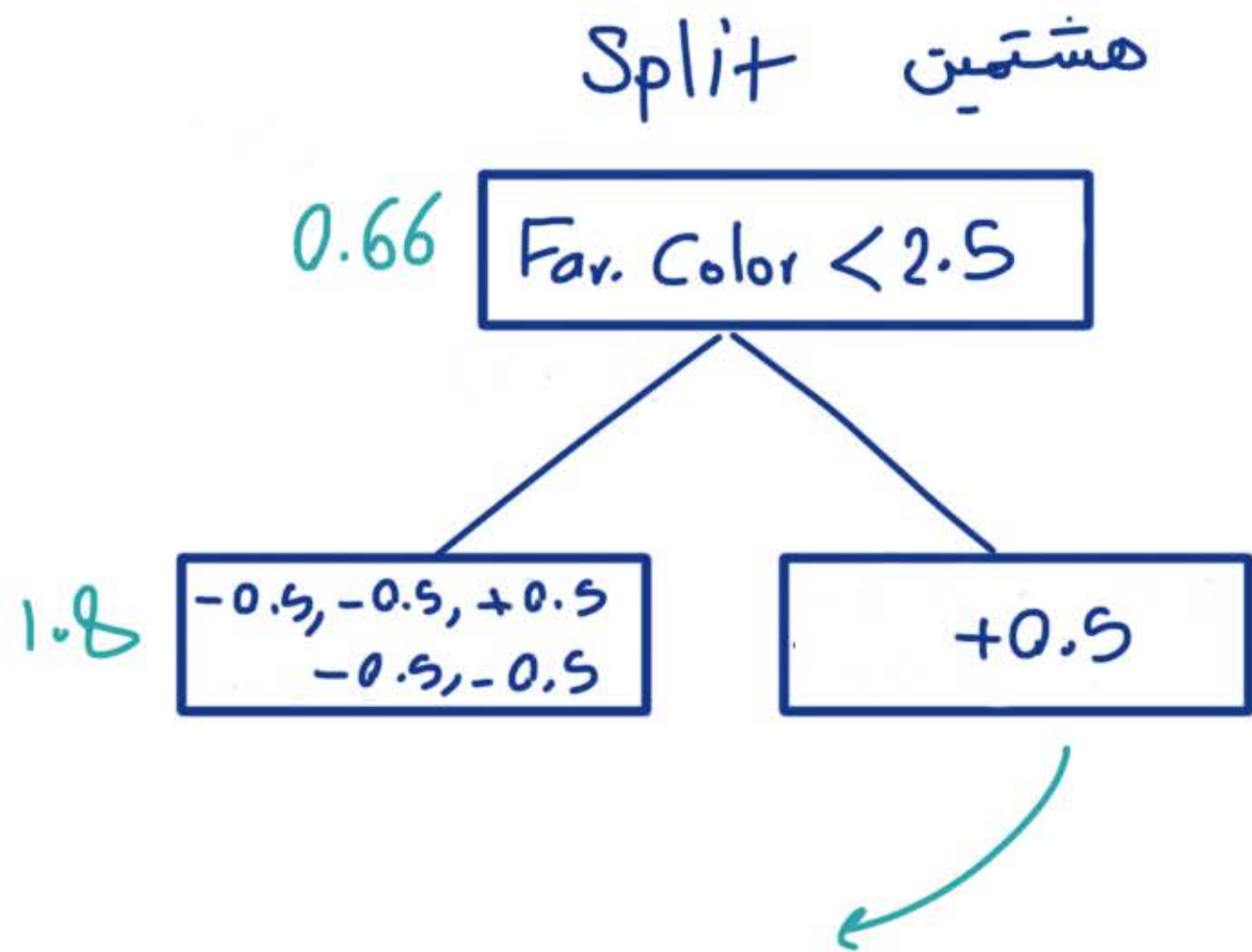


Likes Popcorn	Age	Favorite Color	Loves Troll 2
1	12	1	1
1	87	2	1
0	44	1	0
1	19	3	0
0	32	2	1
0	14	1	1

Residual  
 -0.5  
 -0.5  
 +0.5  
 +0.5  
 -0.5  
 -0.5

درخت یس بی اولیه

0.5



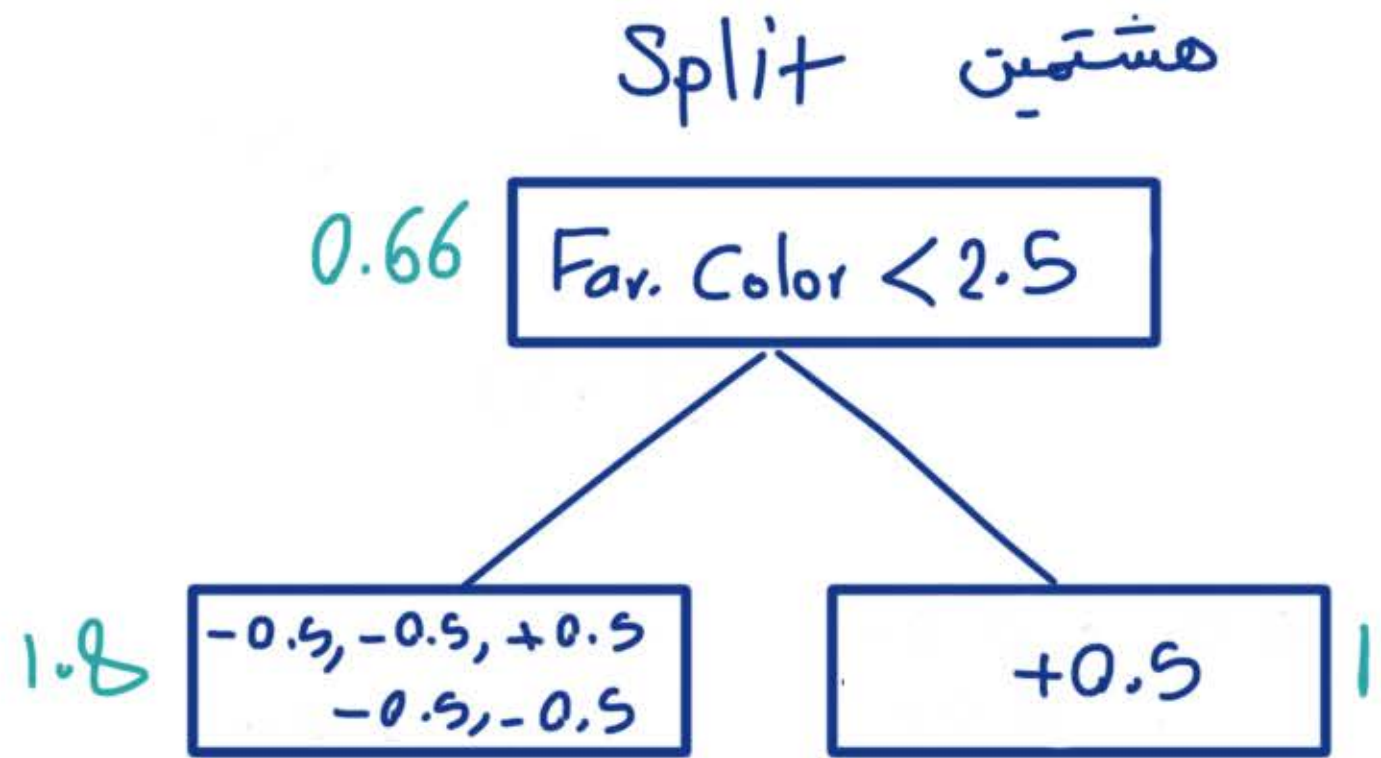
Sim score =  $\frac{(0.5)^2}{0.25} = 1$

Likes Popcorn	Age	Favorite Color	Loves Troll 2
1	12	1	1
1	87	2	1
0	44	1	0
1	19	3	0
0	32	2	1
0	14	1	1

Residual  
-0.5  
-0.5  
+0.5  
+0.5  
-0.5  
-0.5

درخت یس بی اولیه

0.5

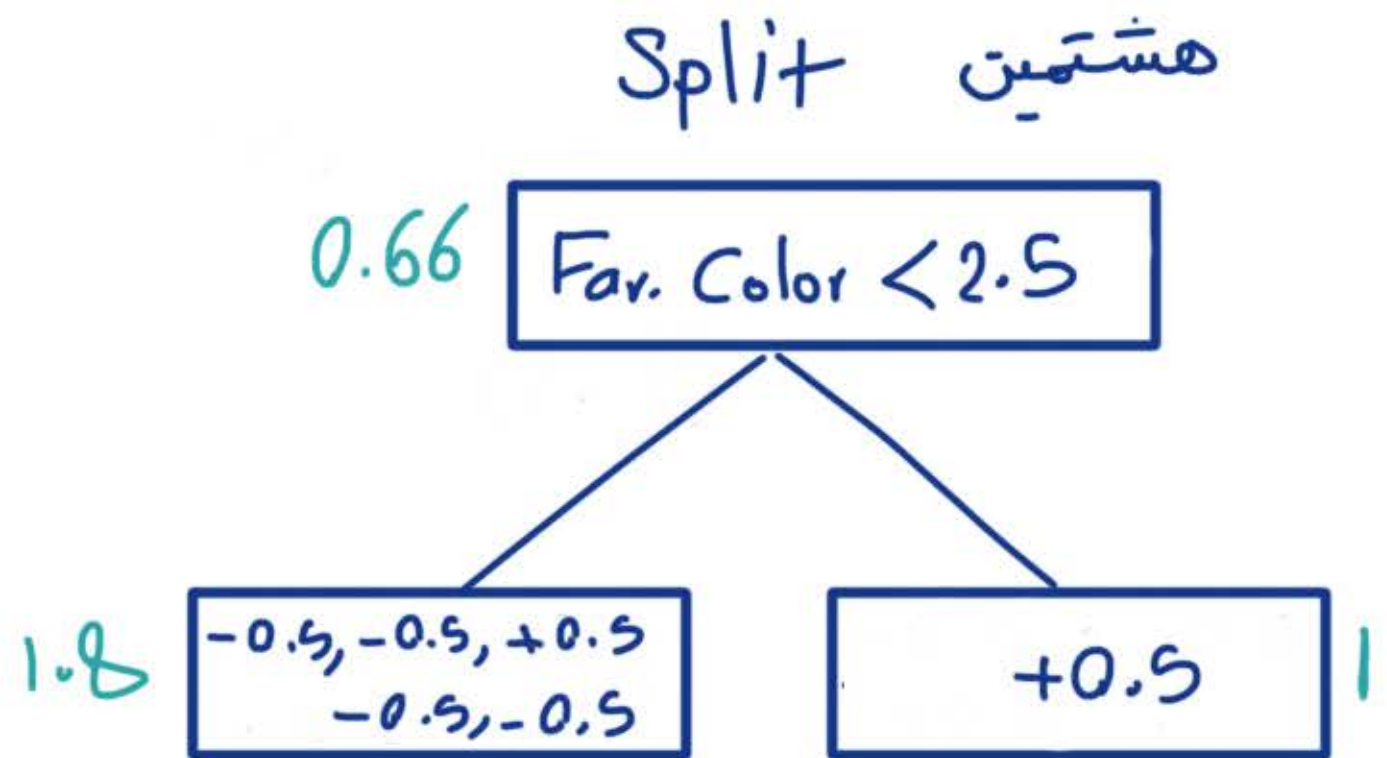


Gain = 1.8 + 1 - 0.66 = 2.14

Likes Popcorn	Age	Favorite Color	Loves Troll 2
1	12	1	1
1	87	2	1
0	44	1	0
1	19	3	0
0	32	2	1
0	14	1	1

Residual  
 -0.5  
 -0.5  
 +0.5  
 +0.5  
 -0.5  
 -0.5

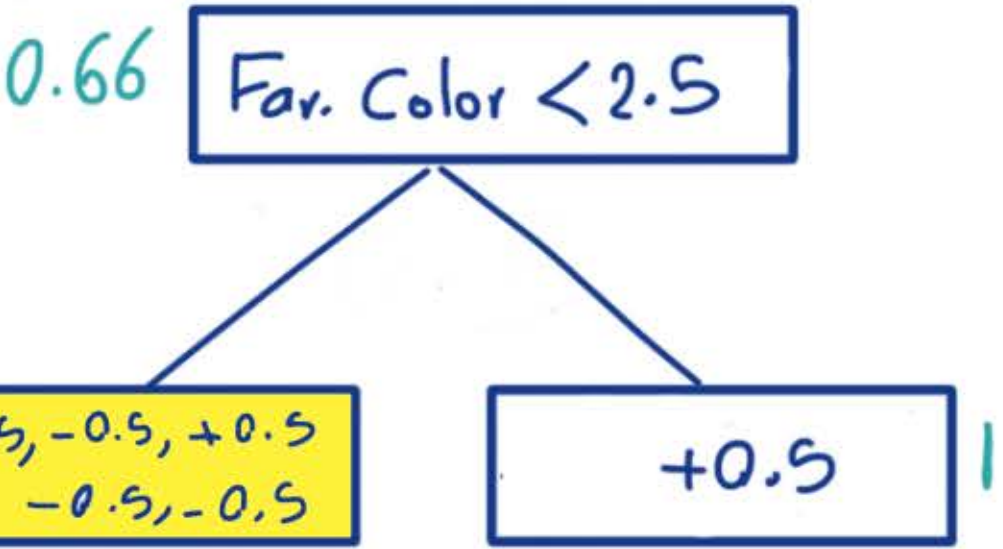
درخت یس بی اولیه  
 0.5



پس این Split بیشترین Gain را نصیب ما می کنه.  
 حال باید به سراغ شاخه سمت چپ رفته و آن را  
 split کنیم.

Likes Popcorn	Age	Favorite Color	Loves Troll 2
1	12	1	1
1	87	2	1
0	44	1	0
1	19	3	0
0	32	2	1
0	14	1	1

Residual  
-0.5  
-0.5  
+0.5  
+0.5  
-0.5  
-0.5



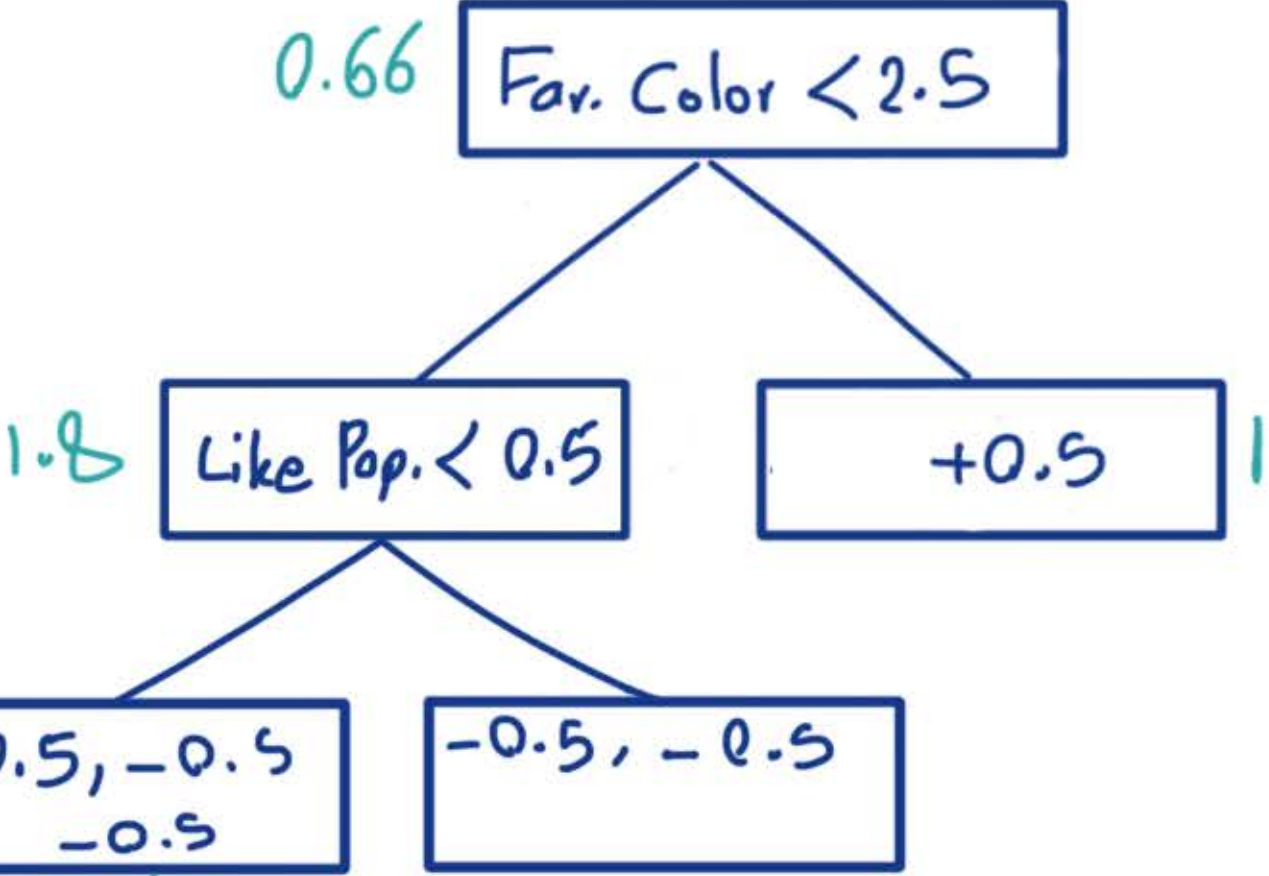
درخت ییس بینی اولیه  
0.5

0.5 ← Like popcorns  
65.5, 38, 23 و 13 ← Age  
1.5 ← Far. color  
Split های ممکن ساخته جیب



Likes Popcorn	Age	Favorite Color	Loves Troll 2
1	12	1	1
1	87	2	1
0	44	1	0
1	19	3	0
0	32	2	1
0	14	1	1

Residual  
-0.5  
-0.5  
+0.5  
+0.5  
-0.5  
-0.5

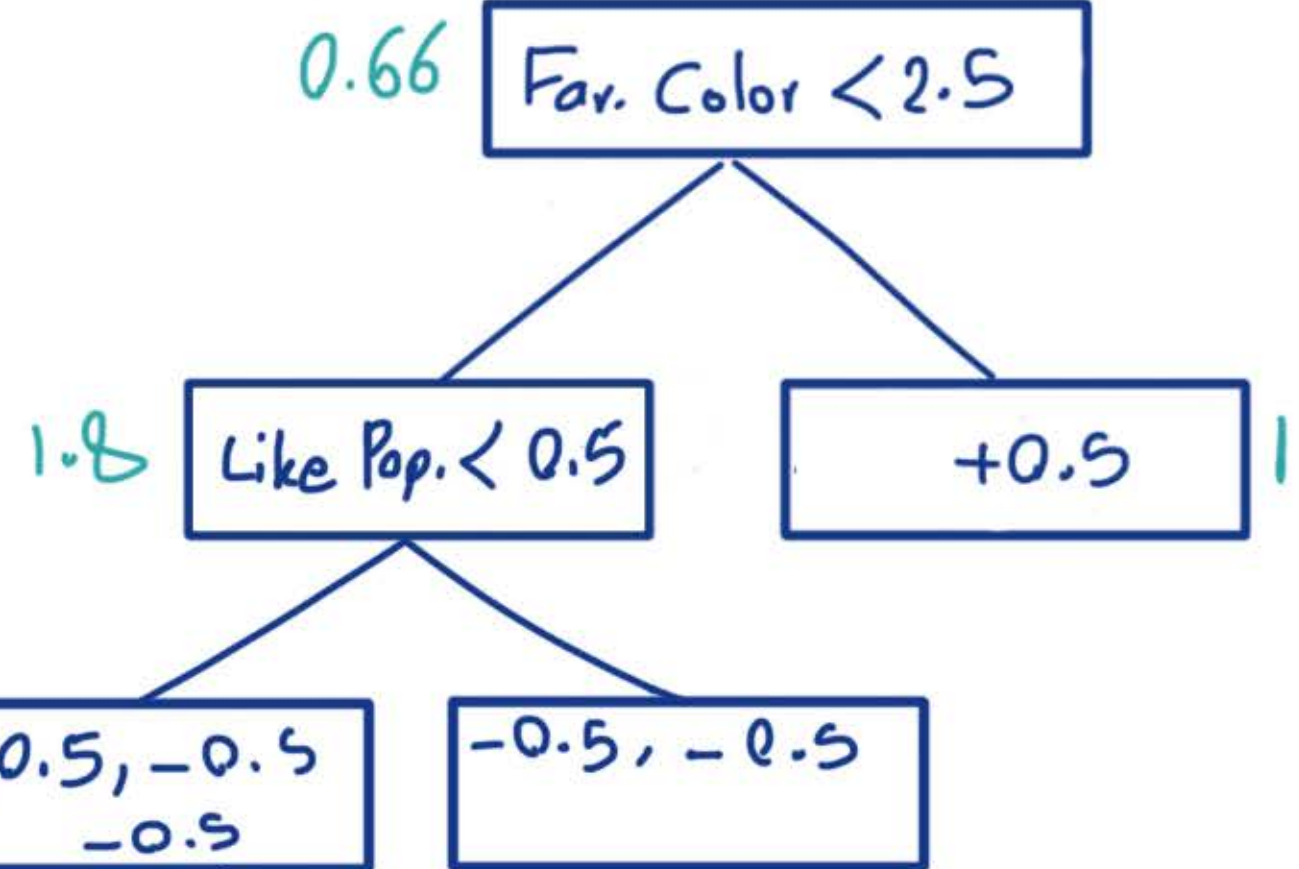


درخت یس بی اولیه  
0.5

Sim. score =  $\frac{(0.5 - 0.5 - 0.5)^2}{3 \times 0.25} = 0.33$

Likes Popcorn	Age	Favorite Color	Loves Troll 2
1	→ 12	1	1
1	→ 87	2	1
0	→ 44	1	0
1	19	3	0
0	→ 32	2	1
0	→ 14	1	1

Residual  
-0.5  
-0.5  
+0.5  
+0.5  
-0.5  
-0.5

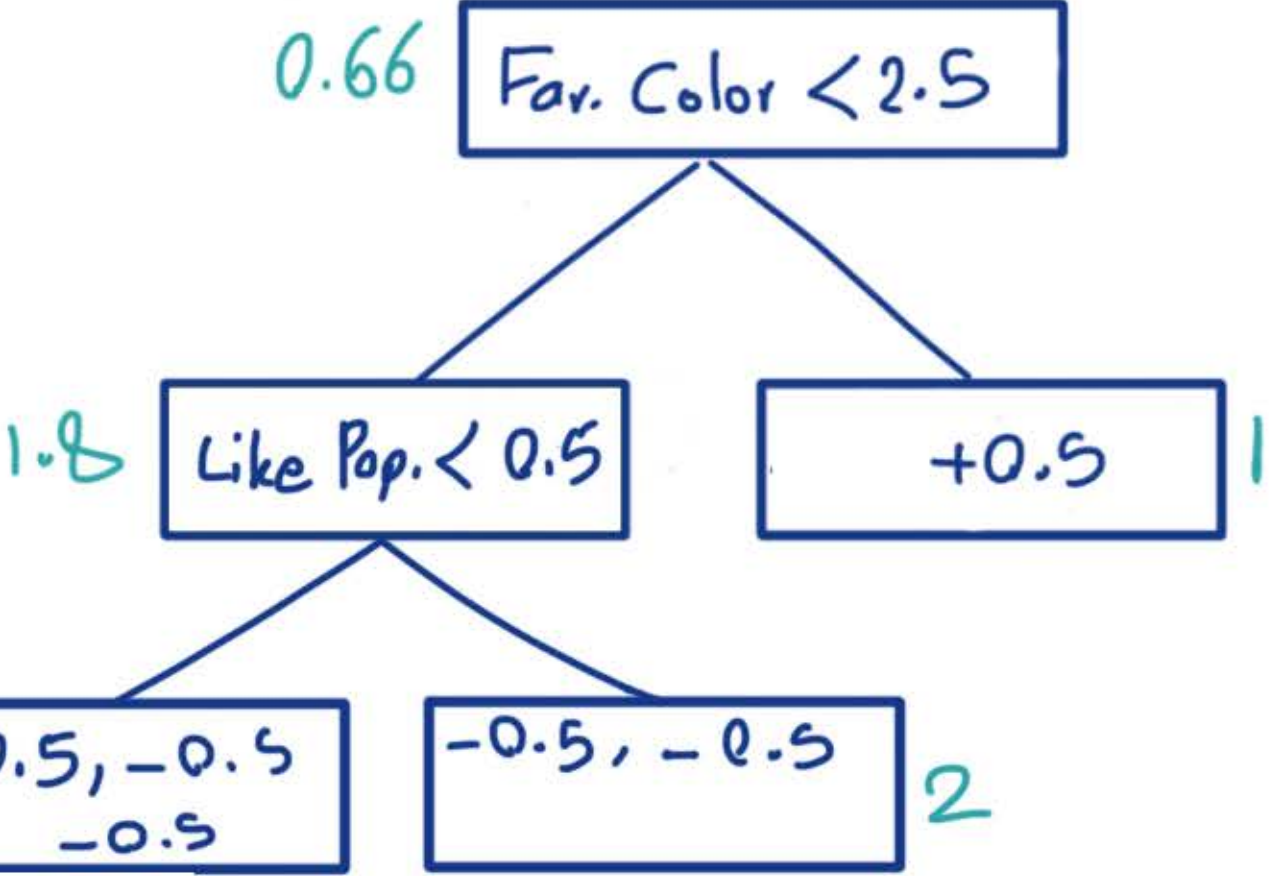


درخت یس بی اولیه  
0.5

$$\text{Sim. score} = \frac{(-0.5 - 0.5)^2}{2 \times 0.25} = 2$$

Likes Popcorn	Age	Favorite Color	Loves Troll 2
1	12	1	1
1	87	2	1
0	44	1	0
1	19	3	0
0	32	2	1
0	14	1	1

Residual  
-0.5  
-0.5  
+0.5  
+0.5  
-0.5  
-0.5



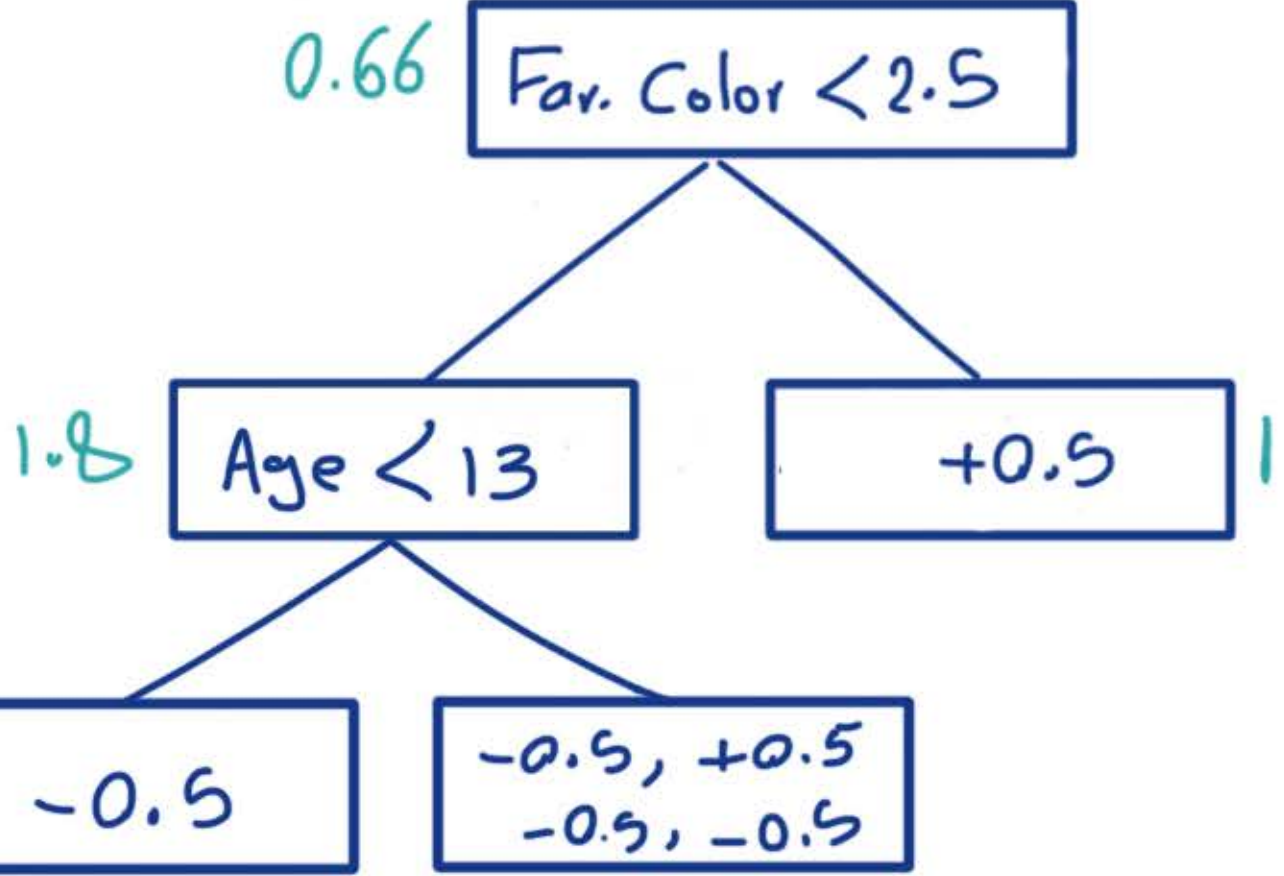
درخت یس بانی اولیه

0.5

$$\text{Gain} = 2 + 0.33 - 1.8 = 0.53$$

Likes Popcorn	Age	Favorite Color	Loves Troll 2
1	12	1	1
1	87	2	1
0	44	1	0
1	19	3	0
0	32	2	1
0	14	1	1

Residual  
 -0.5  
 -0.5  
 +0.5  
 +0.5  
 -0.5  
 -0.5



درخت یس بی اولیه

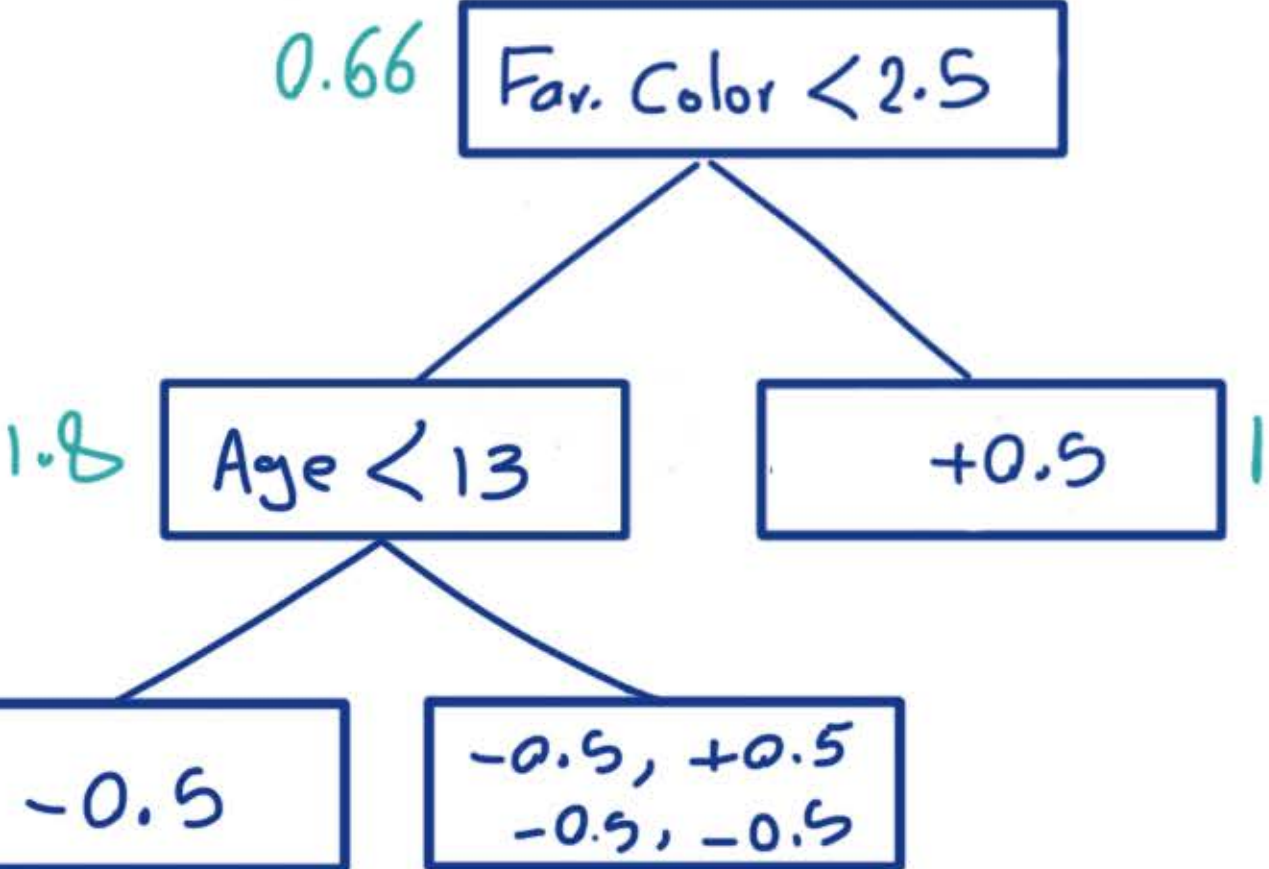
0.5

$$\text{sim. score} = \frac{(-0.5)^2}{0.25} = 1$$



Likes Popcorn	Age	Favorite Color	Loves Troll 2
1	12	1	1
1	87	2	1
0	44	1	0
1	19	3	0
0	32	2	1
0	14	1	1

Residual  
 -0.5  
 -0.5  
 +0.5  
 +0.5  
 -0.5  
 -0.5



درخت یسین بینی اولیه  
 0.5

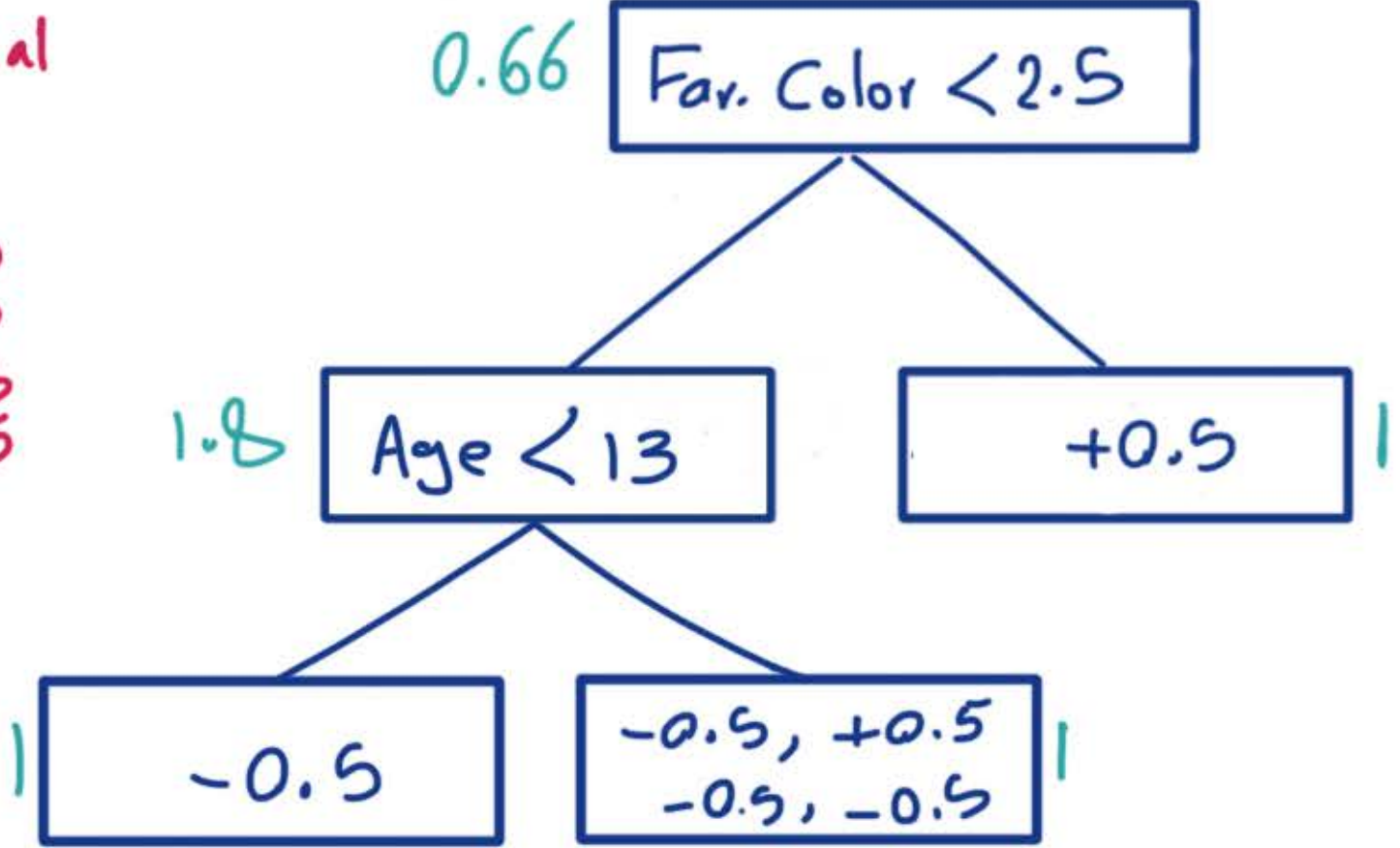
$$\text{sim. score} = \frac{(-0.5 + 0.5 - 0.5 - 0.5)^2}{4 \times 0.25} = 1$$

Likes Popcorn	Age	Favorite Color	Loves Troll 2
1	→ 12	1	1
1	→ 87	2	1
0	→ 44	1	0
1	19	3	0
0	→ 32	2	1
0	→ 14	1	1

Residual  
-0.5  
-0.5  
+0.5  
+0.5  
-0.5  
-0.5

درخت یس بی اولیه

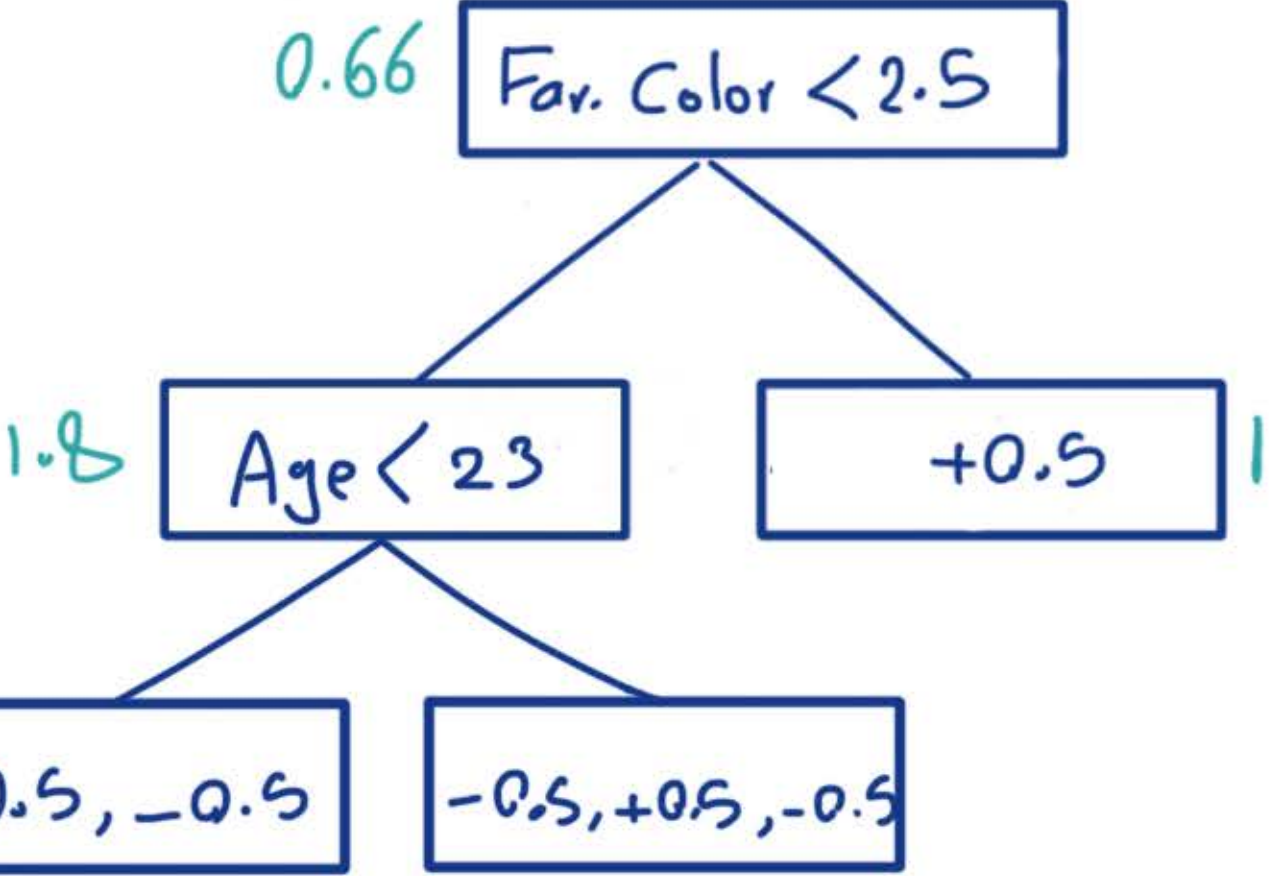
0.5



Gain = 1 + 1 - 1.8 = 0.2

Likes Popcorn	Age	Favorite Color	Loves Troll 2
1	12	1	1
1	87	2	1
0	44	1	0
1	19	3	0
0	32	2	1
0	14	1	1

Residual  
 -0.5  
 -0.5  
 +0.5  
 +0.5  
 -0.5  
 -0.5

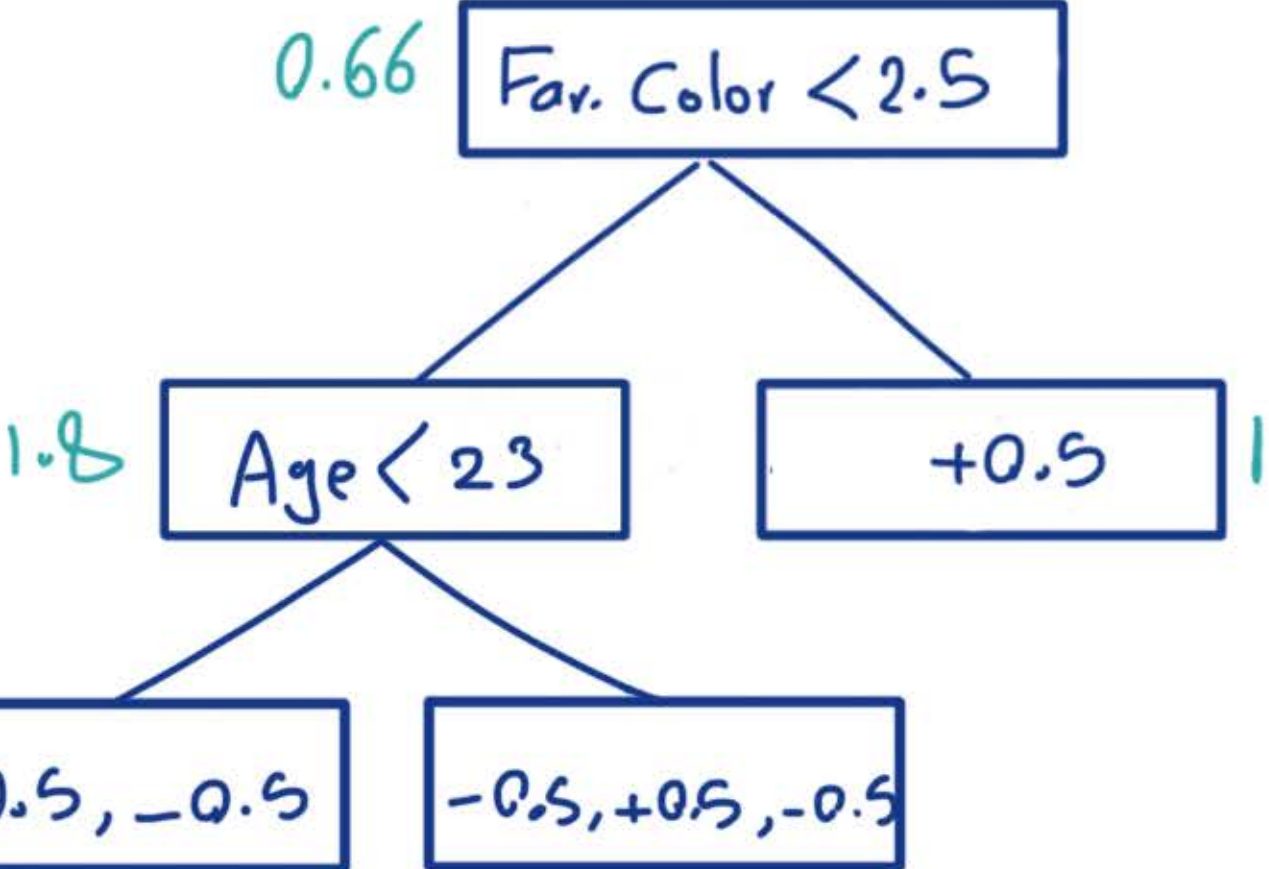


درخت یس بی اولیه  
 0.5

Sim score =  $\frac{(-0.5 - 0.5)^2}{2 \times 0.25} = 2$

Likes Popcorn	Age	Favorite Color	Loves Troll 2
1	12	1	1
1	87	2	1
0	44	1	0
1	19	3	0
0	32	2	1
0	14	1	1

Residual  
 -0.5  
 -0.5  
 +0.5  
 +0.5  
 -0.5  
 -0.5



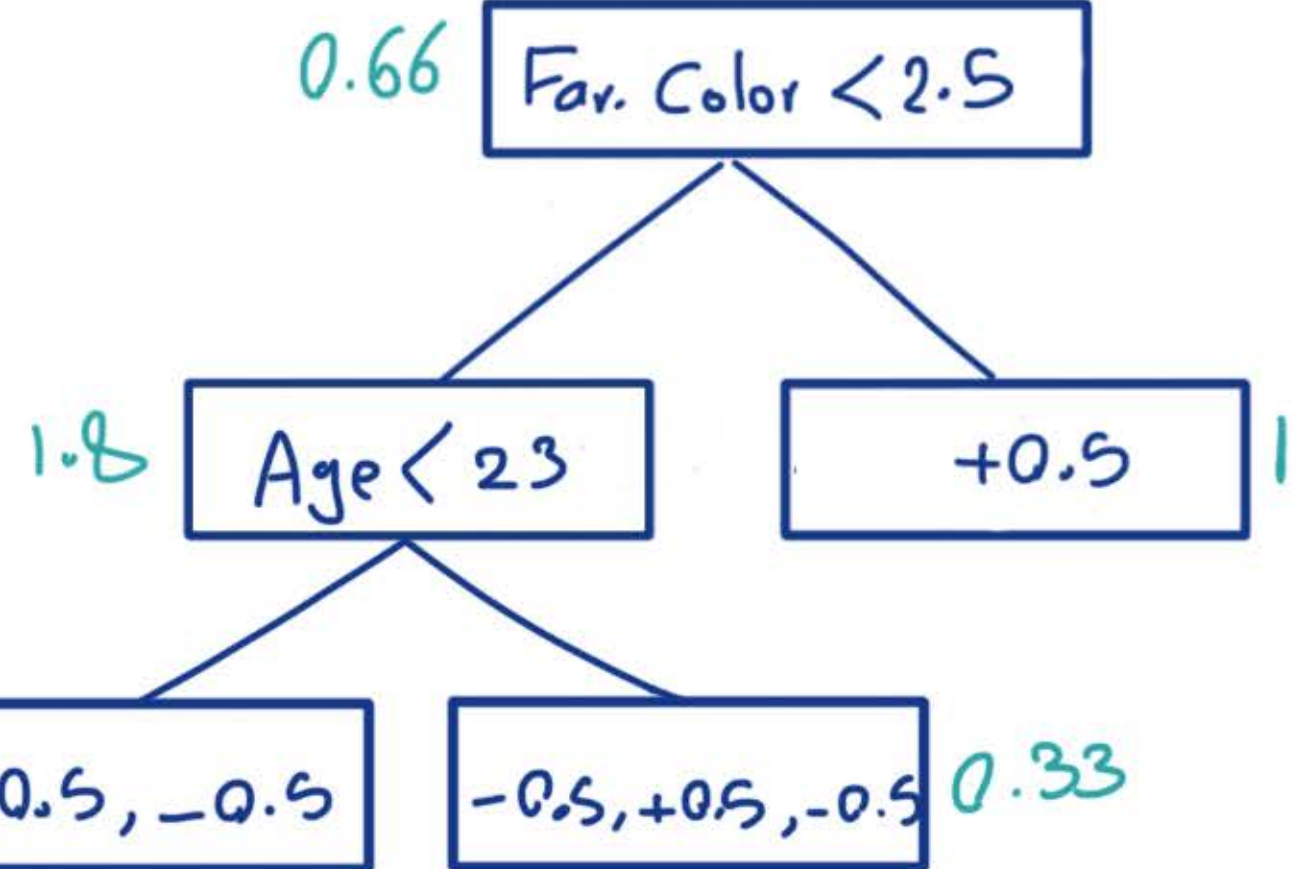
درخت ییس بینی اولیه  
 0.5

$$\text{Sim. score} = \frac{(-0.5 + 0.5 - 0.5)^2}{3 \times 0.25} = 0.33$$



Likes Popcorn	Age	Favorite Color	Loves Troll 2
1	12	1	1
1	87	2	1
0	44	1	0
1	19	3	0
0	32	2	1
0	14	1	1

Residual  
-0.5  
-0.5  
+0.5  
+0.5  
-0.5  
-0.5



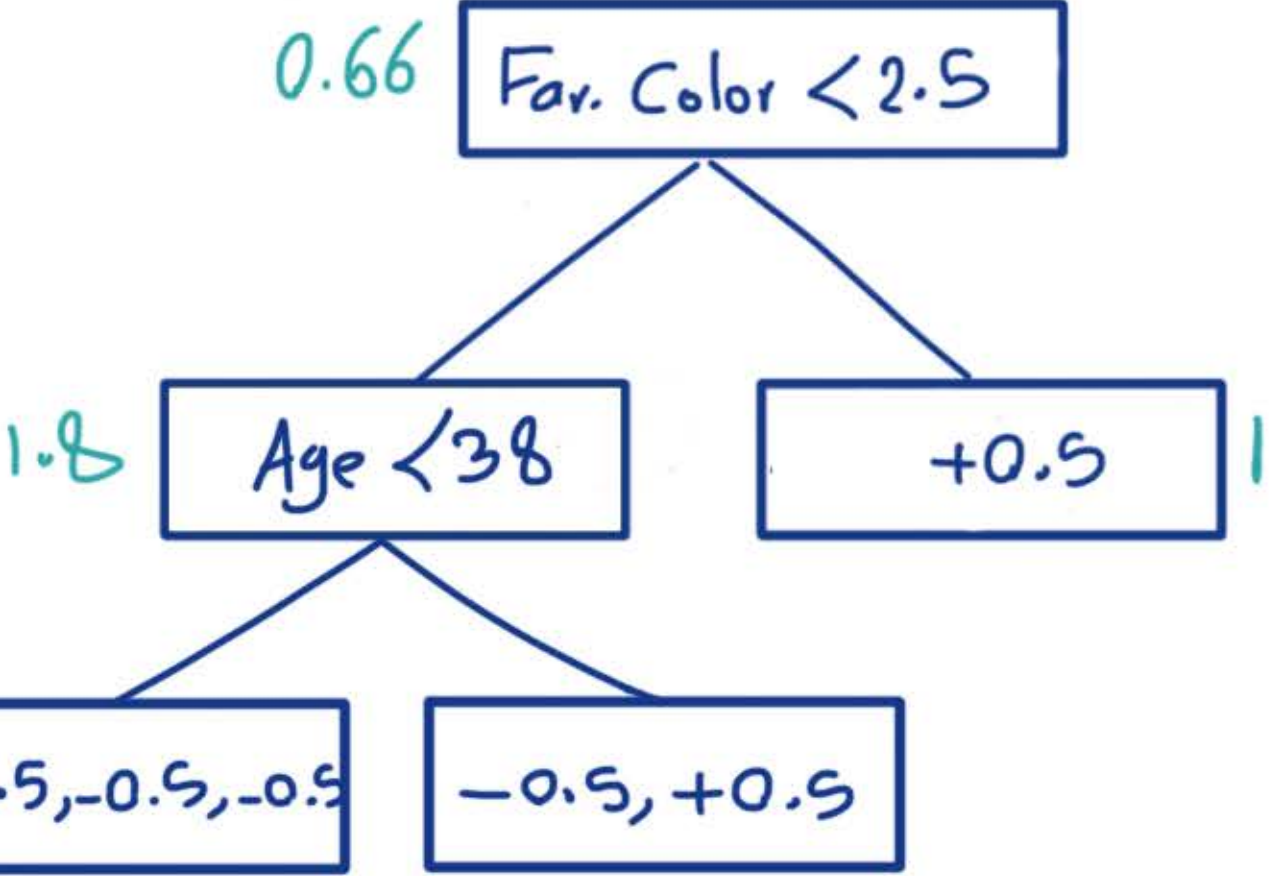
درخت یسین بینی اولیه

0.5

$Gain = 2 + 0.33 - 1.8 = 0.53$

Likes Popcorn	Age	Favorite Color	Loves Troll 2
1	12	1	1
1	87	2	1
0	44	1	0
1	19	3	0
0	32	2	1
0	14	1	1

Residual  
 -0.5  
 -0.5  
 +0.5  
 +0.5  
 -0.5  
 -0.5

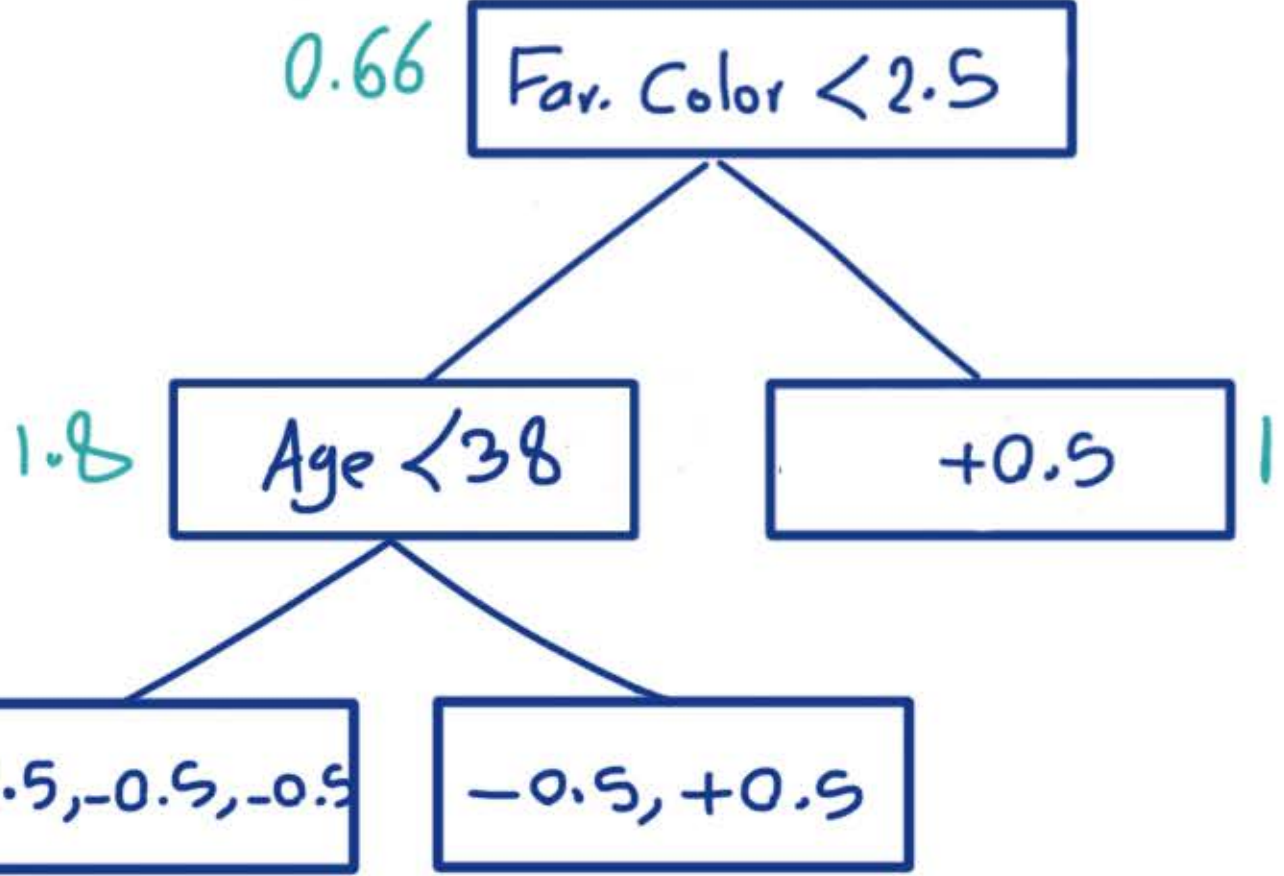


درخت یسین بینی اولیه  
 0.5

Sim score =  $\frac{(-0.5 - 0.5 - 0.5)^2}{3 \times 0.25} = 3$

Likes Popcorn	Age	Favorite Color	Loves Troll 2
1	12	1	1
1	87	2	1
0	44	1	0
1	19	3	0
0	32	2	1
0	14	1	1

Residual  
 -0.5  
 -0.5  
 +0.5  
 +0.5  
 -0.5  
 -0.5

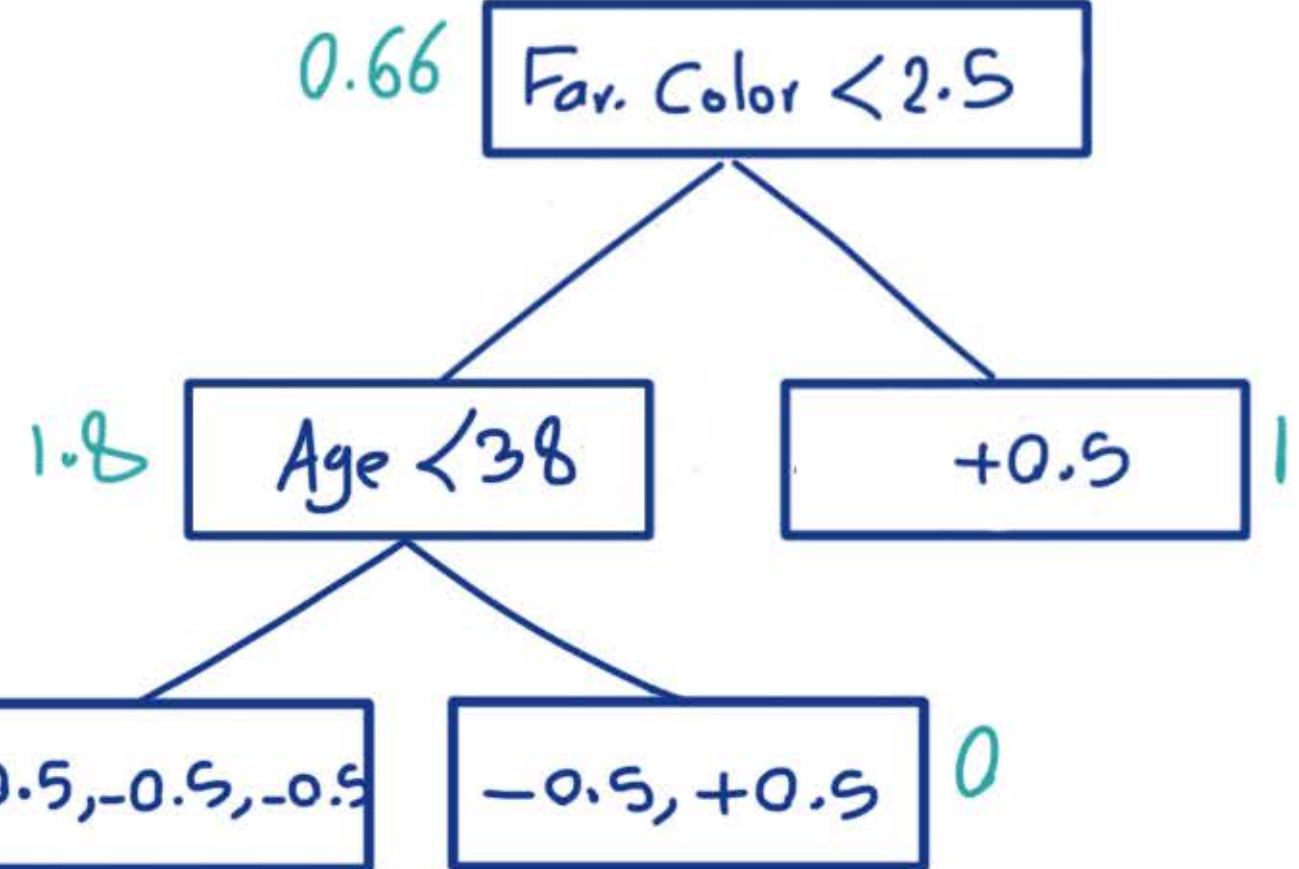


درخت یسین بینی اولیه  
 0.5

$$\text{Sim score} = \frac{(-0.5 + 0.5)^2}{2 \times 0.25} = 0$$

Likes Popcorn	Age	Favorite Color	Loves Troll 2
1	12	1	1
1	87	2	1
0	44	1	0
1	19	3	0
0	32	2	1
0	14	1	1

Residual  
-0.5  
-0.5  
+0.5  
+0.5  
-0.5  
-0.5



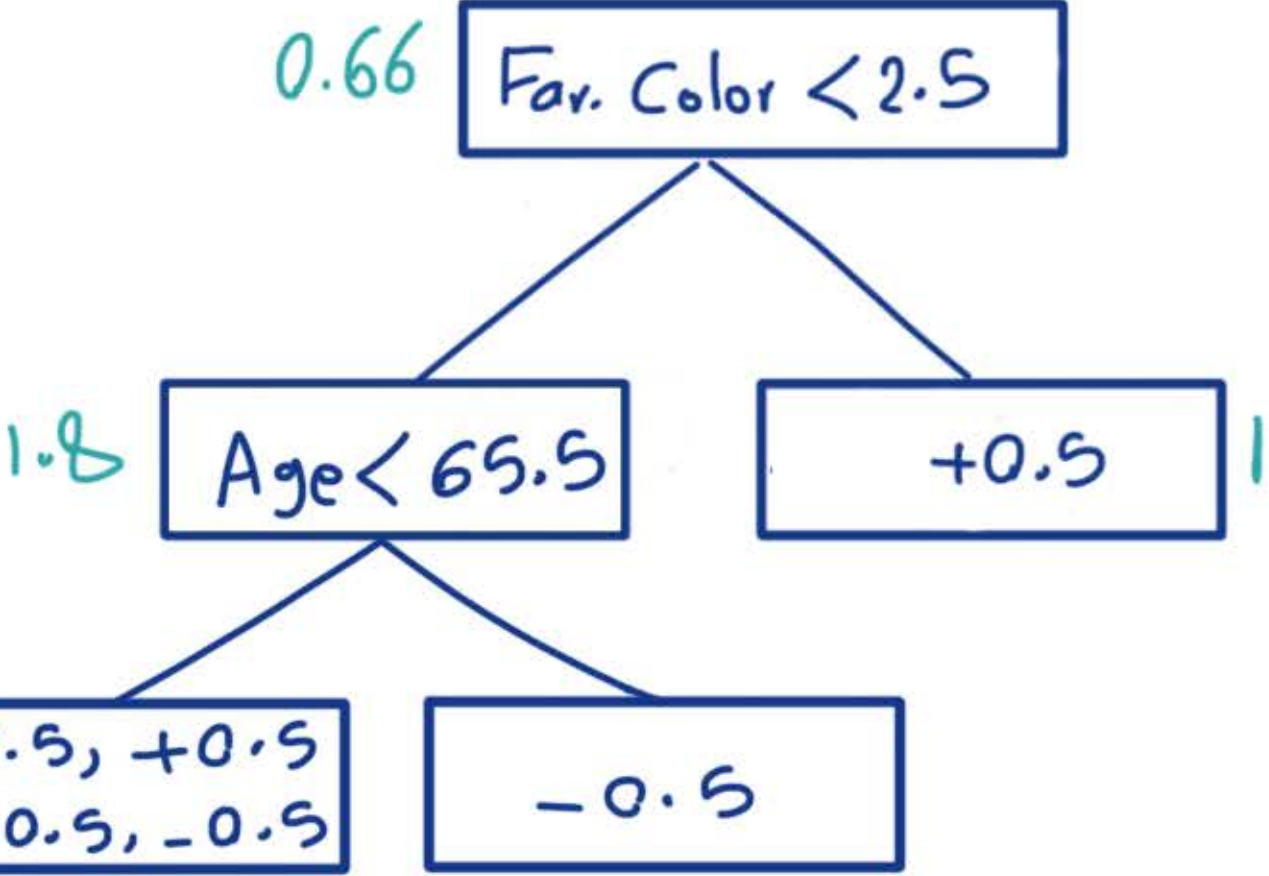
درخت یس بی اولیه  
0.5

Gain = 3 + 0 - 1.8 = 1.2



Likes Popcorn	Age	Favorite Color	Loves Troll 2
1	12	1	1
1	87	2	1
0	44	1	0
1	19	3	0
0	32	2	1
0	14	1	1

Residual  
-0.5  
-0.5  
+0.5  
+0.5  
-0.5  
-0.5

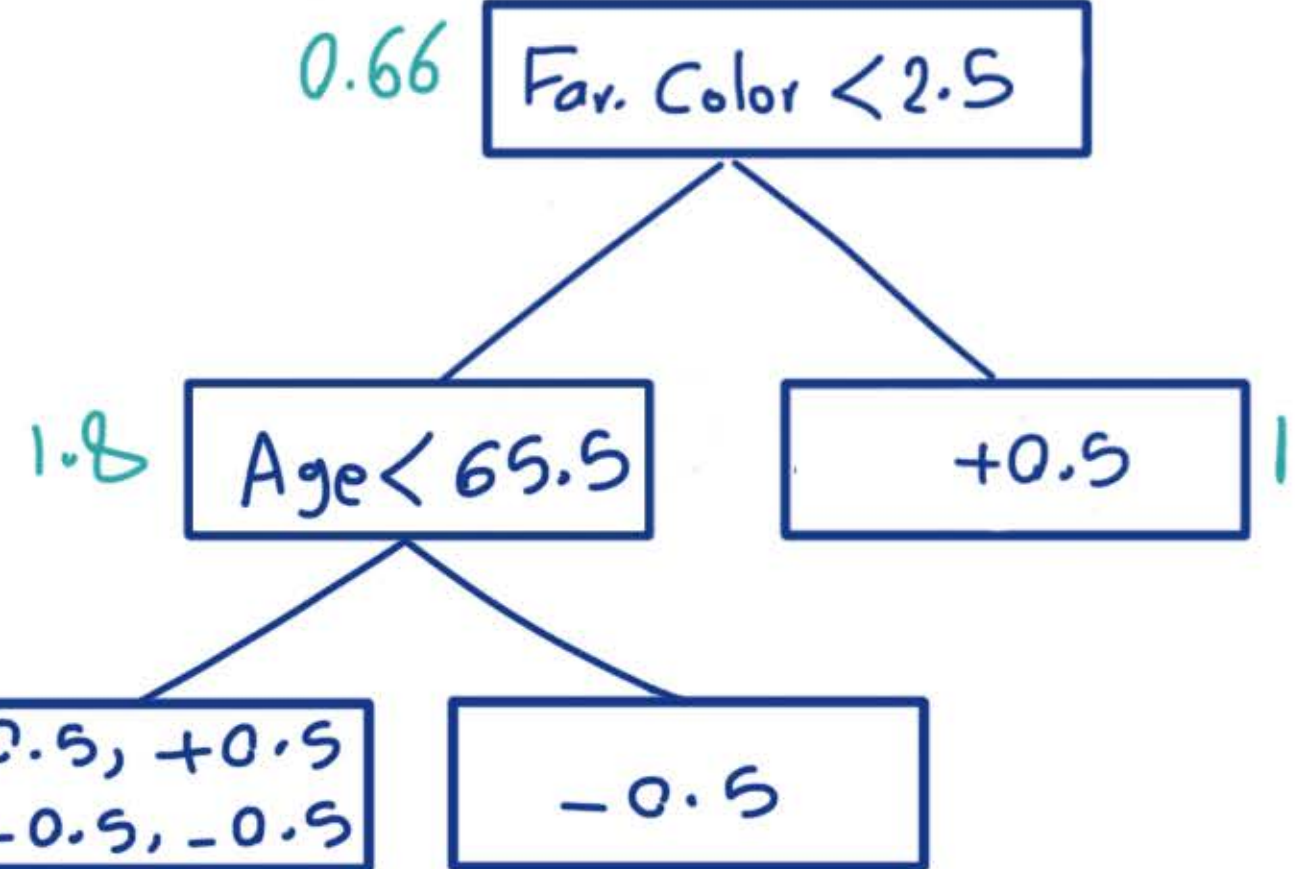


درخت یسین بینی اولیه  
0.5

Sim. score =  $\frac{(-0.5 + 0.5 - 0.5 - 0.5)^2}{4 \times 0.25} = 1$

Likes Popcorn	Age	Favorite Color	Loves Troll 2
1	12	1	1
1	87	2	1
0	44	1	0
1	19	3	0
0	32	2	1
0	14	1	1

Residual  
 -0.5  
 -0.5  
 +0.5  
 +0.5  
 -0.5  
 -0.5

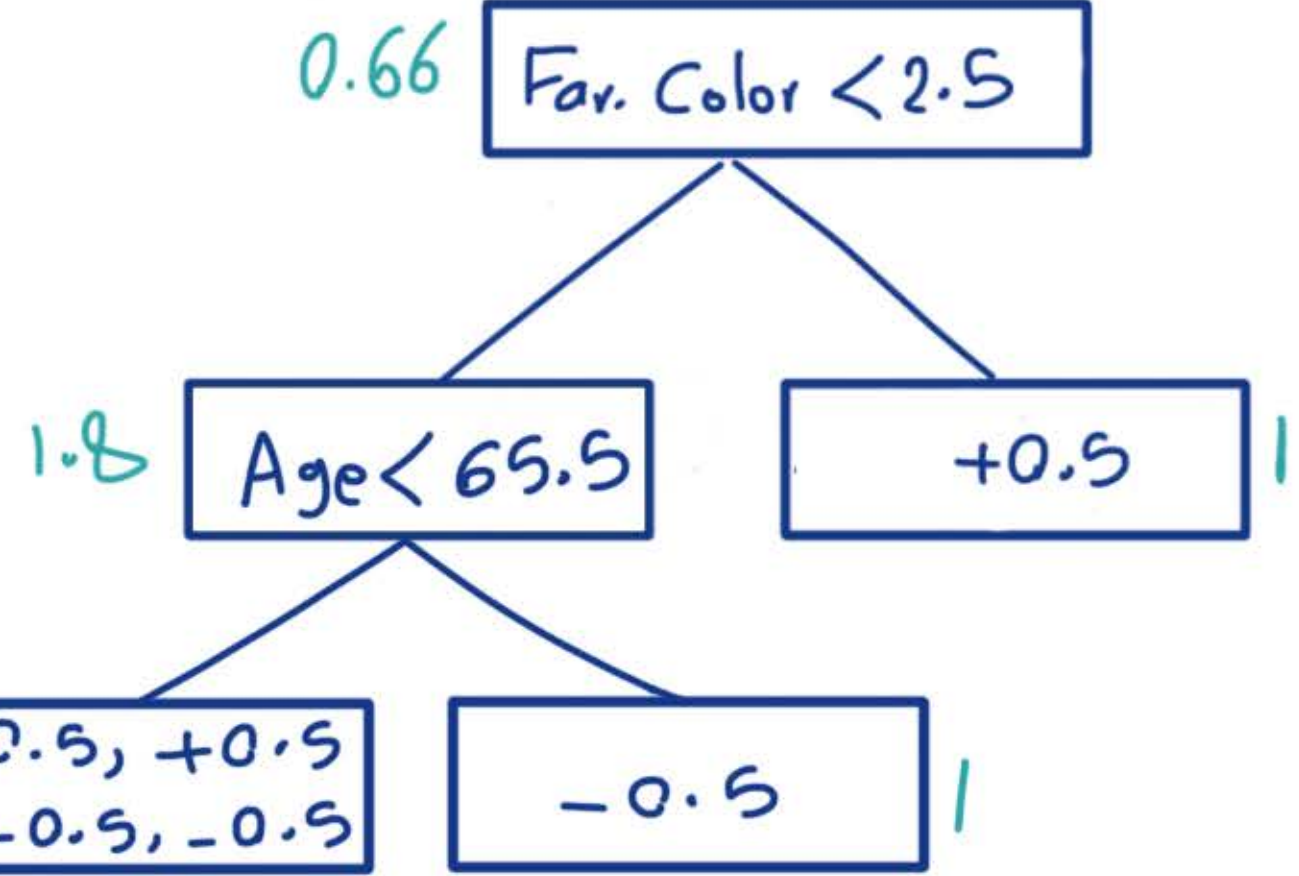


درخت یسین بینی اولیه  
 0.5

$$\text{Sim. Score} = \frac{(-0.5)^2}{0.25} = 1$$

Likes Popcorn	Age	Favorite Color	Loves Troll 2
1	12	1	1
1	87	2	1
0	44	1	0
1	19	3	0
0	32	2	1
0	14	1	1

Residual  
-0.5  
-0.5  
+0.5  
+0.5  
-0.5  
-0.5



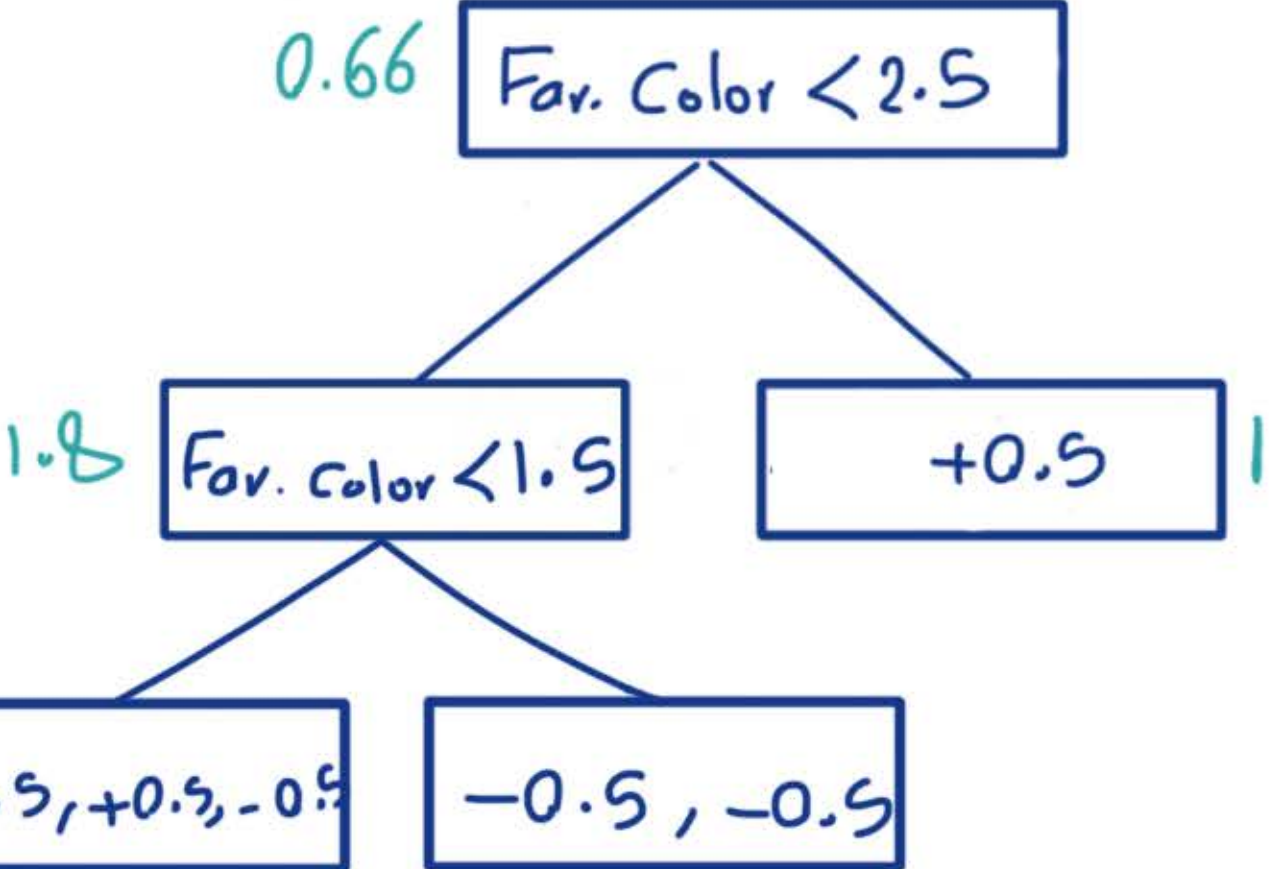
درخت یس بی اولیه

0.5

Gain = 1 + 1 - 1.8 = 0.2

Likes Popcorn	Age	Favorite Color	Loves Troll 2
1	12	1	1
1	87	2	1
0	44	1	0
1	19	3	0
0	32	2	1
0	14	1	1

Residual  
-0.5  
-0.5  
+0.5  
+0.5  
-0.5  
-0.5



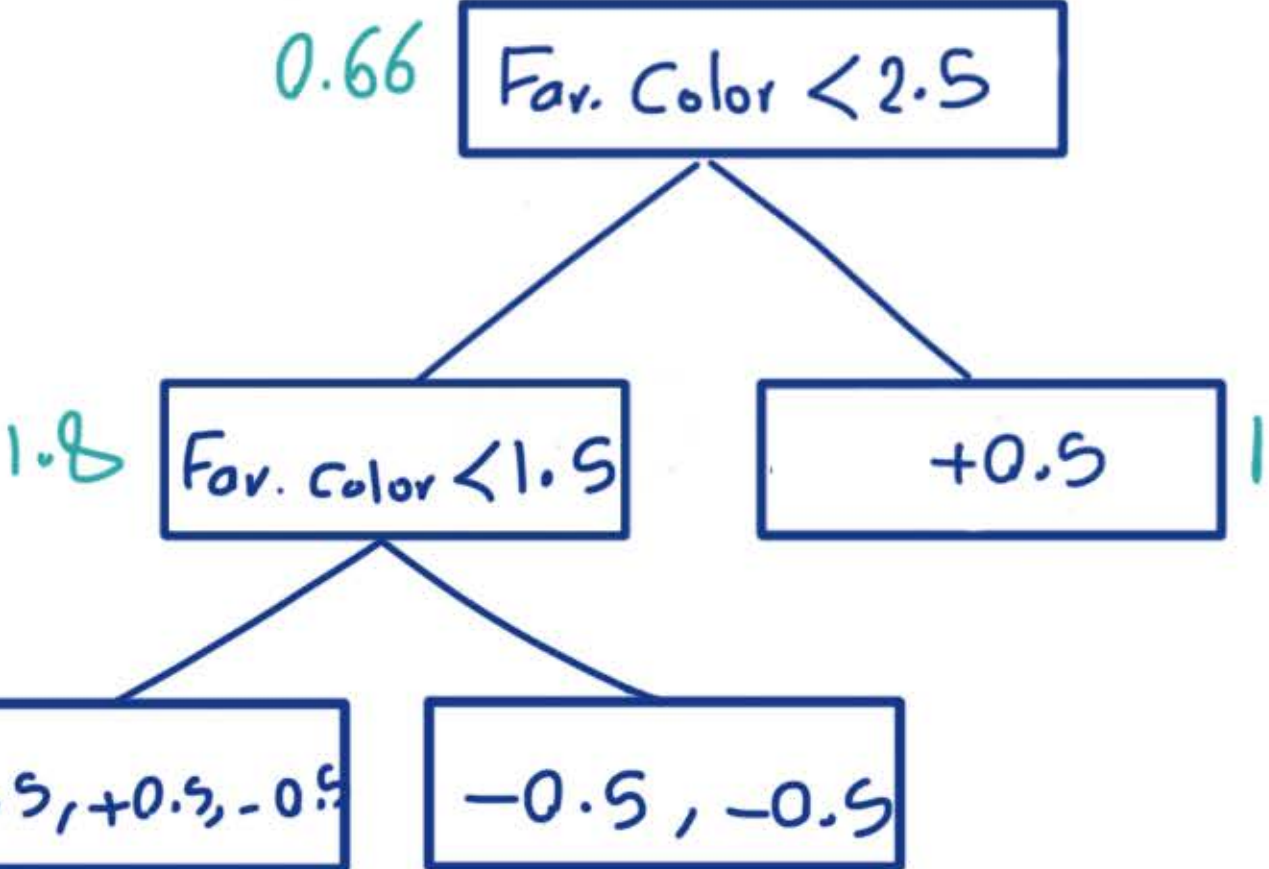
درخت یس بی اولیه  
0.5

$$\text{Sim score} = \frac{(-0.5 + 0.5 - 0.5)^2}{3 \times 0.25} = 0.33$$



Likes Popcorn	Age	Favorite Color	Loves Troll 2
1	→ 12	1	1
1	→ 87	2	1
0	→ 44	1	0
1	19	3	0
0	→ 32	2	1
0	→ 14	1	1

Residual  
 -0.5  
 -0.5  
 +0.5  
 +0.5  
 -0.5  
 -0.5

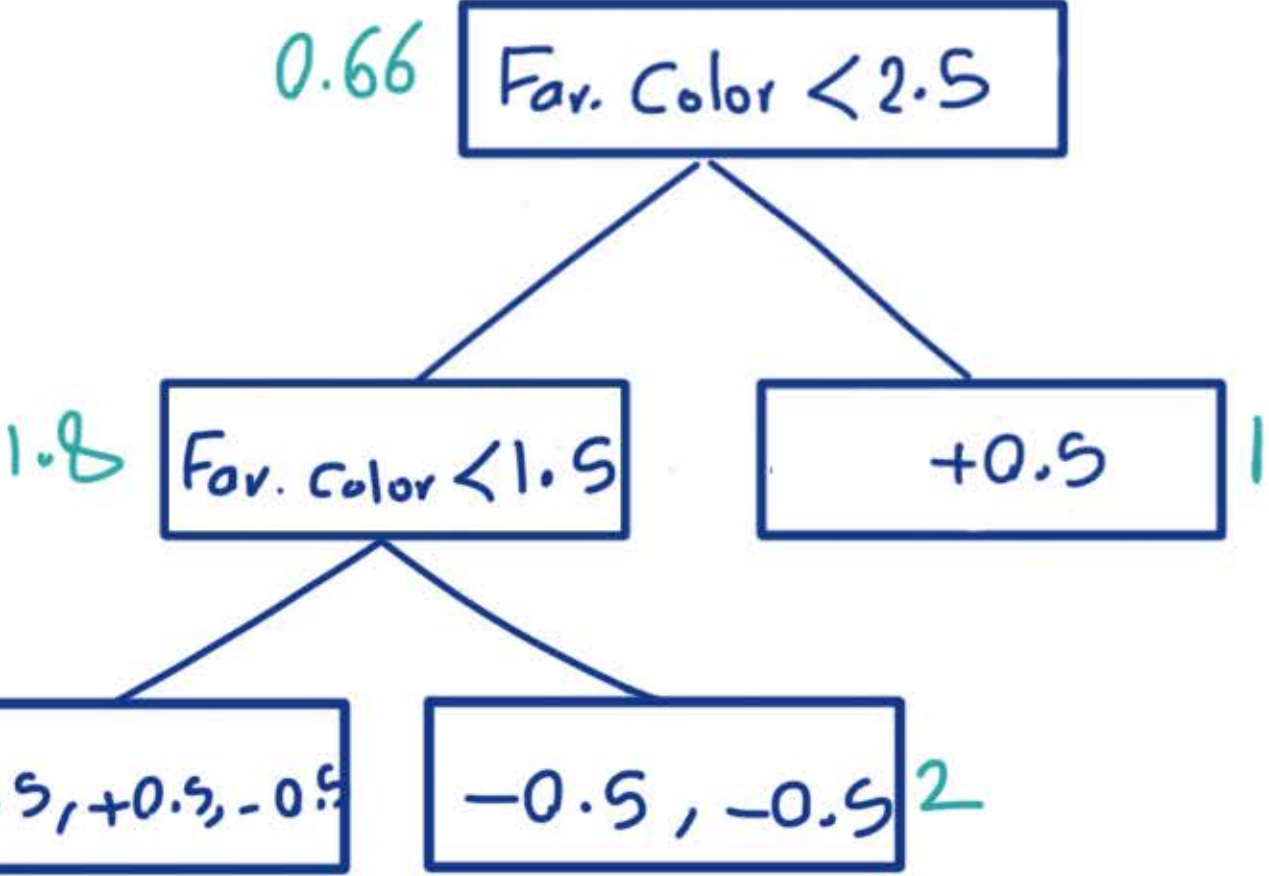


درخت یسین بنی اولیه  
 0.5

$$\text{Sim. score} = \frac{(-0.5 - 0.5)^2}{2 \times 0.25} = 2$$

Likes Popcorn	Age	Favorite Color	Loves Troll 2
1	12	1	1
1	87	2	1
0	44	1	0
1	19	3	0
0	32	2	1
0	14	1	1

Residual  
-0.5  
-0.5  
+0.5  
+0.5  
-0.5  
-0.5

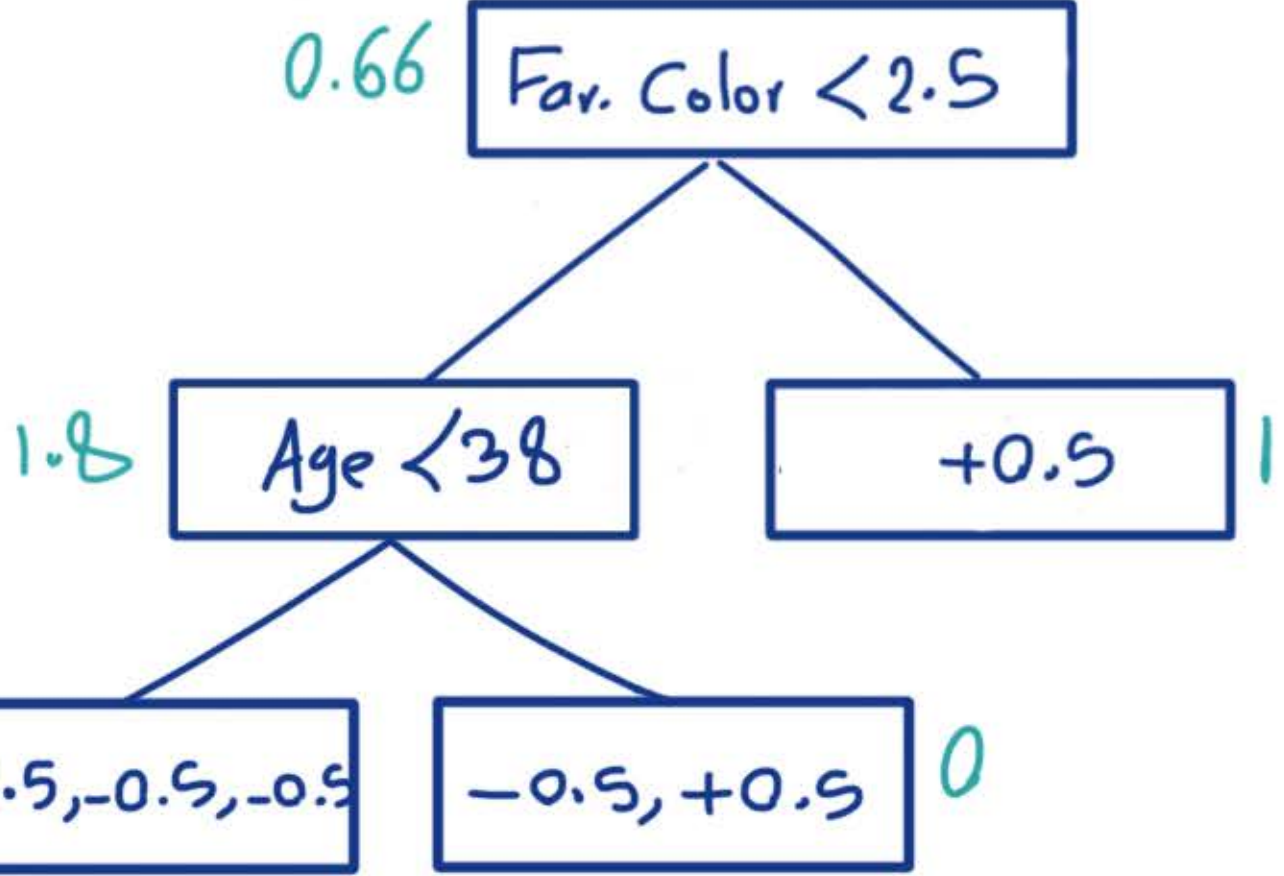


درخت یس بی اولیه  
0.5

$$\text{Gain} = 2 + 0.33 - 1.8 - 0.53$$

Likes Popcorn	Age	Favorite Color	Loves Troll 2
1	12	1	1
1	87	2	1
0	44	1	0
1	19	3	0
0	32	2	1
0	14	1	1

Residual  
 -0.5  
 -0.5  
 +0.5  
 +0.5  
 -0.5  
 -0.5



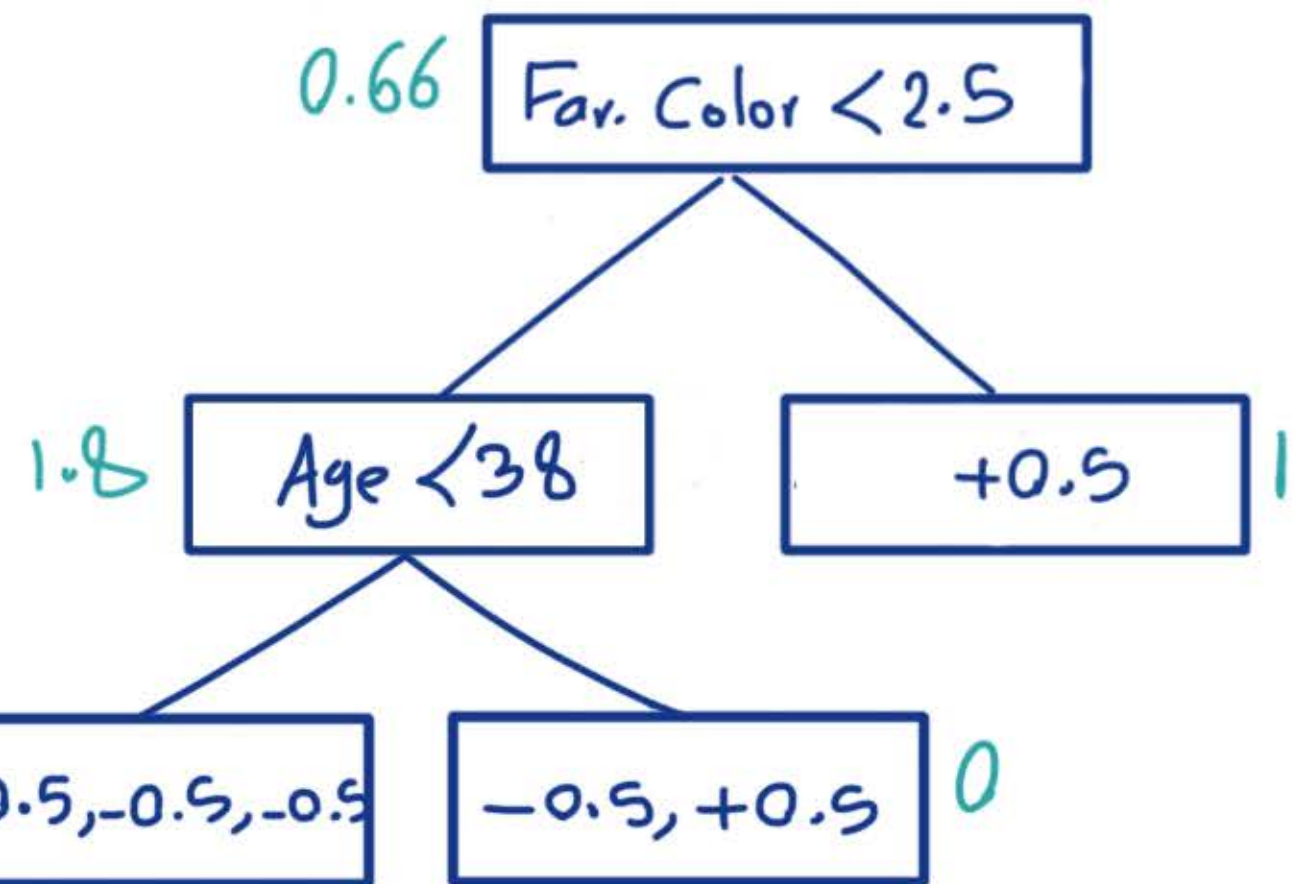
درخت پیس بینی اولیه  
 0.5

پس بیشترین Gain روی این درخت به دست می آید. حال اگر فرضی کنیم حد اکثر محقق مجاز درخت برابر ۲ باشد، کار ساخت این درخت به پایان می رسد و باید درخت جدید بسازیم.



Likes Popcorn	Age	Favorite Color	Loves Troll 2
1	→ 12	1	1
1	→ 87	2	1
0	→ 44	1	0
1	19	3	0
0	→ 32	2	1
0	→ 14	1	1

Residual  
 -0.5  
 -0.5  
 +0.5  
 +0.5  
 -0.5  
 -0.5



درخت یس بی اولیه

0.5

قبل از آن می دلیتم output هر برگ درخت برابر است با :

اگر از اول residual را برابر  $y - \hat{y}$  می گرفتیم به این علامت نیاز نبود

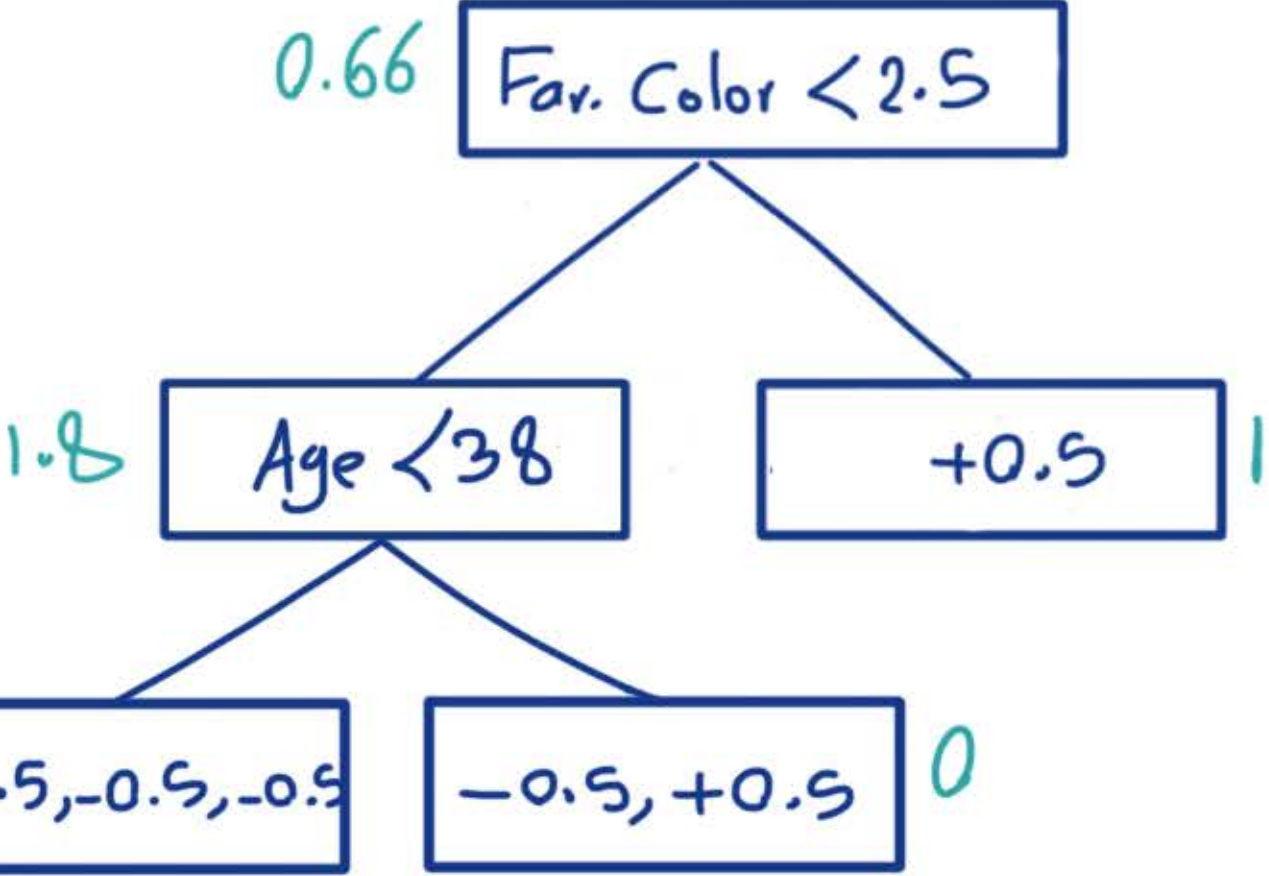
$$-\sum (Residual_i)$$

$$\sum [Previous Prob. i \times (1 - Previous Prob i)] + \lambda$$



Likes Popcorn	Age	Favorite Color	Loves Troll 2
1	12	1	1
1	87	2	1
0	44	1	0
1	19	3	0
0	32	2	1
0	14	1	1

Residual  
 -0.5  
 -0.5  
 +0.5  
 +0.5  
 -0.5  
 -0.5



درخت پیس بینی اولیه

0.5

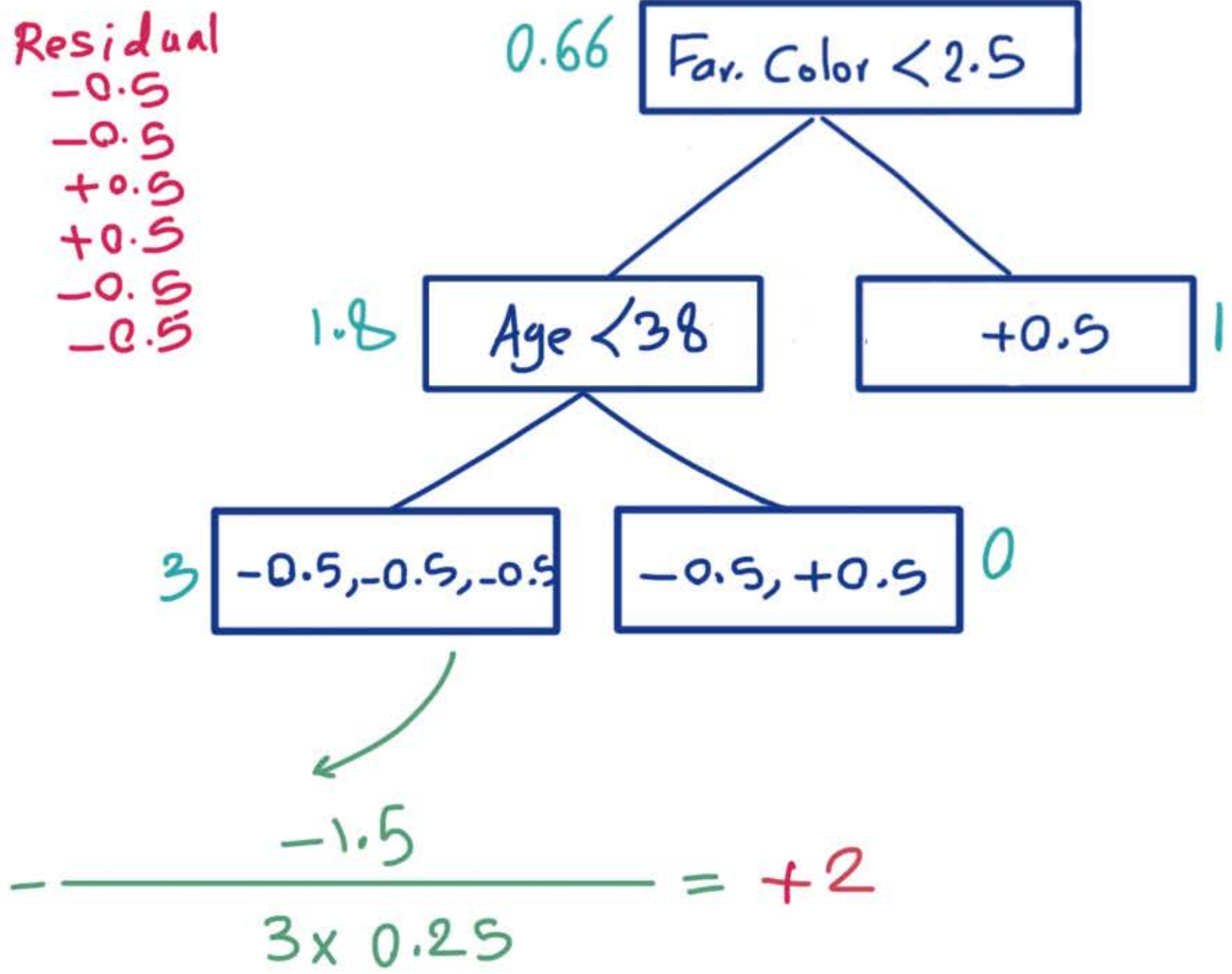
پس خردوبی برگ هلی درخت را محاسبه می کنیم تا از آن برای  
 دسته بندی استفاده کنیم.

Likes Popcorn	Age	Favorite Color	Loves Troll 2
1	→ 12	1	1
1	→ 87	2	1
0	→ 44	1	0
1	19	3	0
0	→ 32	2	1
0	→ 14	1	1

Residual  
 -0.5  
 -0.5  
 +0.5  
 +0.5  
 -0.5  
 -0.5

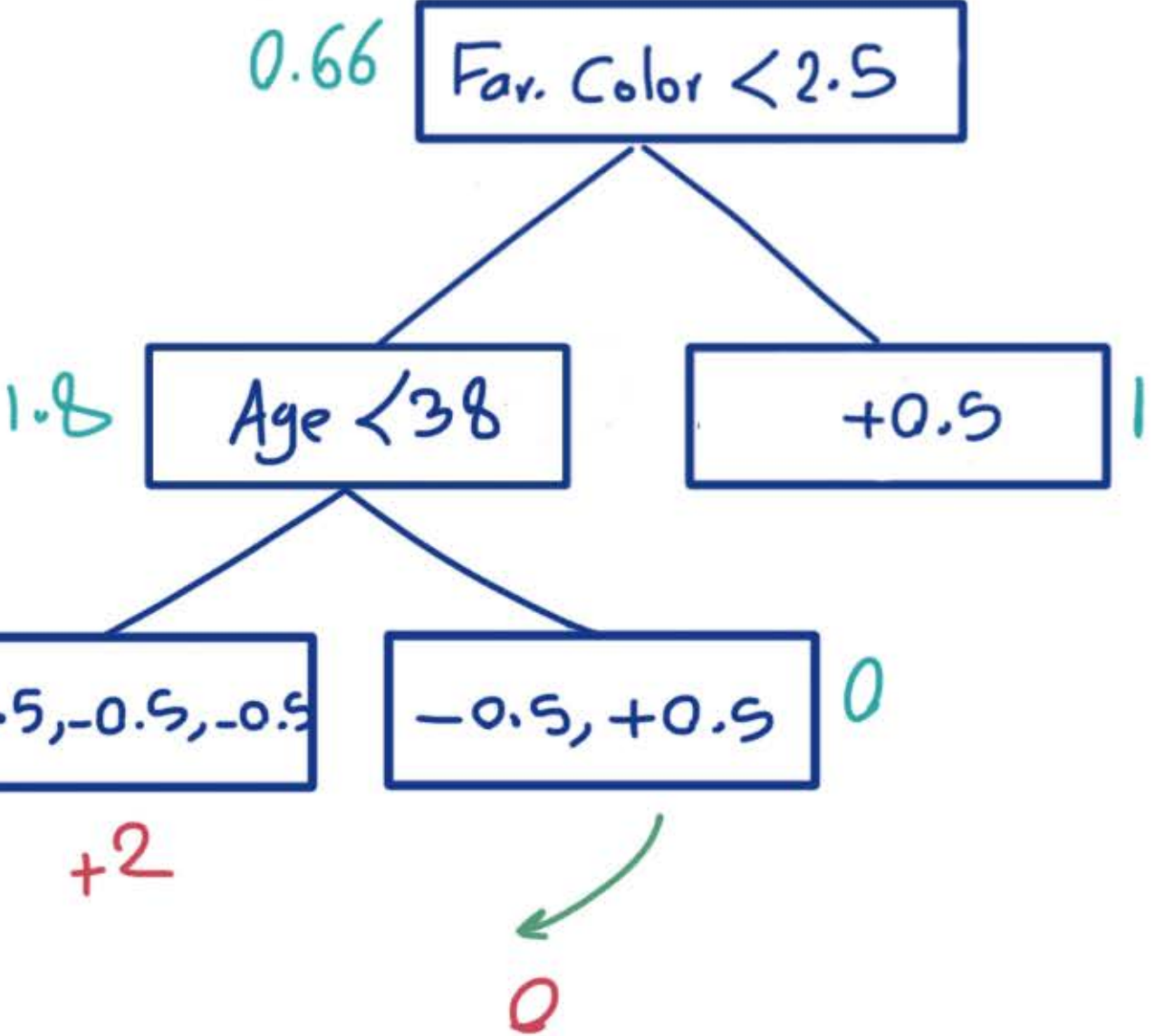
درخت یس بی اولیه

0.5



Likes Popcorn	Age	Favorite Color	Loves Troll 2
1	12	1	1
1	87	2	1
0	44	1	0
1	19	3	0
0	32	2	1
0	14	1	1

Residual  
 -0.5  
 -0.5  
 +0.5  
 +0.5  
 -0.5  
 -0.5

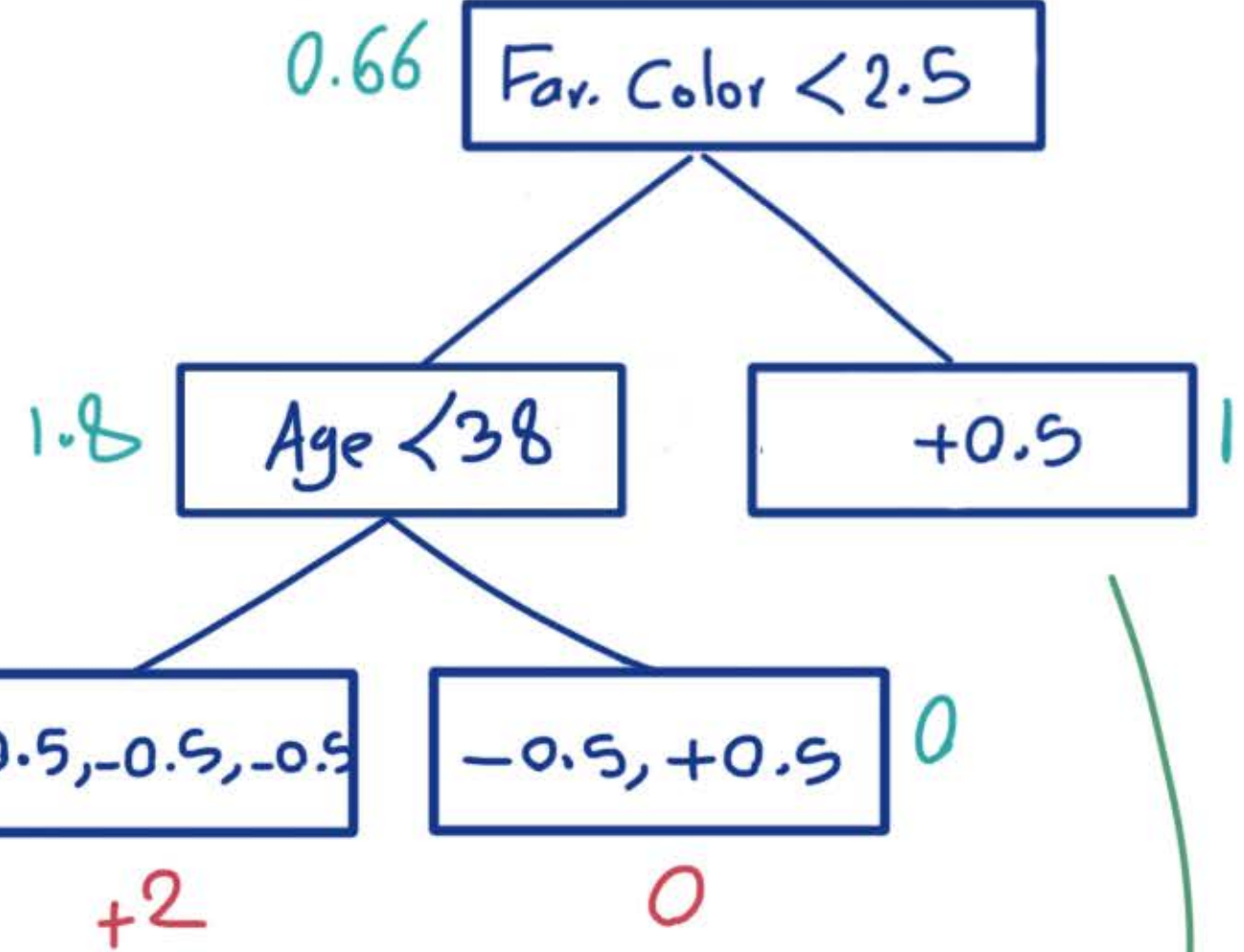


درخت یس بی اولیه

0.5

Likes Popcorn	Age	Favorite Color	Loves Troll 2
1	12	1	1
1	87	2	1
0	44	1	0
1	19	3	0
0	32	2	1
0	14	1	1

Residual  
-0.5  
-0.5  
+0.5  
+0.5  
-0.5  
-0.5



$$-2 = \frac{-0.5}{0.25}$$

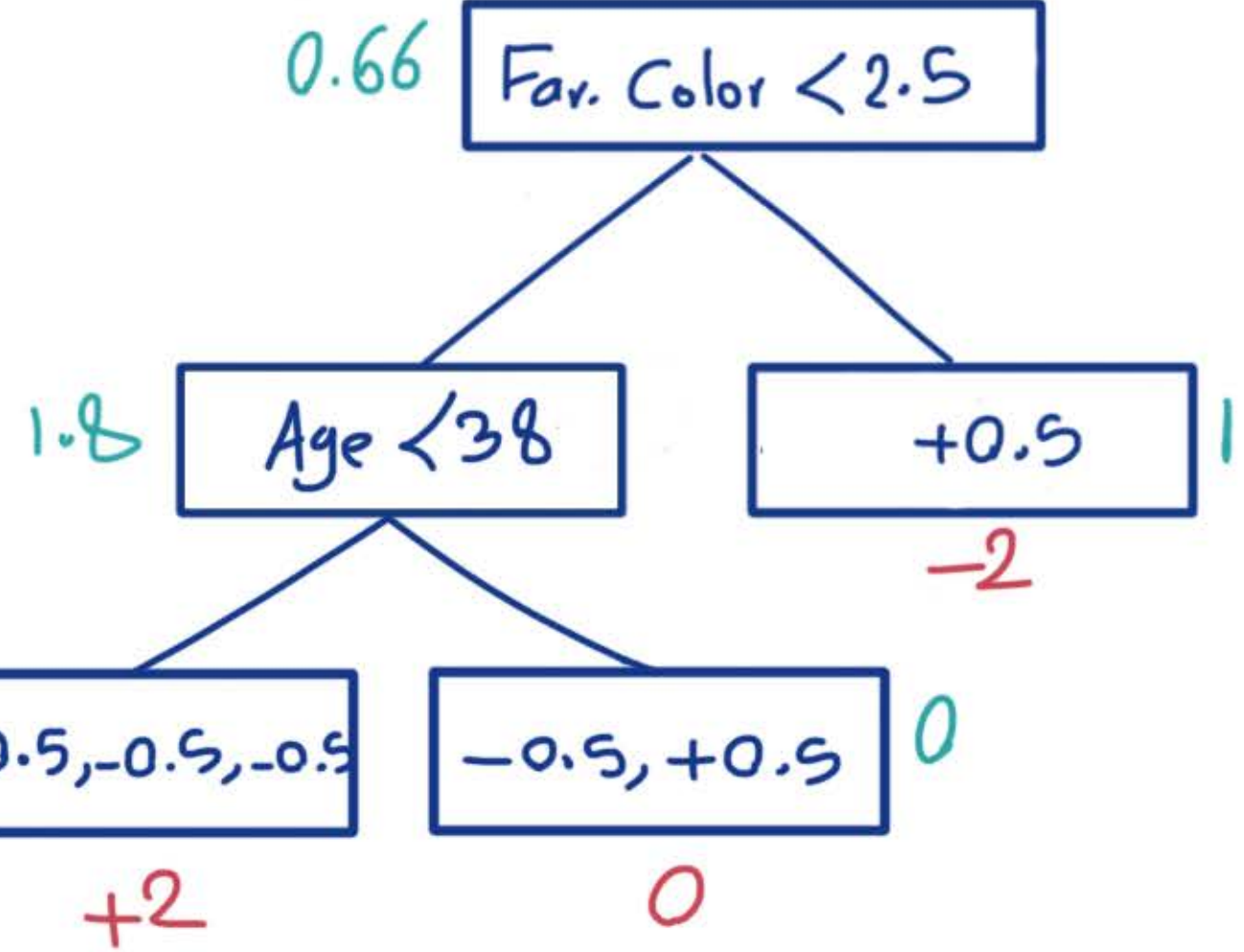
درخت یس بانی اولیه

0.5



Likes Popcorn	Age	Favorite Color	Loves Troll 2
1	12	1	1
1	87	2	1
0	44	1	0
1	19	3	0
0	32	2	1
0	14	1	1

Residual  
 -0.5  
 -0.5  
 +0.5  
 +0.5  
 -0.5  
 -0.5



درخت پیس بینی اولیه

0.5

دس خدوچی در برگ هلی درخت ب شکل بالا می بارید.

به شکل پیش‌فرض برای دسته‌بندی توسط XGBoost، learning rate برابر 0.3 است. اگر بیش‌تر

احتمال  $\sigma$  بودن خروجی برای کلاس داده تا لحظه  $t-1$  را با  $p_{t-1}(x)$  غایی دهیم و خروجی درخت کنونی  $(f_t)$  به ازای داده‌ی  $x$  را توسط  $f_t(x)$  غایی دهیم، احتمال جدید برای  $x$  از این رابطه محاسب می‌شود:

$$p_t(x) = \sigma \left( \log \left( \frac{p_{t-1}(x)}{1 - p_{t-1}(x)} \right) + f_t(x) \right)$$

$$\sigma(x) = \frac{e^x}{1 + e^x}$$

طبق این رابطه احتمال بودن هر داده در کلاس ۱ را محاسب می‌کنیم.

<i>Likes Popcorn</i>	<i>Age</i>	<i>Favorite Color</i>	<i>Loves Troll 2</i>
1	12	1	1
1	87	2	1
0	44	1	0
1	19	3	0
0	32	2	1
0	14	1	1

$$\rightarrow \sigma\left(\log\left(\frac{0.5}{1-0.5}\right) + 0.3 \times 2\right) = 0.65$$

$$\rightarrow \sigma\left(\log\left(\frac{0.5}{1-0.5}\right) + 0.3 \times 0\right) = 0.5$$

$$\rightarrow \sigma\left(\log\left(\frac{0.5}{1-0.5}\right) + 0.3 \times 0\right) = 0.5$$

$$\rightarrow \sigma\left(\log\left(\frac{0.5}{1-0.5}\right) + 0.3 \times -2\right) = 0.35$$

$$\rightarrow \sigma\left(\log\left(\frac{0.5}{1-0.5}\right) + 0.3 \times 2\right) = 0.65$$

$$\rightarrow \sigma\left(\log\left(\frac{0.5}{1-0.5}\right) + 0.3 \times 2\right) = 0.65$$



همانطور که مشاهده می شود از پیش بینی 0.5 به سمت برجسته هر داده حرکت کرده ایم و خط را کاهش داده ایم. با ادامه این روند داده هایی که هنوز خوب در سه نبی شده اند به درسه ی مربوطه نزدیک تر خواهند شد.

برای انبکار ابتدا احتمال معای سه ی هر داده - همراه residual در جدول قرار می گیرد.



Likes Popcorn	Age	Favorite Color	Loves Troll 2	Prob	Res.
1	12	1	1	0.65	-0.35
1	87	2	1	0.5	-0.5
0	44	1	0	0.5	+0.5
1	19	3	0	0.35	+0.35
0	32	2	1	0.65	-0.35
0	14	1	1	0.65	-0.35

حال مجدداً مقادیر Residual را در یک بزرگ قدر  
فی دهیم و Similarity Score را محاسبه می‌کنیم.

-0.35, -0.5, +0.5, +0.35, -0.35, -0.35

$$\begin{aligned}
 \text{sim score} &= \frac{(-0.35 - 0.5 + 0.5 + 0.35 - 0.35 - 0.35)^2}{3 \times (0.65)(0.35) + 2(0.5)(0.5) + (0.35)(0.65)} \\
 &= \frac{0.49}{1.41} = \underline{0.34}
 \end{aligned}$$

Likes Popcorn	Age	Favorite Color	Loves Troll 2	Prob	Res.
1	12	1	1	0.65	-0.35
1	87	2	1	0.5	-0.5
0	44	1	0	0.5	+0.5
1	19	3	0	0.35	+0.35
0	32	2	1	0.65	-0.35
0	14	1	1	0.65	-0.35

بهین ترتیب جدا هم Split های ممکن  
را بررسی می کنیم و دقیقاً متاب مواردی که گفته شد

درخت جدیدی را می سازیم. از آنجایی که  $St$  و  $x$  می تواند هر مقداری می خواهیم ادامه داشته باشد و  
امکان ندارد همان موارد تا ابد وجود ندارد، همینجا کار را متوقف می کنیم. مراحل الگوریتم به طور  
کامل توضیح داده شده است.