Just like in **Part 1** of this series, we start with a leaf that represents an initial **Prediction** for every individual.

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

When we use **Gradient Boost for Classification**, the initial **Prediction** for every individual is the **log(odds)**.

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

Inshorts
Should a person sleep for 8 hours or divide it into two phases?

When we use **Gradient Boost for Classification**, the initial **Prediction** for every individual is the **log(odds)**.

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

I like to think of the **log(odds)** as the **Logistic Regression** equivalent of the average.

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

…then the **log(odds)** that someone **Loves Troll 2** is…

$$\log\left(\frac{4}{2}\right) = 0.7$$

$\log(4/2) = 0.7$ ← So this is the **Initial Prediction**.

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

log(4/2) = 0.7

Just like with **Logistic Regression**, the easiest way to use the **log(odds)** for **Classification** is to convert it to a **Probability**…

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

log(4/2) = 0.7

Just like with **Logistic Regression**, the easiest way to use the **log(odds)** for **Classification** is to convert it to a **Probability**…

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

…and we do that with a **Logistic Function**.

Probability of Loving Troll 2 $= \dfrac{e^{\log(\text{odds})}}{1 + e^{\log(\text{odds})}}$

$\log(4/2) = 0.7$

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

…and we get **0.7** as the **Probability of Loving Troll 2**.

$$\textbf{Probability of Loving Troll 2} = \frac{e^{\log(4/2)}}{1 + e^{\log(4/2)}} = \boxed{0.7}$$

log(4/2) = 0.7

Probability of **Loving Troll 2** = 0.7

Since the **Probability** of **Loving Troll 2** is greater than **0.5**, we can **Classify** everyone in the **Training Dataset** as someone who **Loves Troll 2**.

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

log(4/2) = 0.7

Probability of **Loving Troll 2** = 0.7

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

Since the **Probability** of **Loving Troll 2** is greater than **0.5**, we can **Classify** everyone in the **Training Dataset** as someone who **Loves Troll 2**.

NOTE: While **0.5** is a very common threshold for making **Classification** decisions based on **Probability**, we could have just as easily used a different value.

log(4/2) = 0.7

Probability of **Loving Troll 2** = 0.7

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

Now, **Classifying** everyone in the **Training Dataset** as someone who **Loves Troll 2** is pretty lame, because two of the people do not love the movie.

log(4/2) = 0.7

Probability of **Loving Troll 2** = 0.7

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

We can measure how bad the initial **Prediction** is by calculating **Pseudo Residuals**, the difference between the **Observed** and the **Predicted** values.

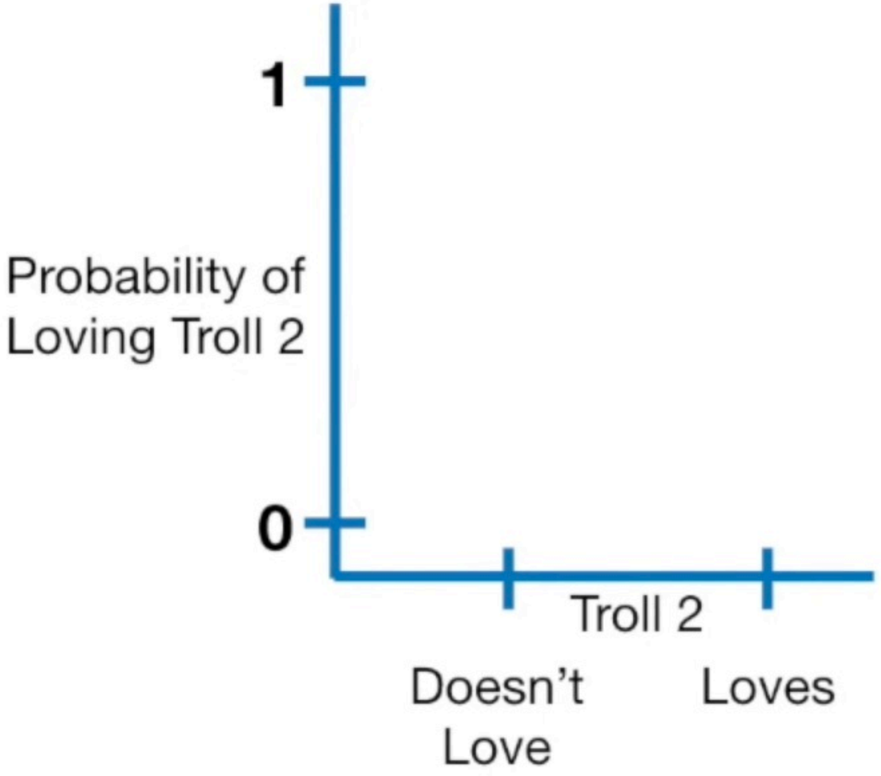
Residual = (**Observed** - **Predicted**)

log(4/2) = 0.7

Probability of **Loving Troll 2** = 0.7

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

Although the math is easy, I think it's easier to grasp what's going on if we draw the **Residuals** on a graph.

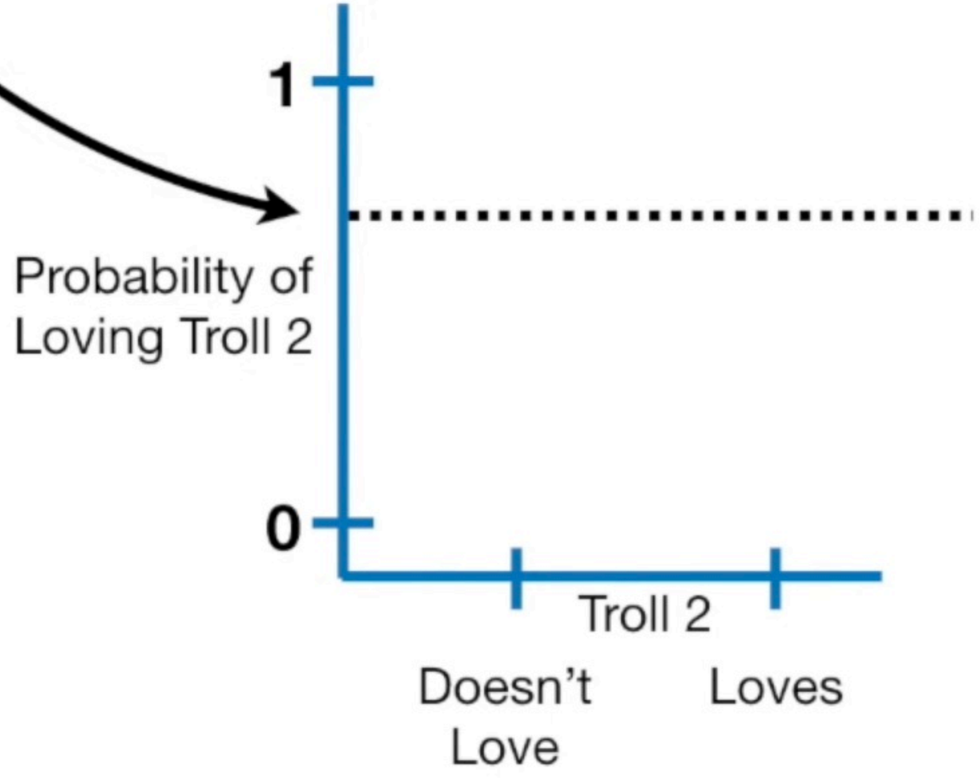


log(4/2) = 0.7

Probability of **Loving Troll 2** = 0.7

The **Predicted Probability** of **Loving Troll 2** is **0.7**.

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

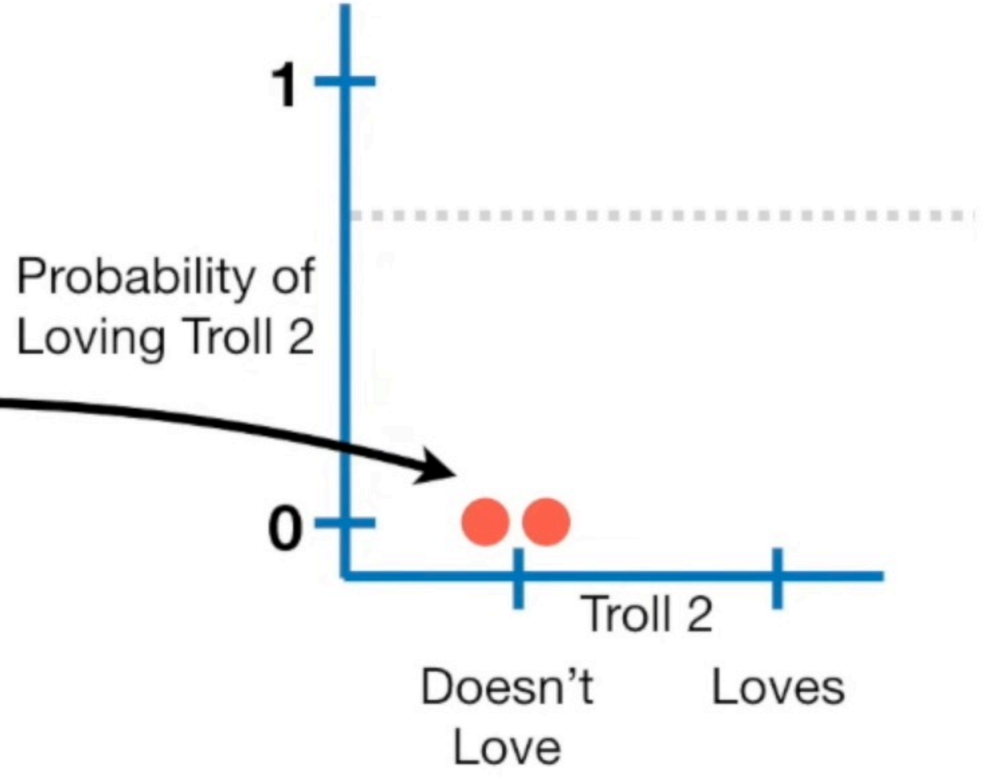


log(4/2) = 0.7

Probability of **Loving Troll 2** = 0.7

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

The **Red Dots**, with the **Probability of Loving Troll 2 = 0**, represent the two people that ***do not*** **Love Troll 2**…

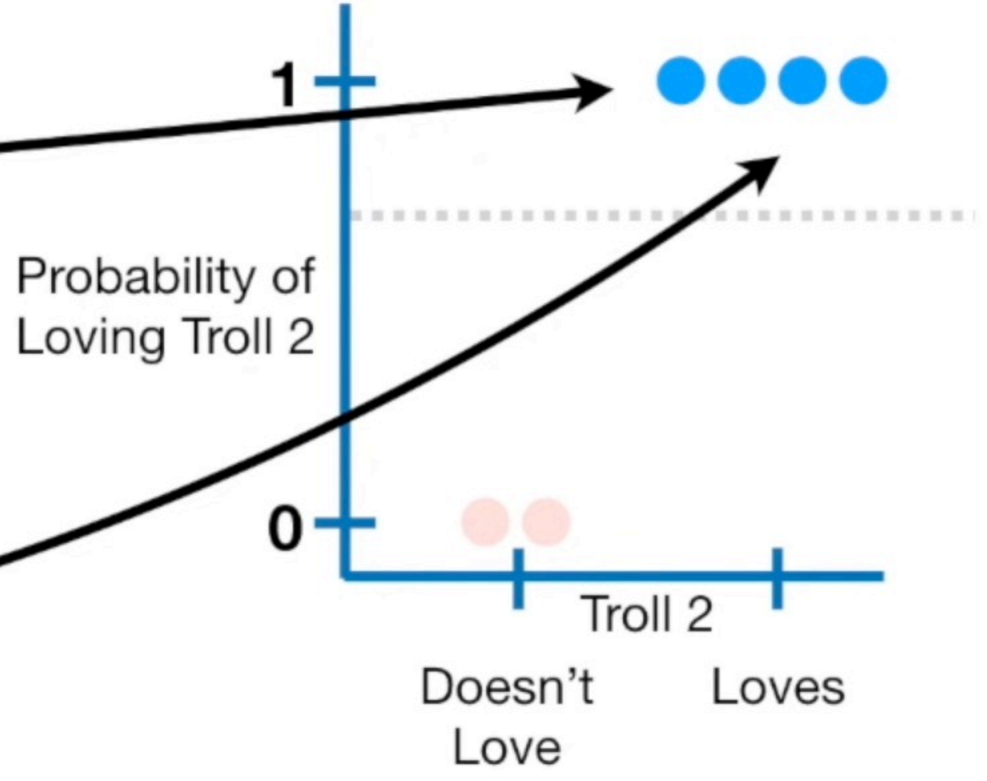


log(4/2) = 0.7

Probability of **Loving Troll 2** = 0.7

…and the **Blue Dots**, with a **Probability of Loving Troll 2 = 1**, represent the four people that **Love Troll 2**.

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

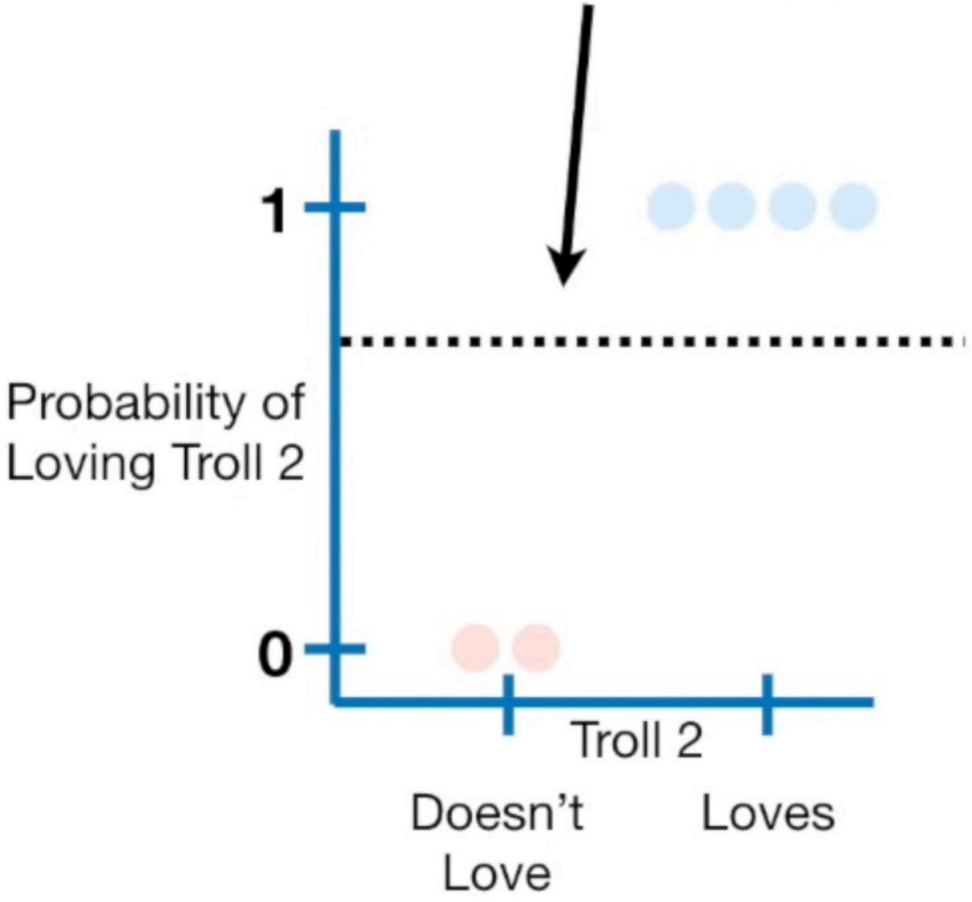


log(4/2) = 0.7

Probability of **Loving Troll 2** = 0.7

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

…and the dotted line is the **Predicted** value.



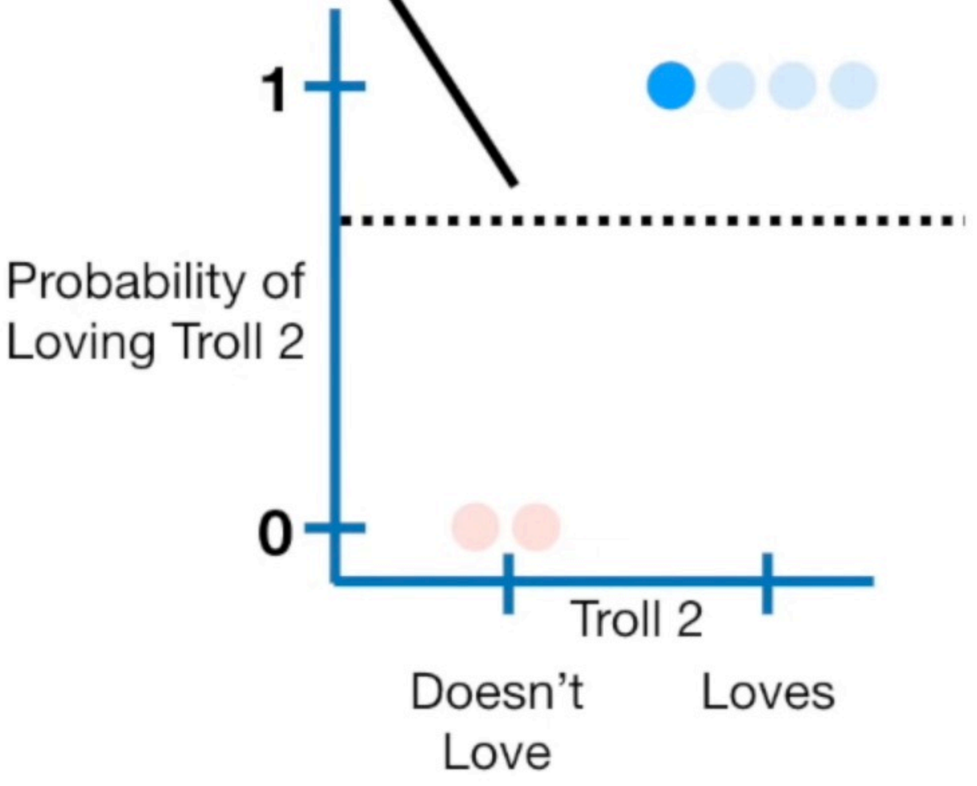
log(4/2) = 0.7

Probability of **Loving Troll 2** = 0.7

…and **0.7** for the **Predicted** value…

Residual = (**1** - **0.7**)

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



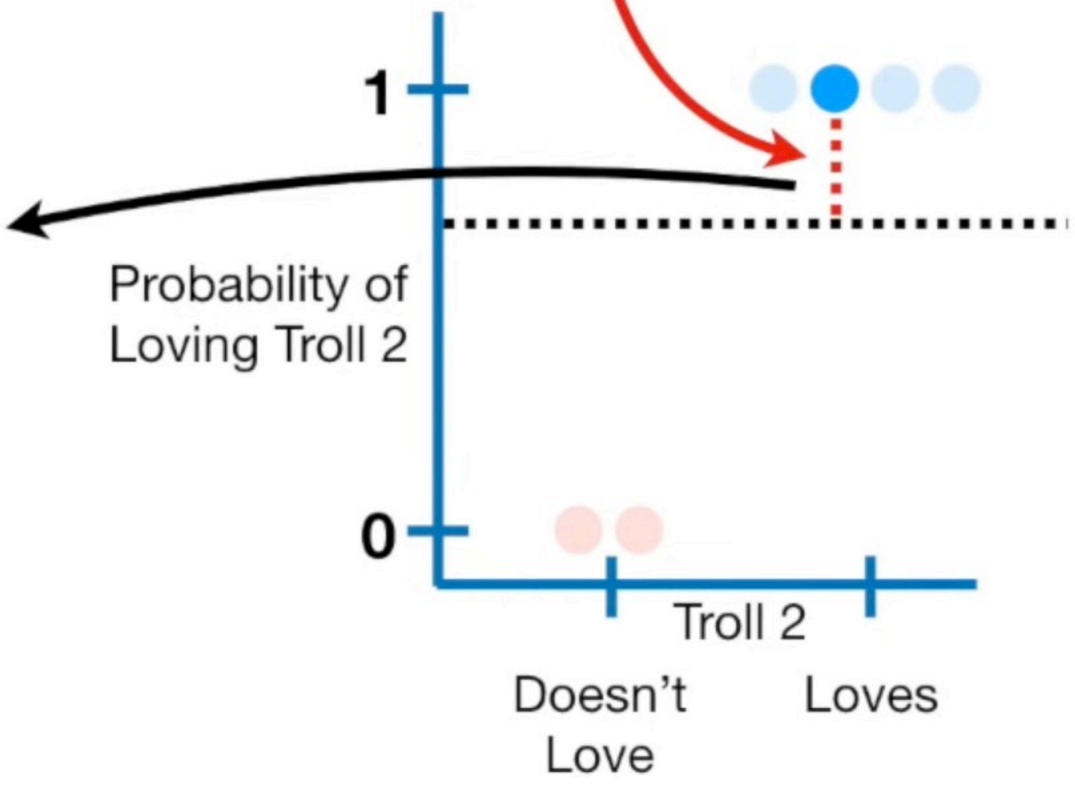
log(4/2) = 0.7

Probability of **Loving Troll 2** = 0.7

Then we calculate the rest of the **Residuals**…

Residual = (**1** - **0.7**) = 0.3

Likes Popcorn	Age	Favorite Color	Loves Troll 2	Residual
Yes	12	Blue	Yes	0.3
Yes	87	Green	Yes	0.3
No	44	Blue	No	
Yes	19	Red	No	
No	32	Green	Yes	
No	14	Blue	Yes	



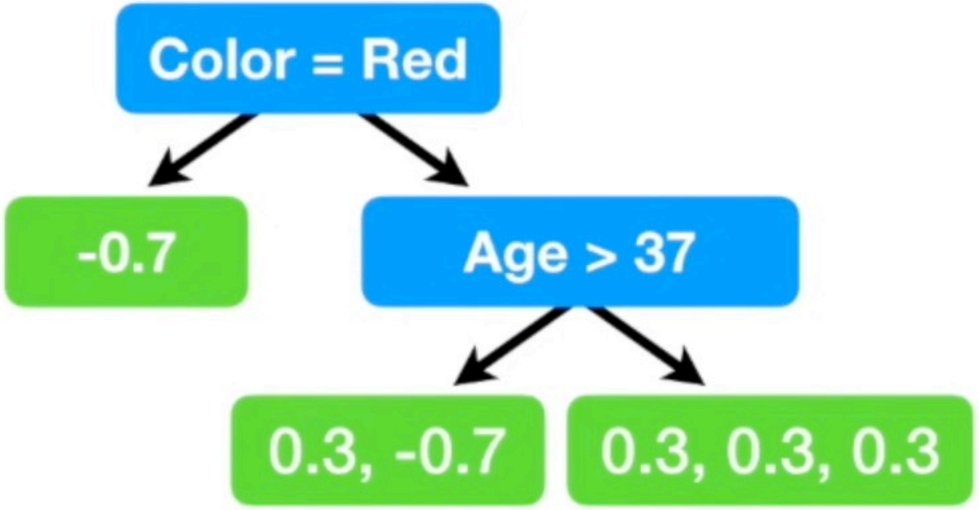
log(4/2) = 0.7

Probability of **Loving Troll 2** = 0.7

Likes Popcorn	Age	Favorite Color	Loves Troll 2	Residual
Yes	12	Blue	Yes	0.3
Yes	87	Green	Yes	0.3
No	44	Blue	No	-0.7
Yes	19	Red	No	-0.7
No	32	Green	Yes	0.3
No	14	Blue	Yes	0.3

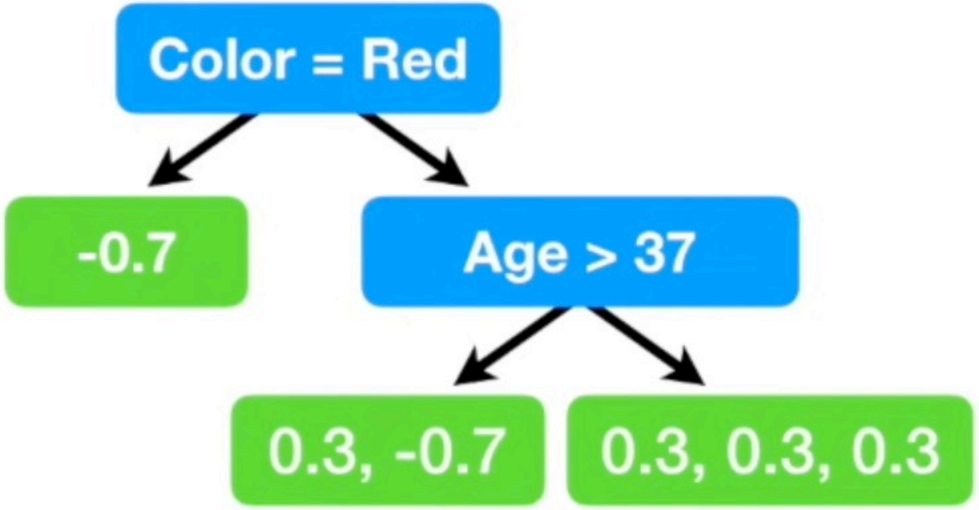
Hooray! We've calculated the **Residuals** for the leaf's initial **Prediction**.

Likes Popcorn	Age	Favorite Color	Loves Troll 2	Residual
Yes	12	Blue	Yes	0.3
Yes	87	Green	Yes	0.3
No	44	Blue	No	-0.7
Yes	19	Red	No	-0.7
No	32	Green	Yes	0.3
No	14	Blue	Yes	0.3



NOTE: Just like when we used **Gradient Boost for Regression**, we are limiting the number of leaves that we will allow in the tree.

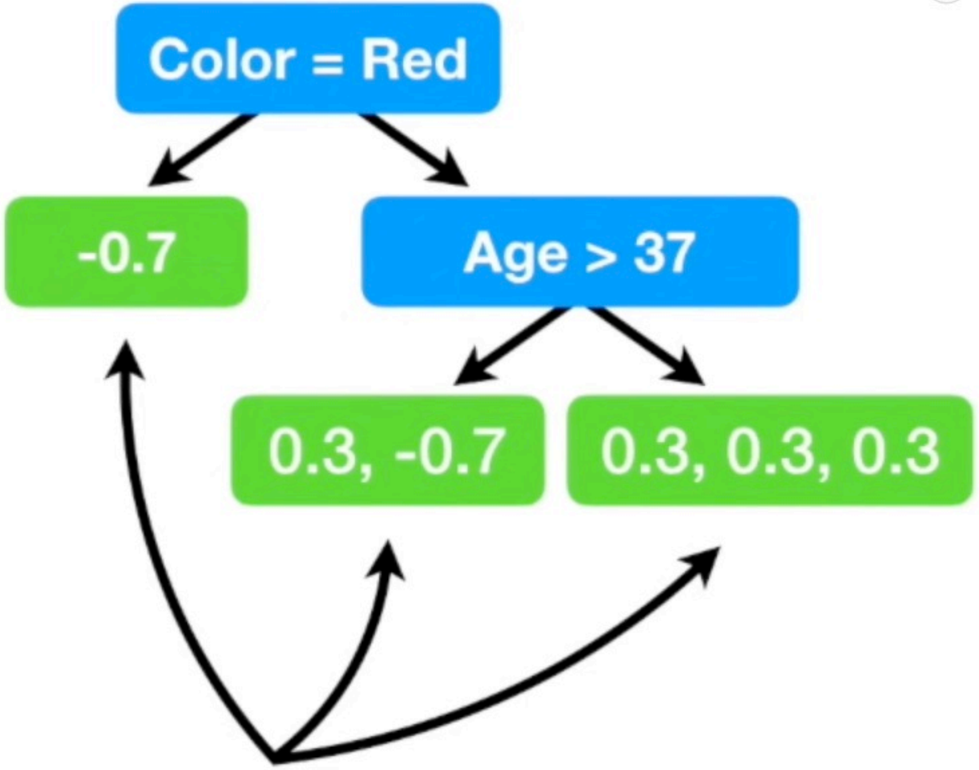
Likes Popcorn	Age	Favorite Color	Loves Troll 2	Residual
Yes	12	Blue	Yes	0.3
Yes	87	Green	Yes	0.3
No	44	Blue	No	-0.7
Yes	19	Red	No	-0.7
No	32	Green	Yes	0.3
No	14	Blue	Yes	0.3



In this simple example, we are limiting the number of leaves to **3**.

In practice people often set the maximum number of leaves to be between **8** and **32**

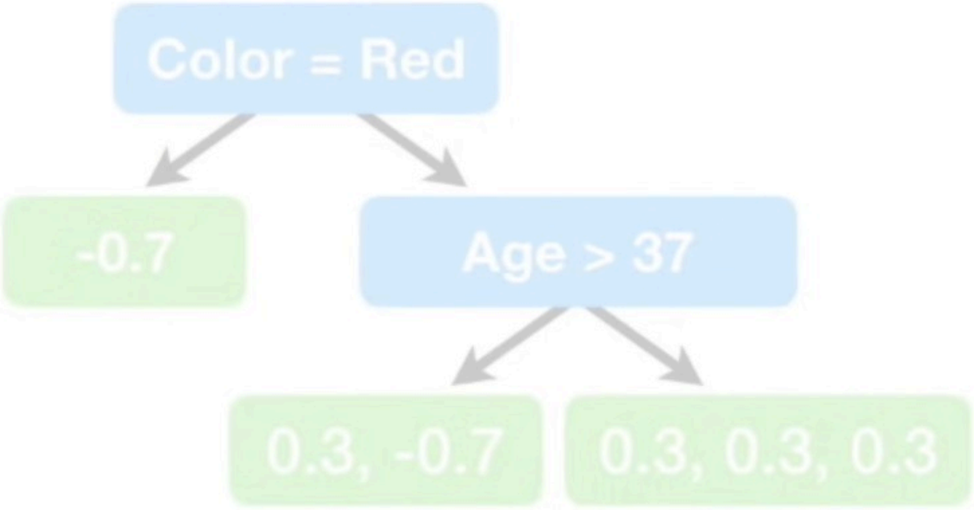
Likes Popcorn	Age	Favorite Color	Loves Troll 2	Residual
Yes	12	Blue	Yes	0.3
Yes	87	Green	Yes	0.3
No	44	Blue	No	-0.7
Yes	19	Red	No	-0.7
No	32	Green	Yes	0.3
No	14	Blue	Yes	0.3



Now let's calculate the **Output Values** for the leaves.

5:51 / 17:02
Trees

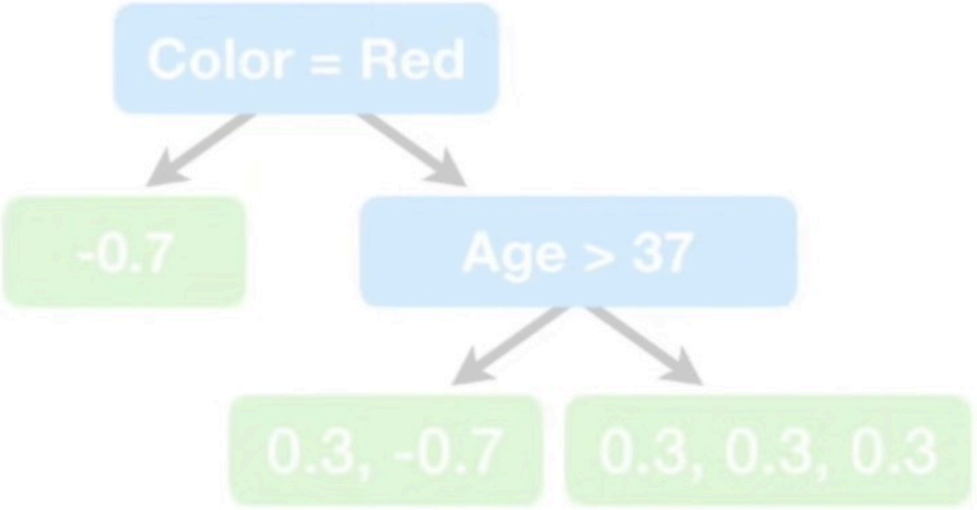
When we used **Gradient Boost** for **Regression**, a leaf with single **Residual** had an **Output Value** equal to that **Residual**.



In contrast, when we use **Gradient Boost** for **Classification**, the situation is a little more complex.

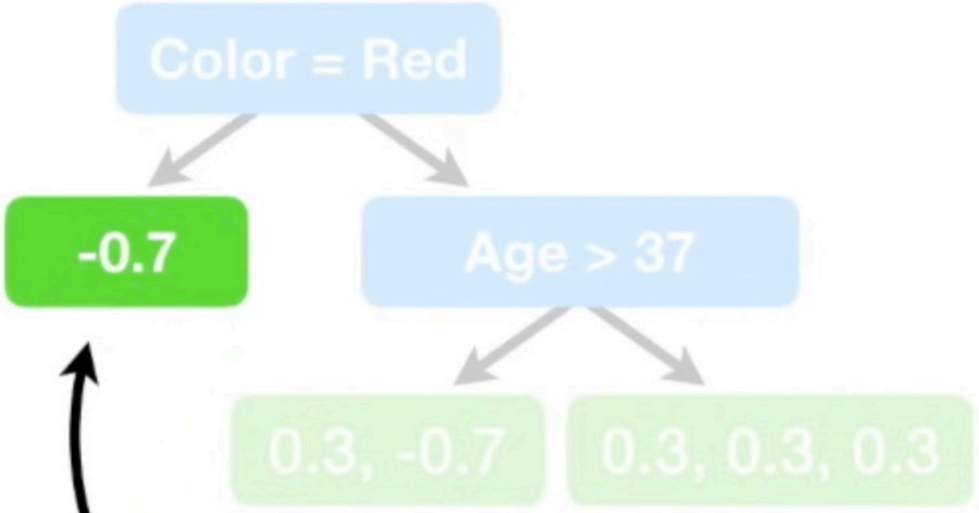
log(4/2) = 0.7

This is because the **Predictions** are in terms of the **log(odds)**…

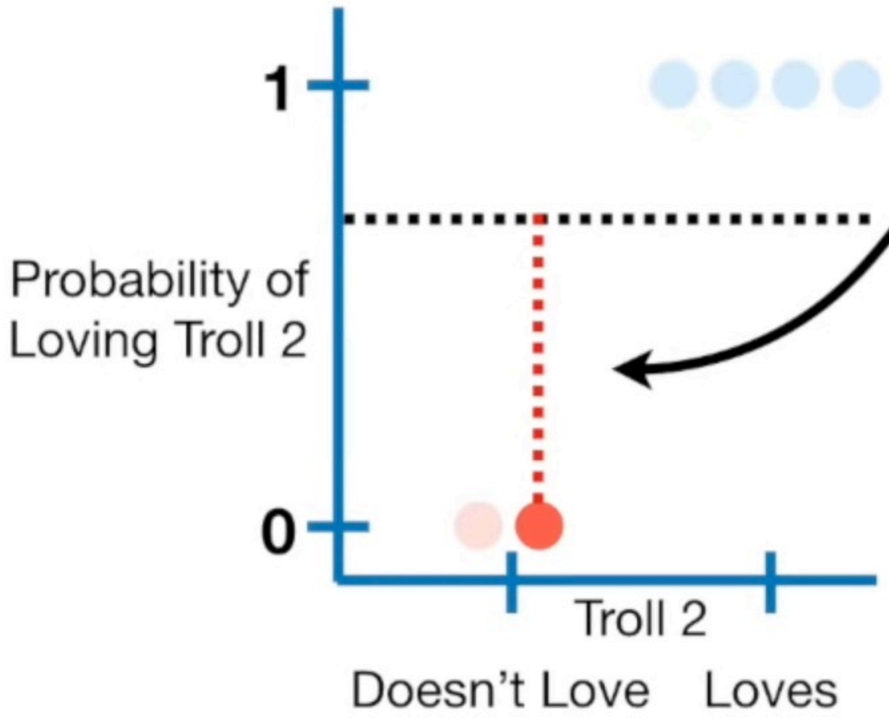


log(4/2) = 0.7

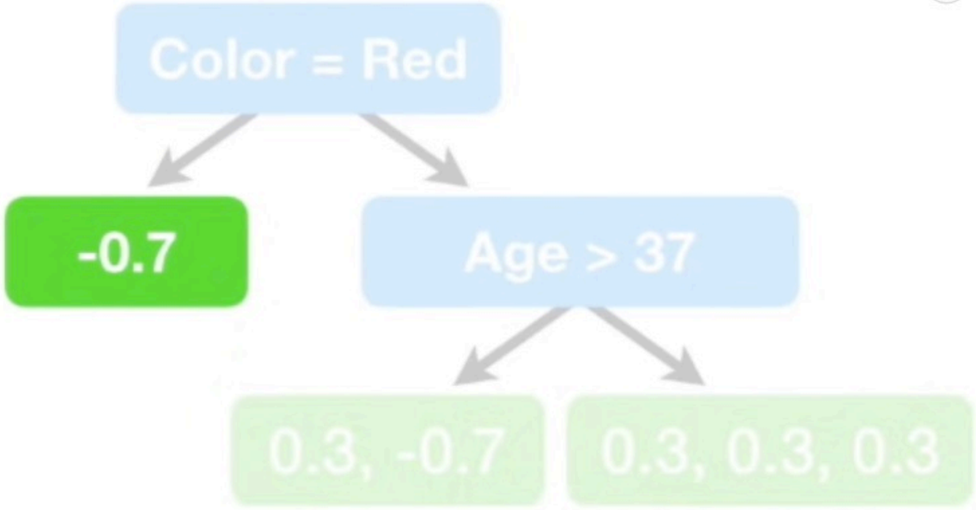
This is because the **Predictions** are in terms of the **log(odds)**…



…and this leaf is derived from a **Probability**…



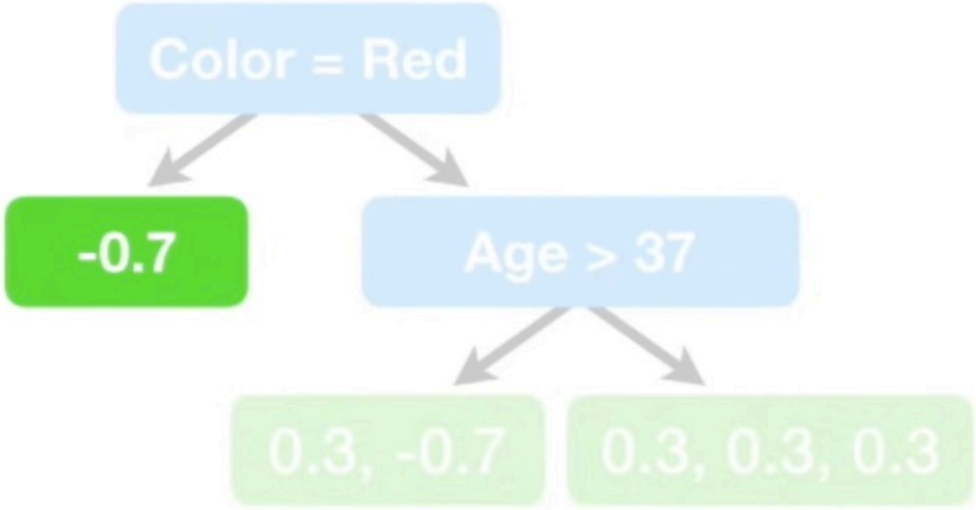
log(4/2) = 0.7

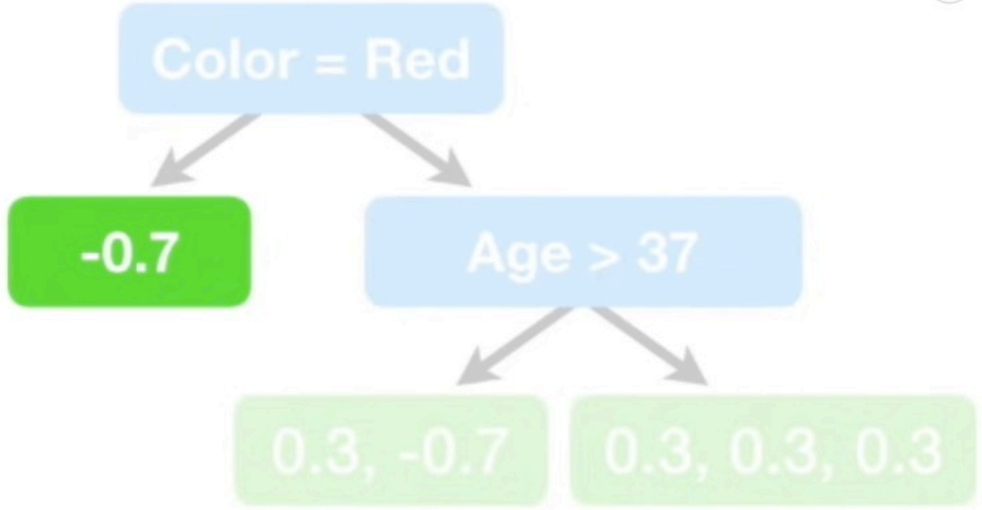


…so we can't just add them together to get a new **log(odds) Prediction** without some sort of transformation.

When we use **Gradient Boost** for **Classification**, the most common transformation is the following formula.

$$\frac{\sum \text{Residual}_i}{\sum \left[\text{Previous Probability}_i \times (1 - \text{Previous Probability}_i)\right]}$$



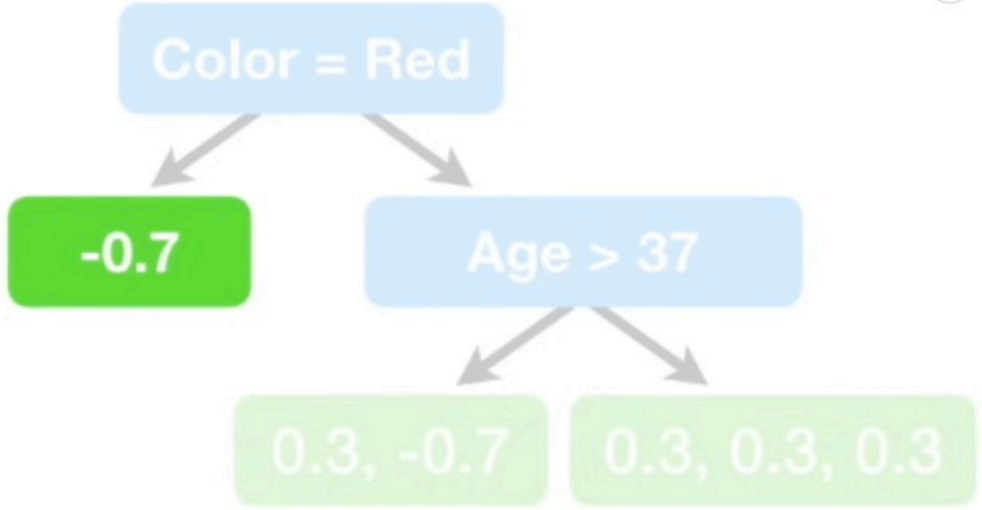


The numerator is the sum of the all of the **Residuals** in the leaf…

$$\frac{\sum \text{Residual}_i}{\sum [\text{Previous Probability}_i \times (1 - \text{Previous Probability}_i)]}$$

6:50 / 17:02
Transformations

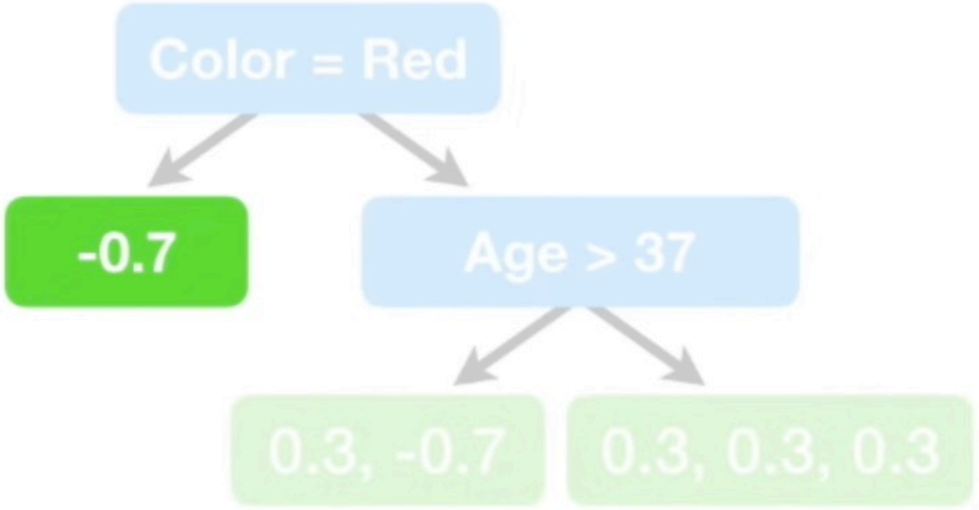
…and the denominator is the sum of the previously predicted probabilities for each **Residual** times **1** minus the same predicted probability.



$$\frac{\sum \text{Residual}_i}{\sum \left[\text{Previous Probability}_i \times (1 - \text{Previous Probability}_i)\right]}$$

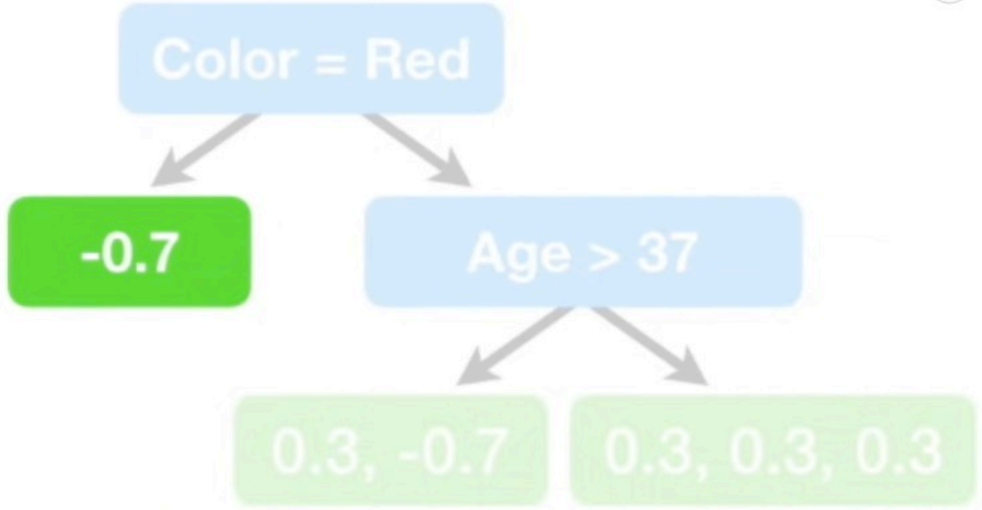
6:58 / 17:02
Transformations

NOTE: The derivation of this formula is quite technical, so I'm saving it for **Part 4** of this series when we get into the nitty gritty details of **Gradient Boost for Classification**.



$$\frac{\sum \text{Residual}_i}{\sum \left[\text{Previous Probability}_i \times (1 - \text{Previous Probability}_i)\right]}$$

Since there is only one **Residual** in this leaf, we can ignore these summation signs for now.

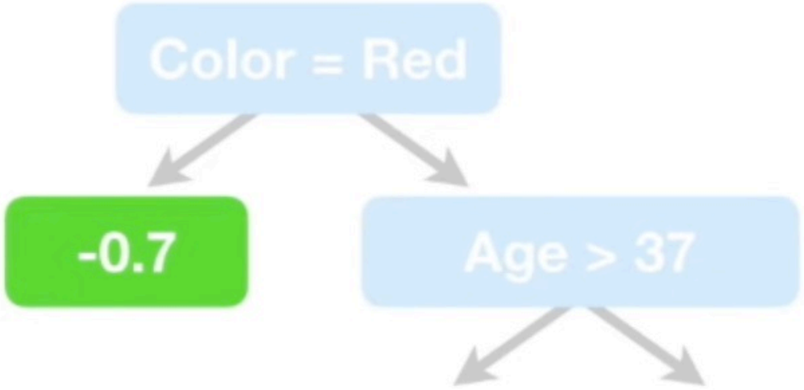


$$\frac{\sum \text{Residual}_i}{\sum \left[\text{Previous Probability}_i \times (1 - \text{Previous Probability}_i)\right]}$$

+5
7:24 / 17:02
Transformations

log(4/2) = 0.7

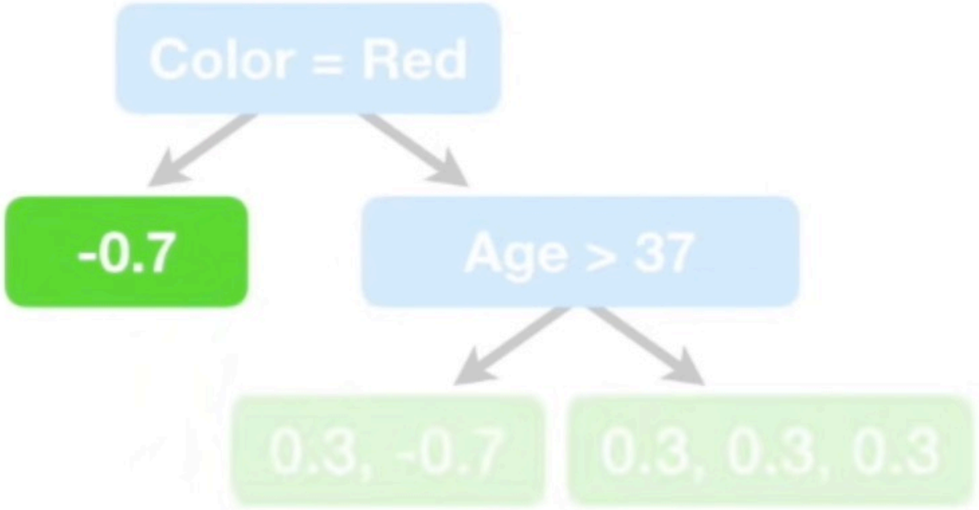
Probability of Loving Troll 2 = 0.7



…and, since we are building the first tree, the **Previous Probability** refers to the probability from the initial leaf…

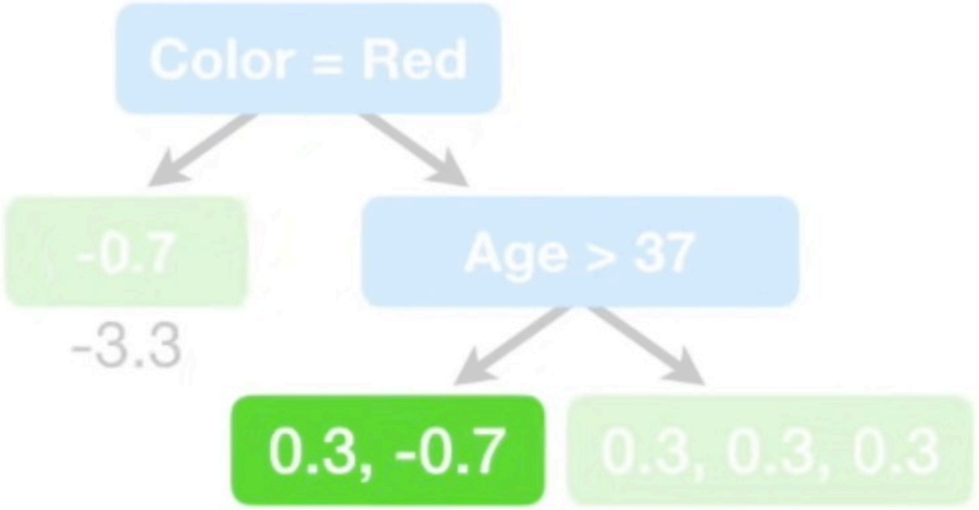
-0.7

$$\sum \left[\text{Previous Probability}_i \times (1 - \text{Previous Probability}_i)\right]$$

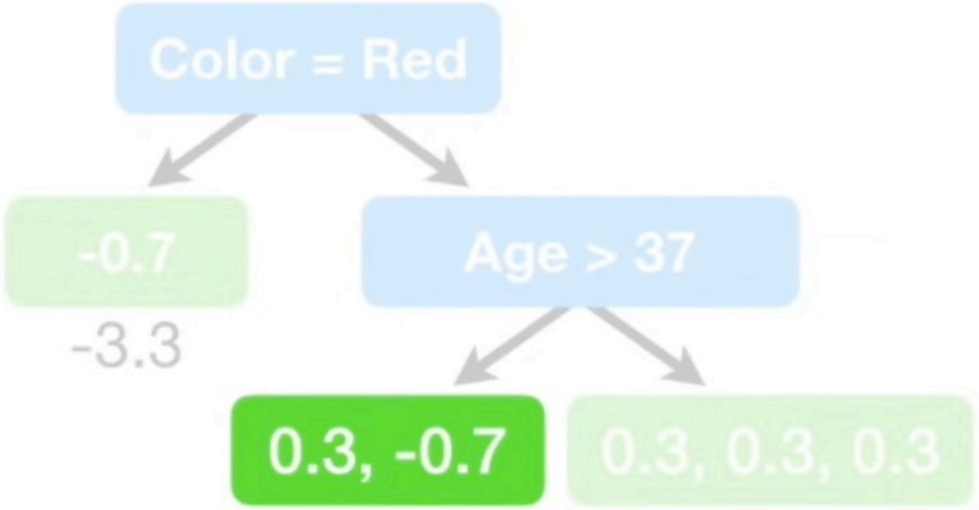


…do the math…

$$\frac{-0.7}{0.7 \times (1 - 0.7)} = -3.3$$

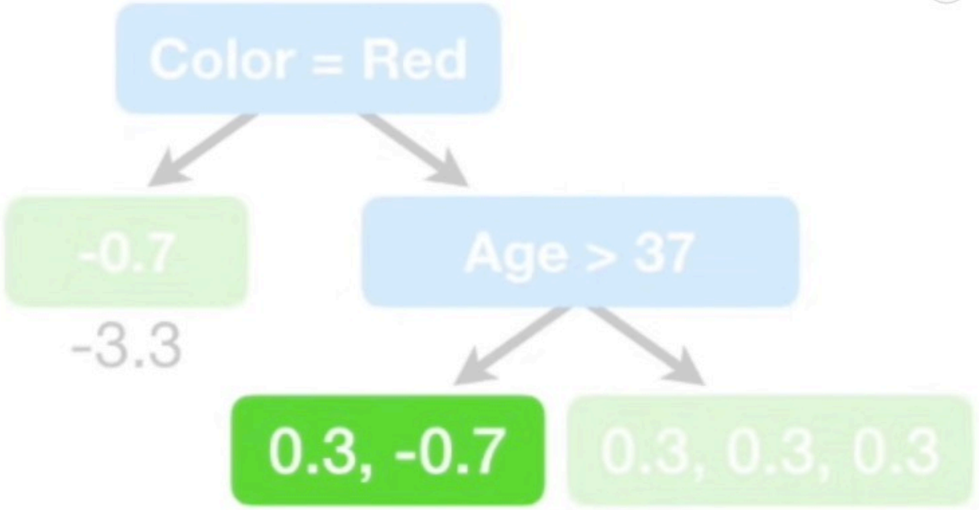


Now we need to calculate the **Output Value** for this leaf.



$$\frac{0.3 + -0.7}{\sum \left[\text{Previous Probability}_i \times (1 - \text{Previous Probability}_i)\right]}$$

Since we have two **Residuals** in the leaf, we'll add them together in the numerator…



$$\frac{0.3 + -0.7}{\sum \left[\text{Previous Probability}_i \times (1 - \text{Previous Probability}_i)\right]}$$

…and in the denominator we just add up…

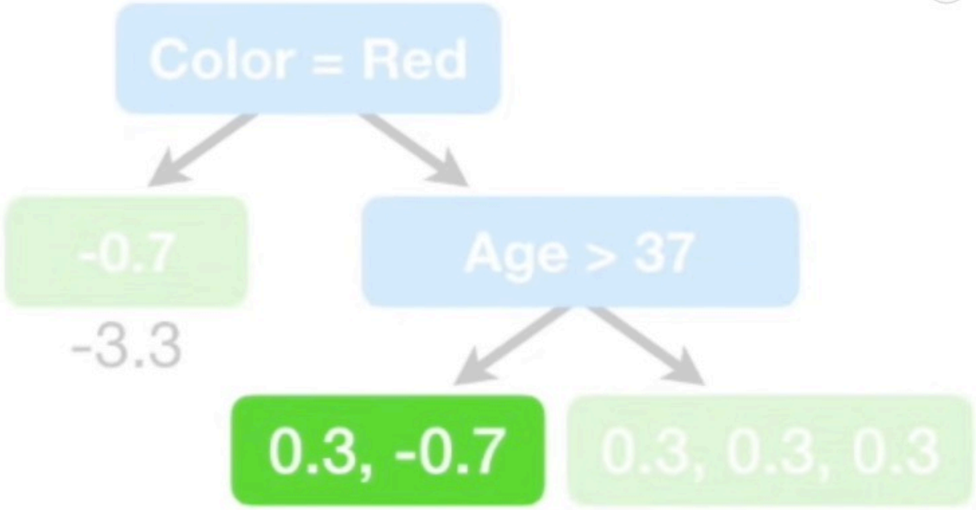
Previous Probability x (1 - Previous Probability)

8:08 / 17:02
Transformations

log(4/2) = 0.7

Initial Probability of **Loving Troll 2** = 0.7

$$\frac{0.3 + -0.7}{(0.7 \times (1 - 0.7)) + (0.7 \times (1 - 0.7))}$$



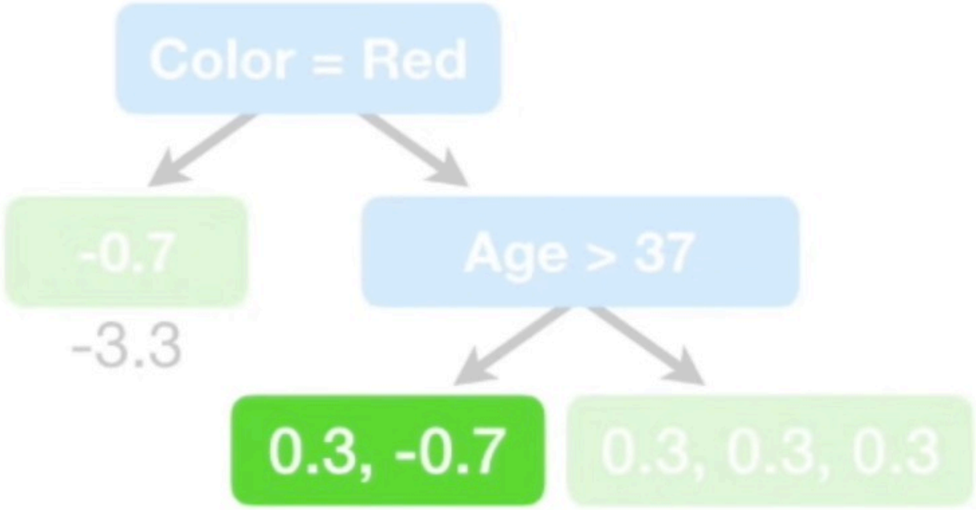
So we plug in the **Previous Probability** for each **Residual**.

8:15 / 17:02
Transformations

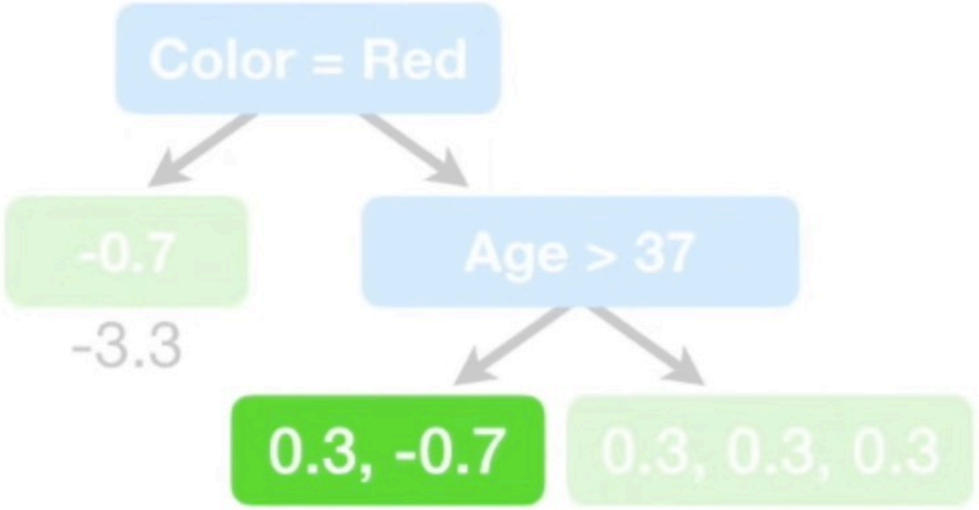
log(4/2) = 0.7

Initial Probability of **Loving Troll 2** = 0.7

$$\frac{0.3 + -0.7}{(0.7 \times (1 - 0.7)) + (0.7 \times (1 - 0.7))}$$

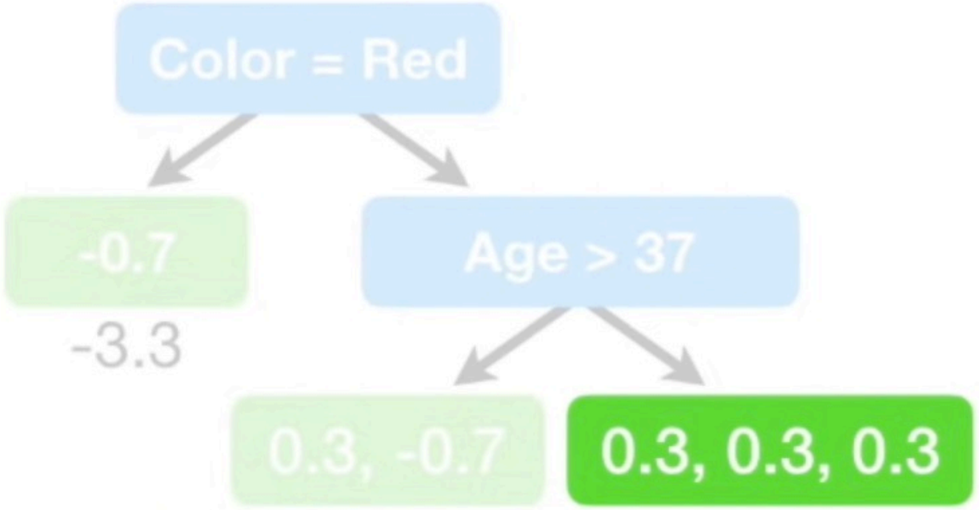


NOTE: For now, the **Previous Probabilities** are the same for all of the **Residuals**, but this will change when we build the next tree.



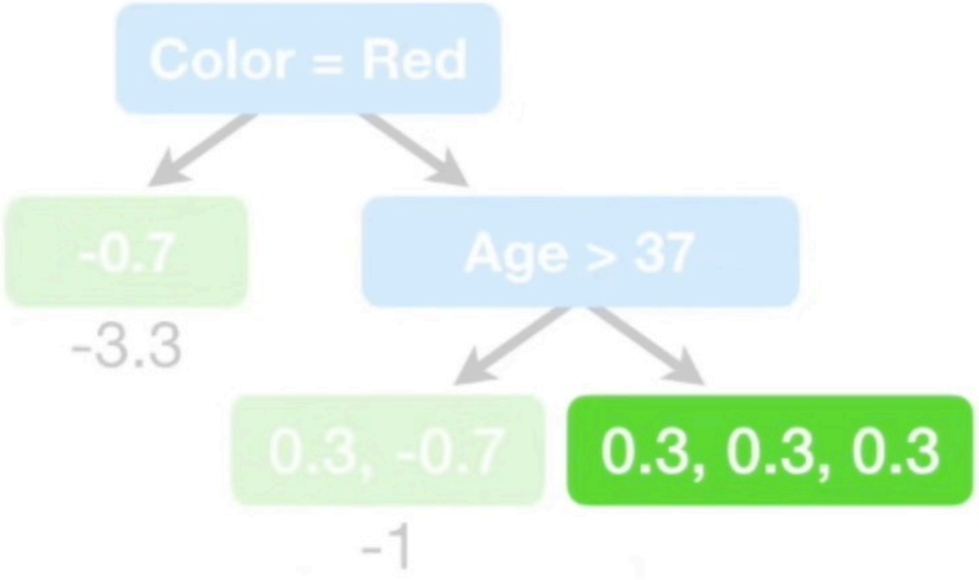
$$\frac{0.3 + -0.7}{2(0.7 \times (1 - 0.7))} = \boxed{-1}$$

…and the **Output Value** for this leaf is **-1**.



$$\frac{\sum \text{Residual}_i}{\sum \text{Previous Probability}_i \times (1 - \text{Previous Probability}_i)}$$

We plug the **Residuals** into the formula…



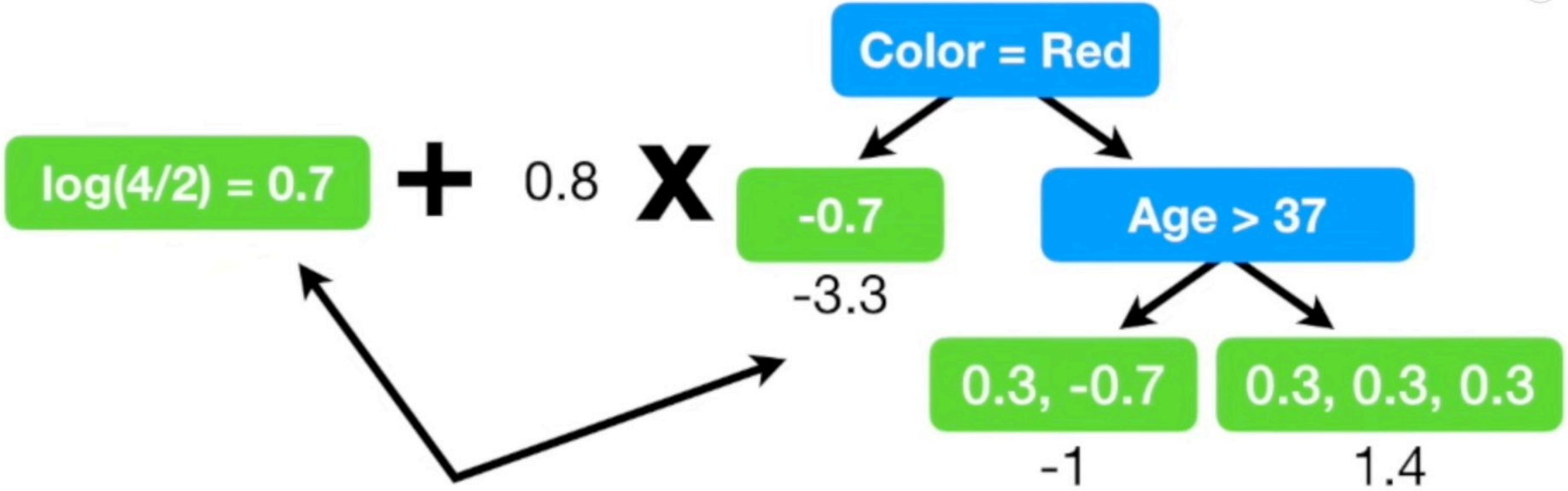
$$\frac{3 \times 0.3}{3 \times (0.7 \times (1 - 0.7))}$$

…and do the math…

Want to ace some interviews?
www.scaler.com

$$\frac{0.3}{(0.7 \times (1 - 0.7))} = \boxed{1.4}$$

…and the **Output Value** for this leaf is **1.4**.

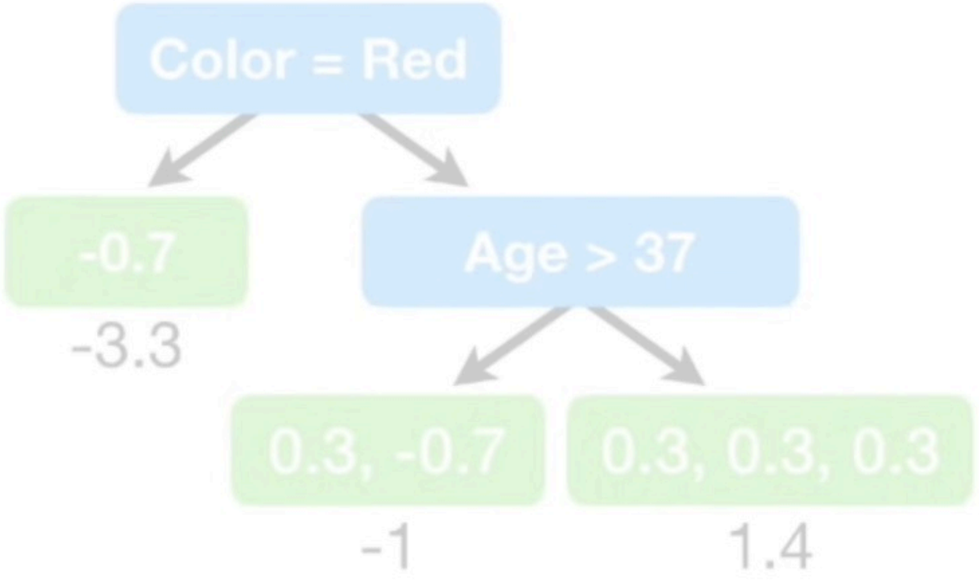


Now we are ready to update our **Predictions** by combining the initial leaf with the new tree.

9:02 / 17:02
Updating predictions

NOTE: Just like before, the new tree is scaled by a **Learning Rate**.

log(4/2) = 0.7 + 0.8 X

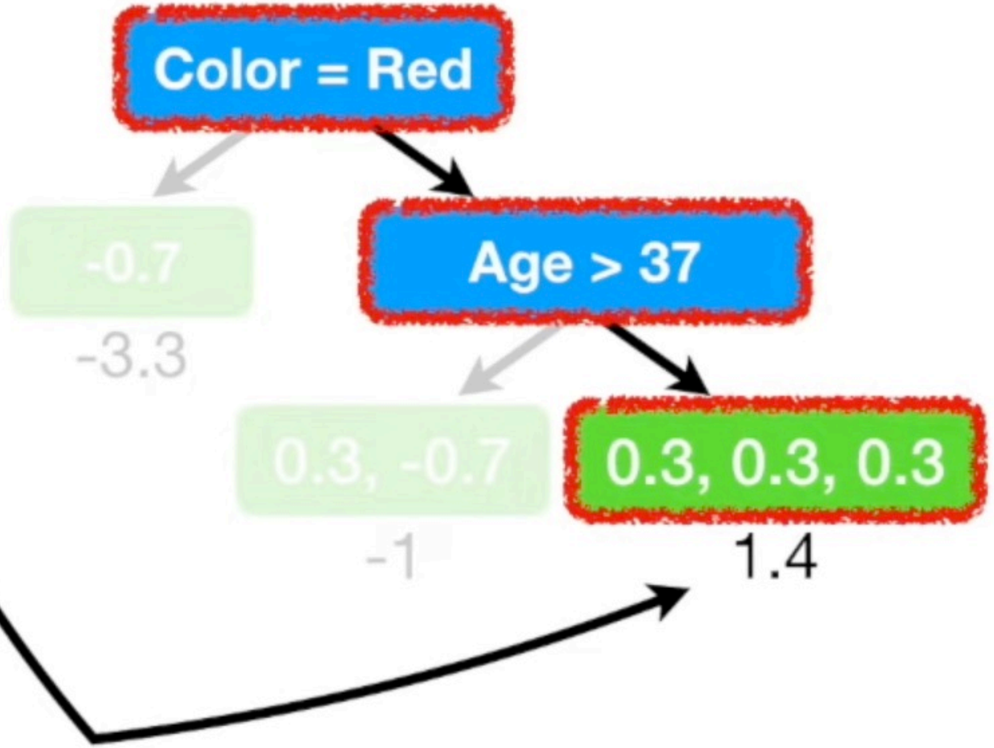


Likes Popcorn	Age	Favorite Color	Loves Troll 2
Yes	12	Blue	Yes
Yes	87	Green	Yes
No	44	Blue	No
Yes	19	Red	No
No	32	Green	Yes
No	14	Blue	Yes

The **log(odds) Prediction** is the previous **Prediction**…

log(odds) Prediction = 0.7

log(4/2) = 0.7 + 0.8 X



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

…plus the **Output Value** from the tree scaled by the **Learning Rate**…

log(odds) Prediction = 0.7 + (0.8 × 1.4)

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

Now we convert the new **log(odds) Prediction** into a **Probability**…

$$\text{Probability} = \frac{e^{1.8}}{1 + e^{1.8}} = 0.9$$

log(odds) Prediction = 0.7 + (0.8 × 1.4) = 1.8

log(4/2) = 0.7

Initial Probability of **Loving Troll 2** = 0.7

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

…so we are taking a small step in the right direction since this person **Loves Troll 2**.

$$\text{Probability} = \frac{e^{1.8}}{1 + e^{1.8}} = \boxed{0.9}$$

log(odds) Prediction = 0.7 + (0.8 × 1.4) = 1.8

Likes Popcorn	Age	Favorite Color	Loves Troll 2	Predicted Prob.
Yes	12	Blue	Yes	0.9
Yes	87	Green	Yes	
No	44	Blue	No	
Yes	19	Red	No	
No	32	Green	Yes	
No	14	Blue	Yes	

Now we convert the new **log(odds) Prediction** into a **Probability**…

$$\text{Probability} = \frac{e^{\log(\text{odds})}}{1 + e^{\log(\text{odds})}}$$

$$\log(\text{odds}) \text{ Prediction} = 0.7 + (0.8 \times -1) = -0.1$$

Likes Popcorn	Age	Favorite Color	Loves Troll 2	Predicted Prob.
Yes	12	Blue	Yes	0.9
Yes	87	Green	Yes	0.5
No	44	Blue	No	
Yes	19	Red	No	
No	32	Green	Yes	
No	14	Blue	Yes	

…and save the new **Predicted Probability**, **0.5**, here…

$$\text{Probability} = \frac{e^{-0.1}}{1 + e^{-0.1}} = \boxed{0.5}$$

log(odds) Prediction = 0.7 + (0.8 × -1) = -0.1

Likes Popcorn	Age	Favorite Color	Loves Troll 2	Predicted Prob.
Yes	12	Blue	Yes	0.9
Yes	87	Green	Yes	0.5
No	44	Blue	No	
Yes	19	Red	No	
No	32	Green	Yes	
No	14	Blue	Yes	

NOTE: This new predicted probability is worse than before, and this is one reason why we build a lot of trees, and not just one.

$$\text{Probability} = \frac{e^{-0.1}}{1 + e^{-0.1}} = \boxed{0.5}$$

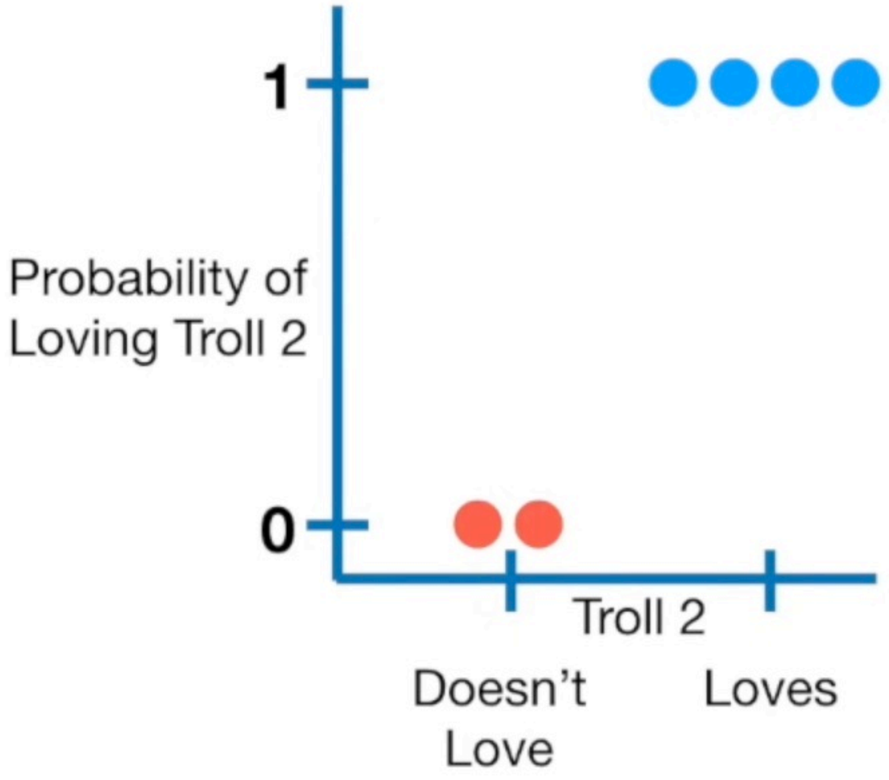
log(odds) Prediction = 0.7 + (0.8 × -1) = -0.1

…and just like before, **Residuals** are the difference between the **Observed** and **Predicted Probabilities**… → **Residual** = (**Observed** - **Predicted**)

Likes Popcorn	Age	Favorite Color	Loves Troll 2	Predicted Prob.	Residual
Yes	12	Blue	Yes	0.9	
Yes	87	Green	Yes	0.5	
No	44	Blue	No	0.5	
Yes	19	Red	No	0.1	
No	32	Green	Yes	0.9	
No	14	Blue	Yes	0.9	

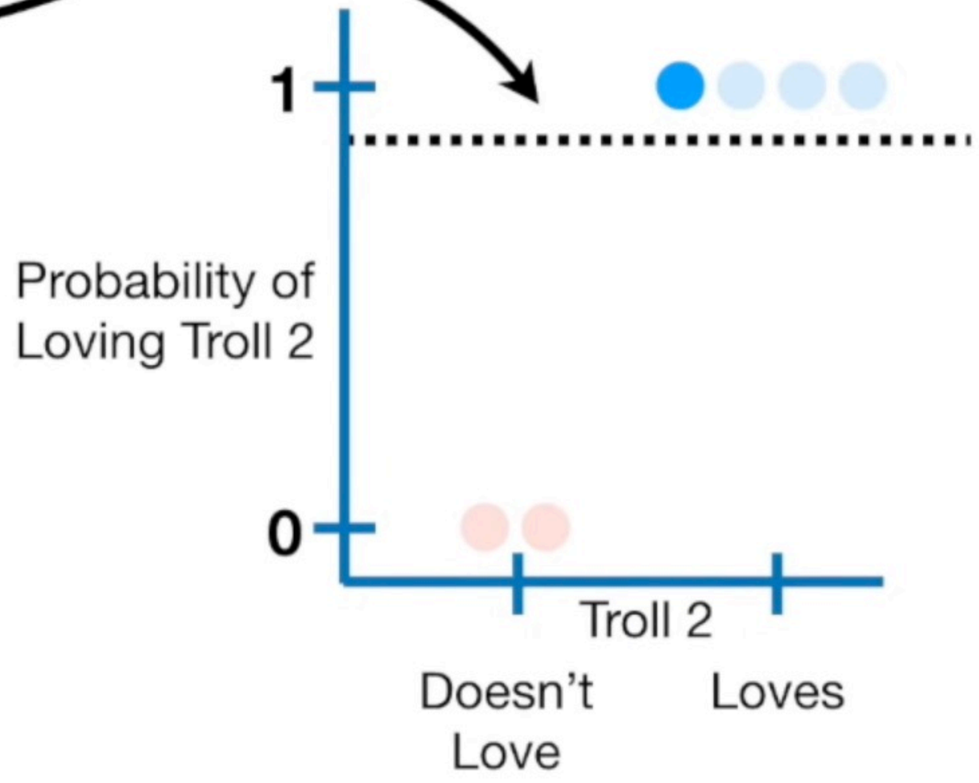
…however, now everyone has a different **Predicted Probability**.

Likes Popcorn	Age	Favorite Color	Loves Troll 2	Predicted Prob.	Residual
Yes	12	Blue	Yes	0.9	
Yes	87	Green	Yes	0.5	
No	44	Blue	No	0.5	
Yes	19	Red	No	0.1	
No	32	Green	Yes	0.9	
No	14	Blue	Yes	0.9	



We plot the **Predicted Probability**…

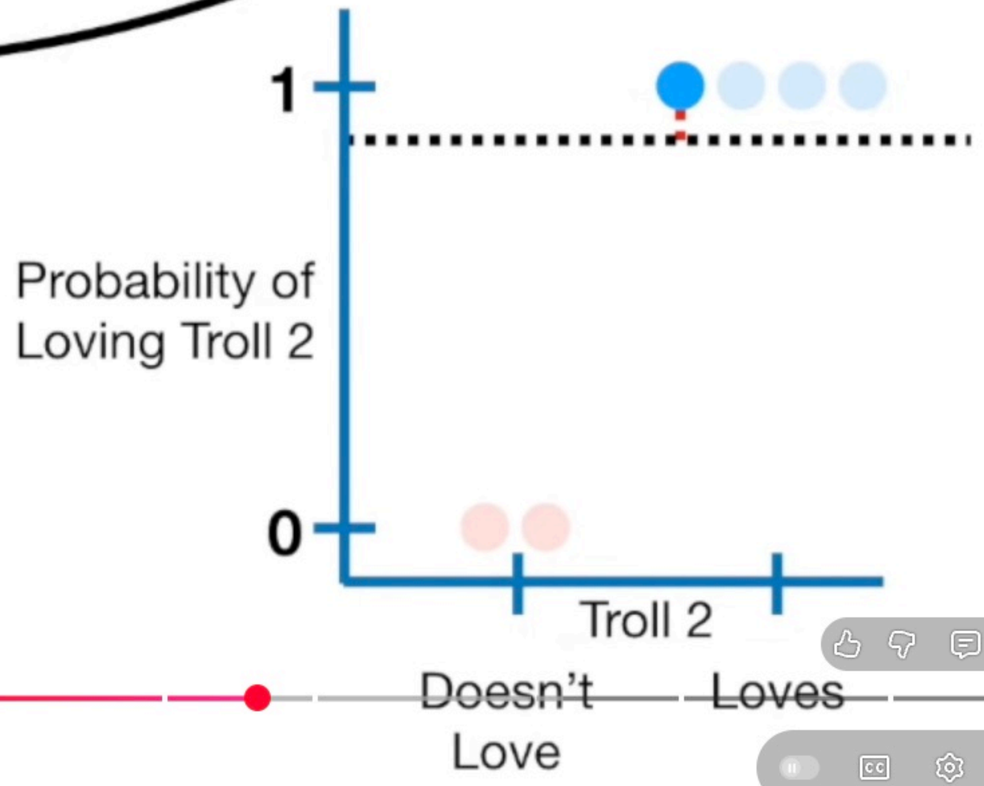
Likes Popcorn	Age	Favorite Color	Loves Troll 2	Predicted Prob.	Residual
Yes	12	Blue	Yes	0.9	
Yes	87	Green	Yes	0.5	
No	44	Blue	No	0.5	
Yes	19	Red	No	0.1	
No	32	Green	Yes	0.9	
No	14	Blue	Yes	0.9	



And we save that value here.

Residual = (**1** - **0.9**) = 0.1

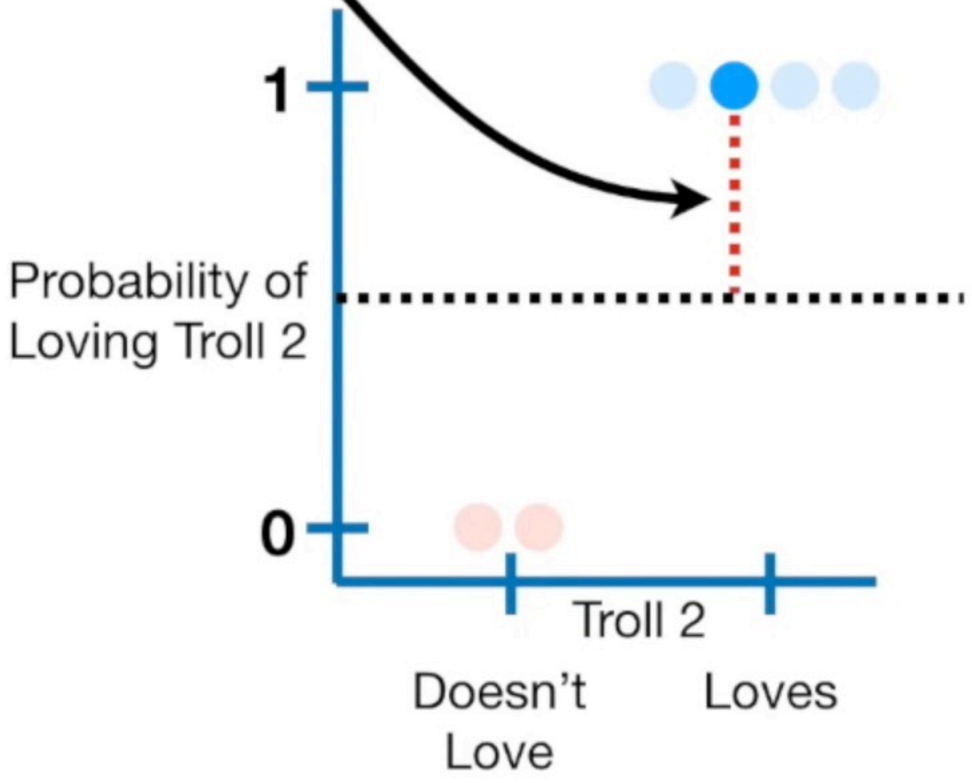
Likes Popcorn	Age	Favorite Color	Loves Troll 2	Predicted Prob.	Residual
Yes	12	Blue	Yes	0.9	0.1
Yes	87	Green	Yes	0.5	
No	44	Blue	No	0.5	
Yes	19	Red	No	0.1	
No	32	Green	Yes	0.9	
No	14	Blue	Yes	0.9	



11:34 / 17:02
Residual

…and the **Residual** is the difference. → **Residual** = (**1** - **0.5**)

Likes Popcorn	Age	Favorite Color	Loves Troll 2	Predicted Prob.	Residual
Yes	12	Blue	Yes	0.9	0.1
Yes	87	Green	Yes	0.5	
No	44	Blue	No	0.5	
Yes	19	Red	No	0.1	
No	32	Green	Yes	0.9	
No	14	Blue	Yes	0.9	

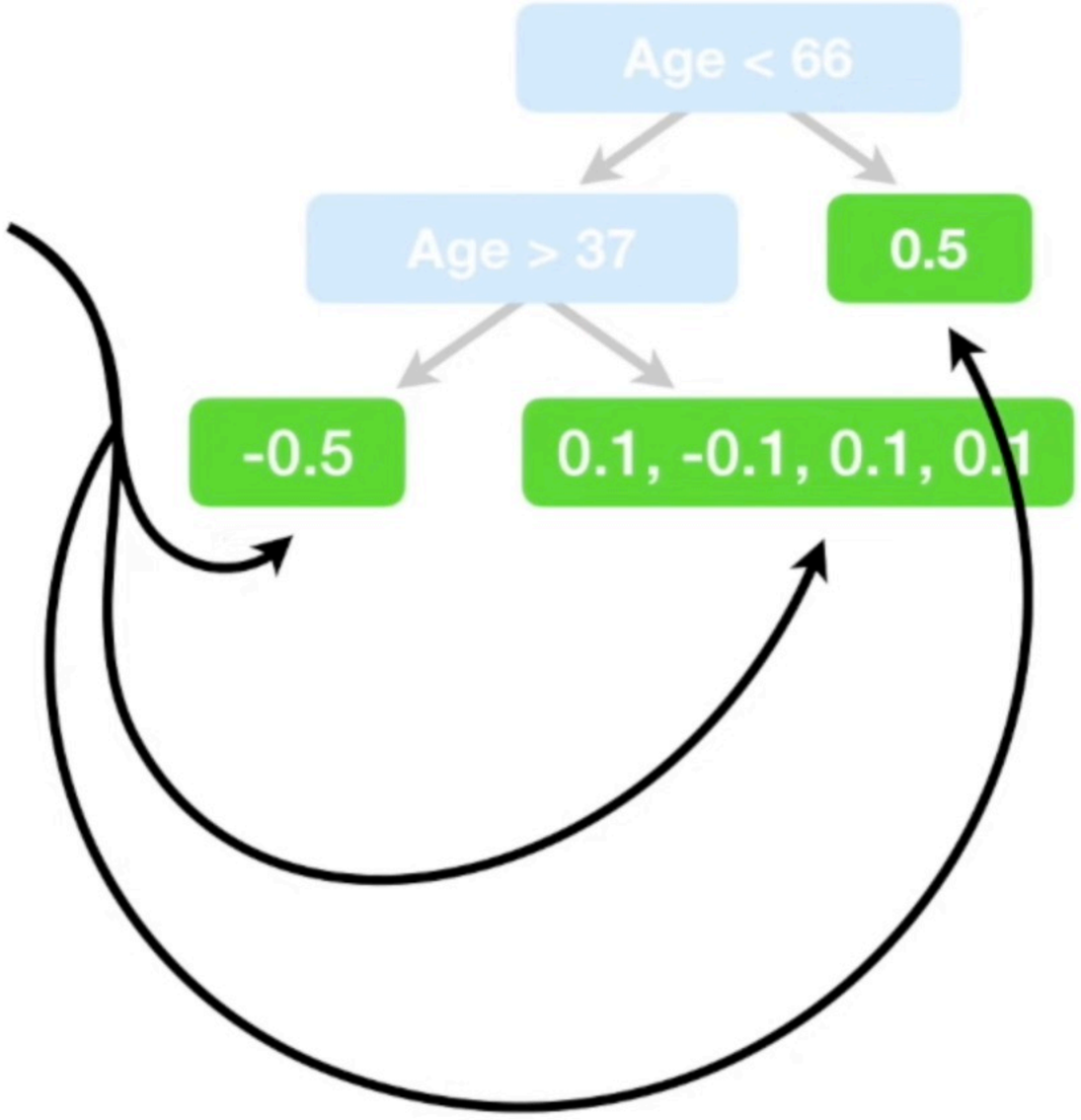


Likes Popcorn	Age	Favorite Color	Loves Troll 2	Predicted Prob.	Residual
Yes	12	Blue	Yes	0.9	0.1
Yes	87	Green	Yes	0.5	0.5
No	44	Blue	No	0.5	-0.5
Yes	19	Red	No	0.1	-0.1
No	32	Green	Yes	0.9	0.1
No	14	Blue	Yes	0.9	0.1

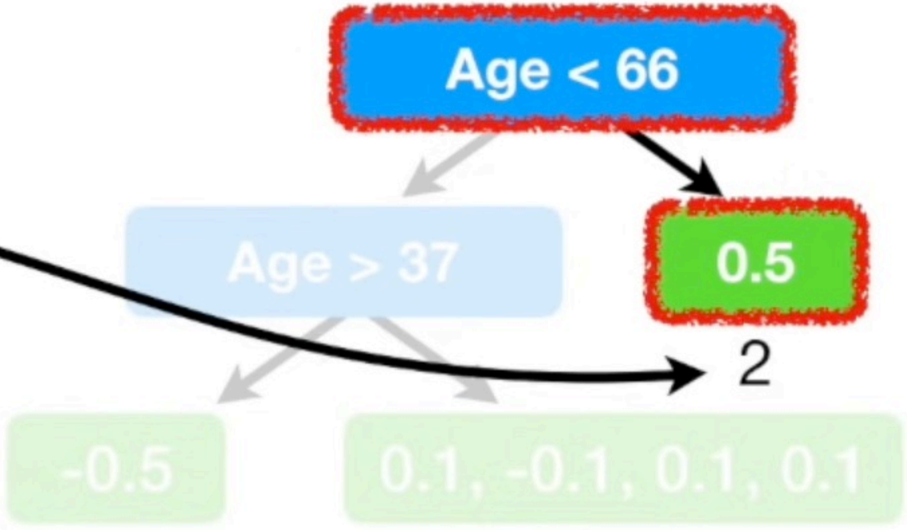
BAM!

…and then we need to calculate the **Output Values** for each leaf.

Likes Popcorn	Age	Favorite Color	Loves Troll 2	Predicted Prob.	Residual
Yes	12	Blue	Yes	0.9	0.1
Yes	87	Green	Yes	0.5	0.5
No	44	Blue	No	0.5	-0.5
Yes	19	Red	No	0.1	-0.1
No	32	Green	Yes	0.9	0.1
No	14	Blue	Yes	0.9	0.1



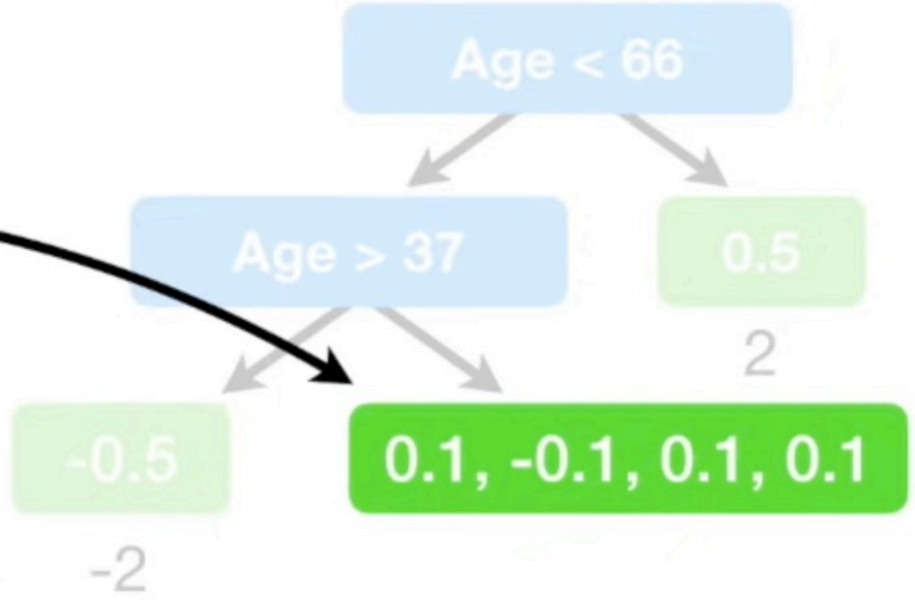
$$\frac{0.5}{0.5 \times (1 - 0.5)} = \boxed{2}$$



Likes Popcorn	Age	Favorite Color	Loves Troll 2	Predicted Prob.	Residual
Yes	12	Blue	Yes	0.9	0.1
Yes	87	Green	Yes	0.5	0.5
No	44	Blue	No	0.5	-0.5
Yes	19	Red	No	0.1	-0.1
No	32	Green	Yes	0.9	0.1
No	14	Blue	Yes	0.9	0.1

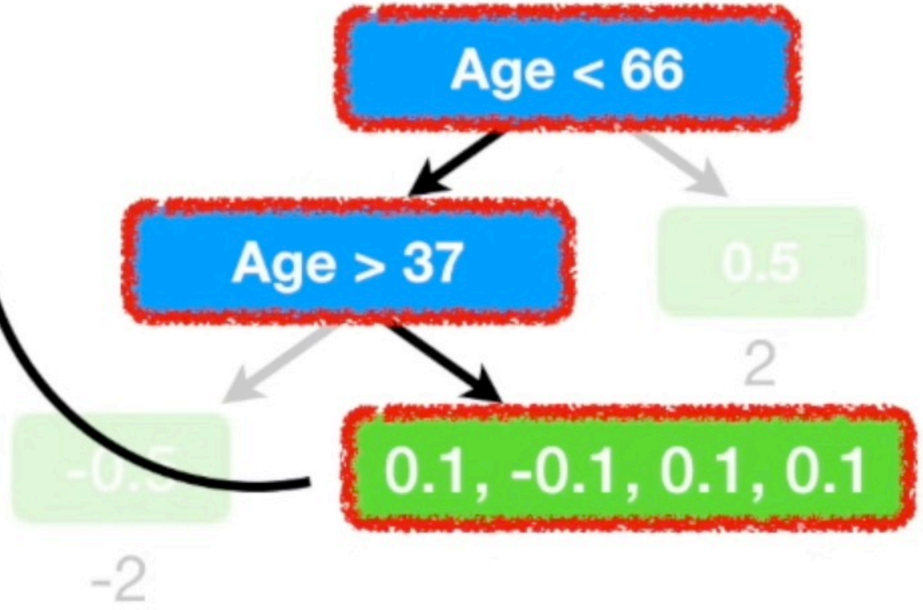
…and the **Output Value** for this leaf is **2**.

Lastly, let's calculate the **Output Value** for this leaf…



Likes Popcorn	Age	Favorite Color	Loves Troll 2	Predicted Prob.	Residual
Yes	12	Blue	Yes	0.9	0.1
Yes	87	Green	Yes	0.5	0.5
No	44	Blue	No	0.5	-0.5
Yes	19	Red	No	0.1	-0.1
No	32	Green	Yes	0.9	0.1
No	14	Blue	Yes	0.9	0.1

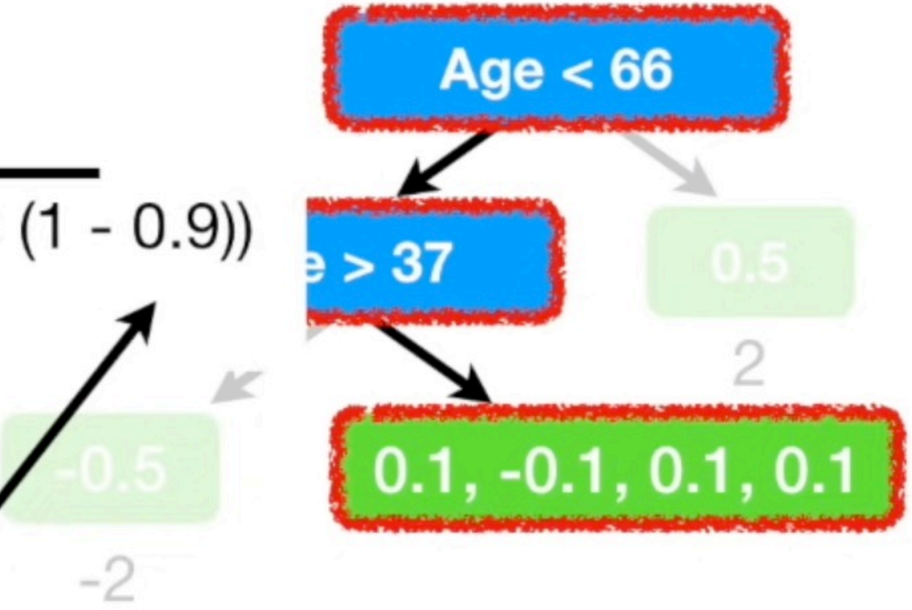
$$\frac{\sum \text{Residual}_i}{\sum \text{Previous Probability}_i \times (1 - \text{Previous Probability}_i)}$$



Likes Popcorn	Age	Favorite Color	Loves Troll 2	Predicted Prob.	Residual
Yes	12	Blue	Yes	0.9	0.1
Yes	87	Green	Yes	0.5	0.5
No	44	Blue	No	0.5	-0.5
Yes	19	Red	No	0.1	-0.1
No	32	Green	Yes	0.9	0.1
No	14	Blue	Yes	0.9	0.1

So we plug the **Residuals** into the formula for the **Output Values**…

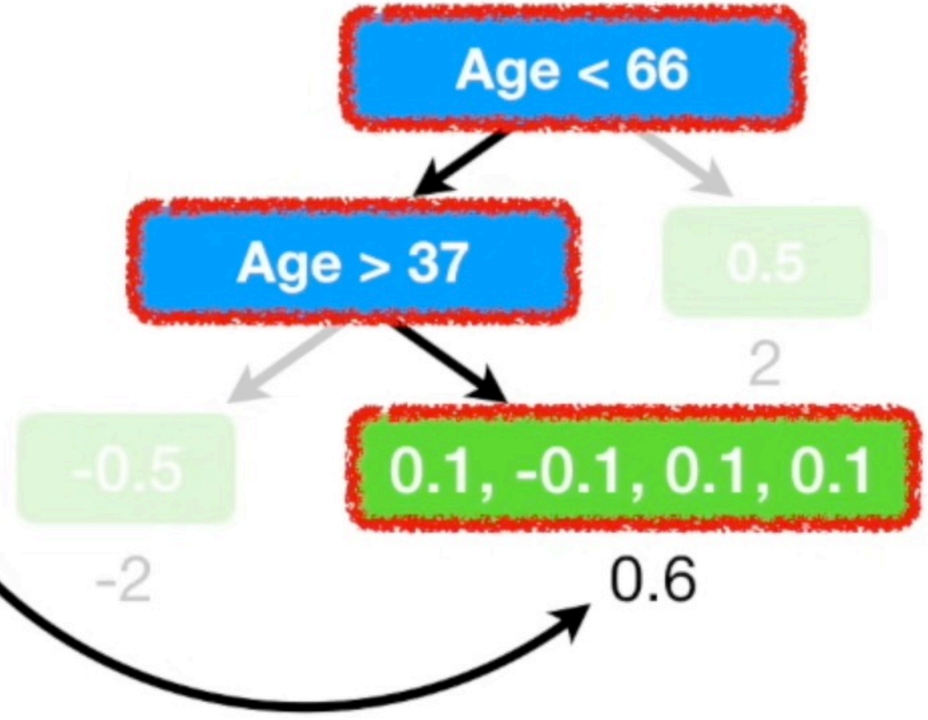
$$\frac{0.1 + -0.1 + 0.1 + 0.1}{(0.9 \times (1 - 0.9)) + (0.1 \times (1 - 0.1)) + (0.9 \times (1 - 0.9)) + (0.9 \times (1 - 0.9))}$$



Likes Popcorn	Age	Favorite Color	Loves Troll 2	Predicted Prob.	Residual
Yes	12	Blue	Yes	0.9	0.1
Yes	87	Green	Yes	0.5	0.5
No	44	Blue	No	0.5	-0.5
Yes	19	Red	No	0.1	0.1
No	32	Green	Yes	0.9	0.1
No	14	Blue	Yes	0.9	0.1

…and we plug in the **Predicted Probability** for each individual in the leaf…

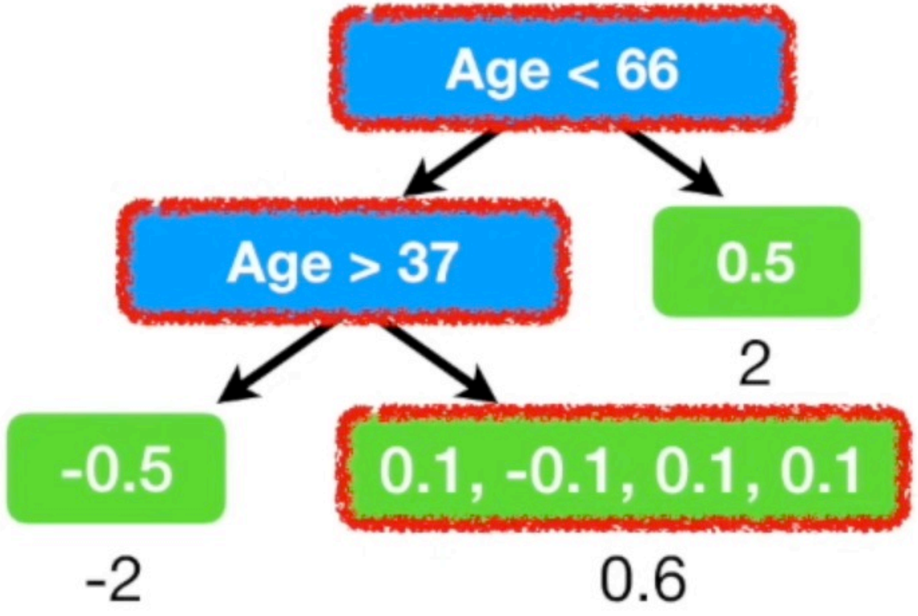
$$\frac{0.2}{0.09 + 0.09 + 0.09 + 0.09} = \boxed{0.6}$$



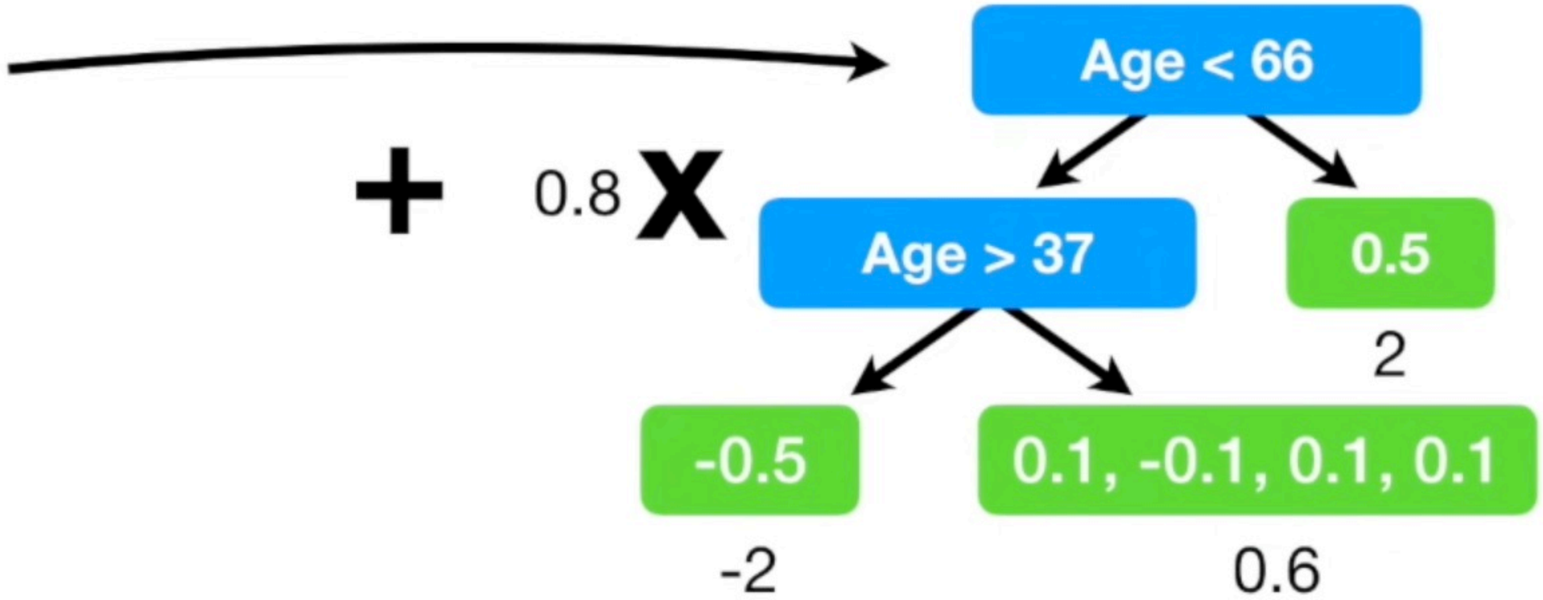
…and the **Output Value** for this leaf is **0.6**.

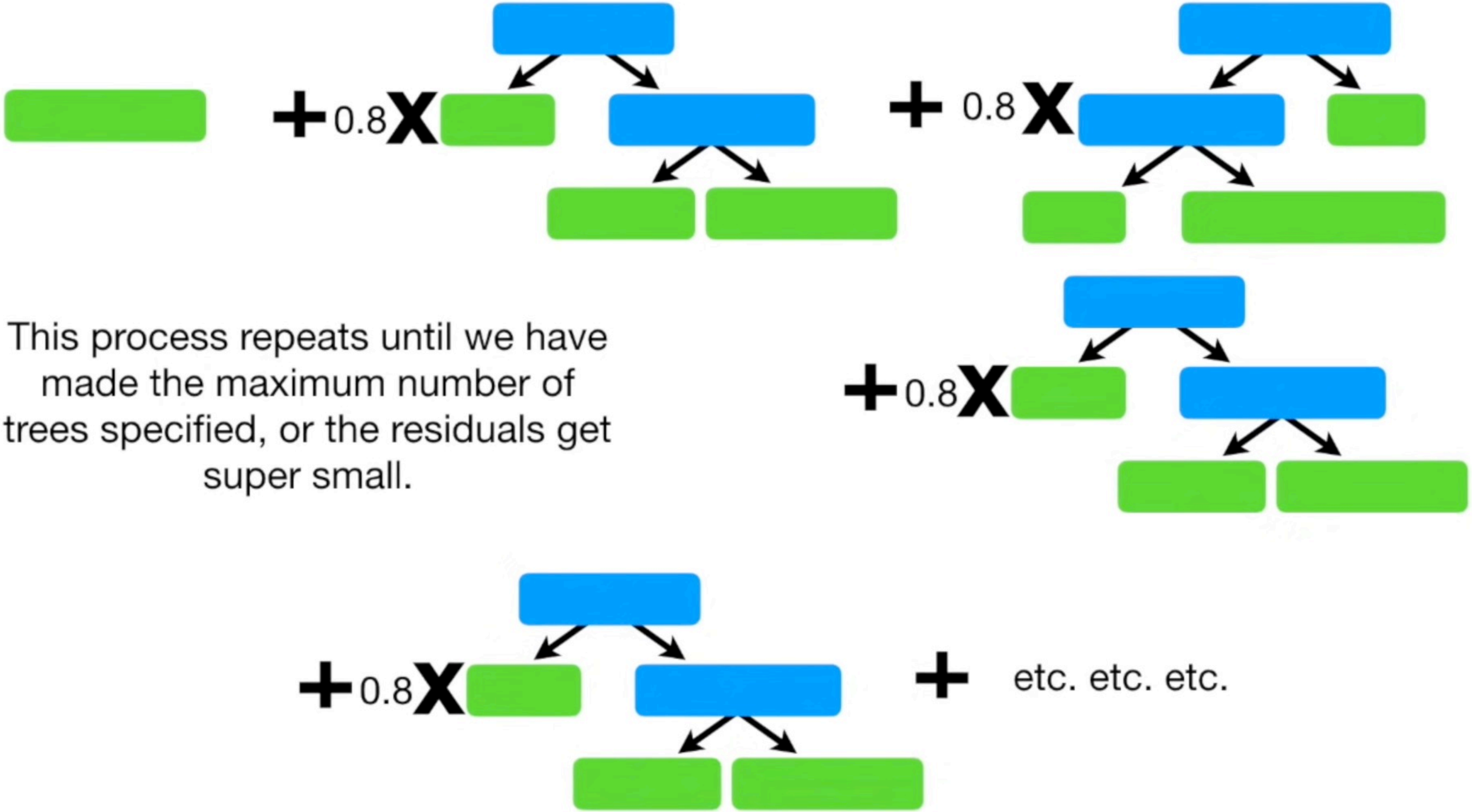
Likes Popcorn	Age	Favorite Color	Loves Troll 2	Predicted Prob.	Residual
Yes	12	Blue	Yes	0.9	0.1
Yes	87	Green	Yes	0.5	0.5
No	44	Blue	No	0.5	-0.5
Yes	19	Red	No	0.1	-0.1
No	32	Green	Yes	0.9	0.1
No	14	Blue	Yes	0.9	0.1

Now that we've calculated all of the **Output Values** for this tree, we can combine it with everything else we've done so far.



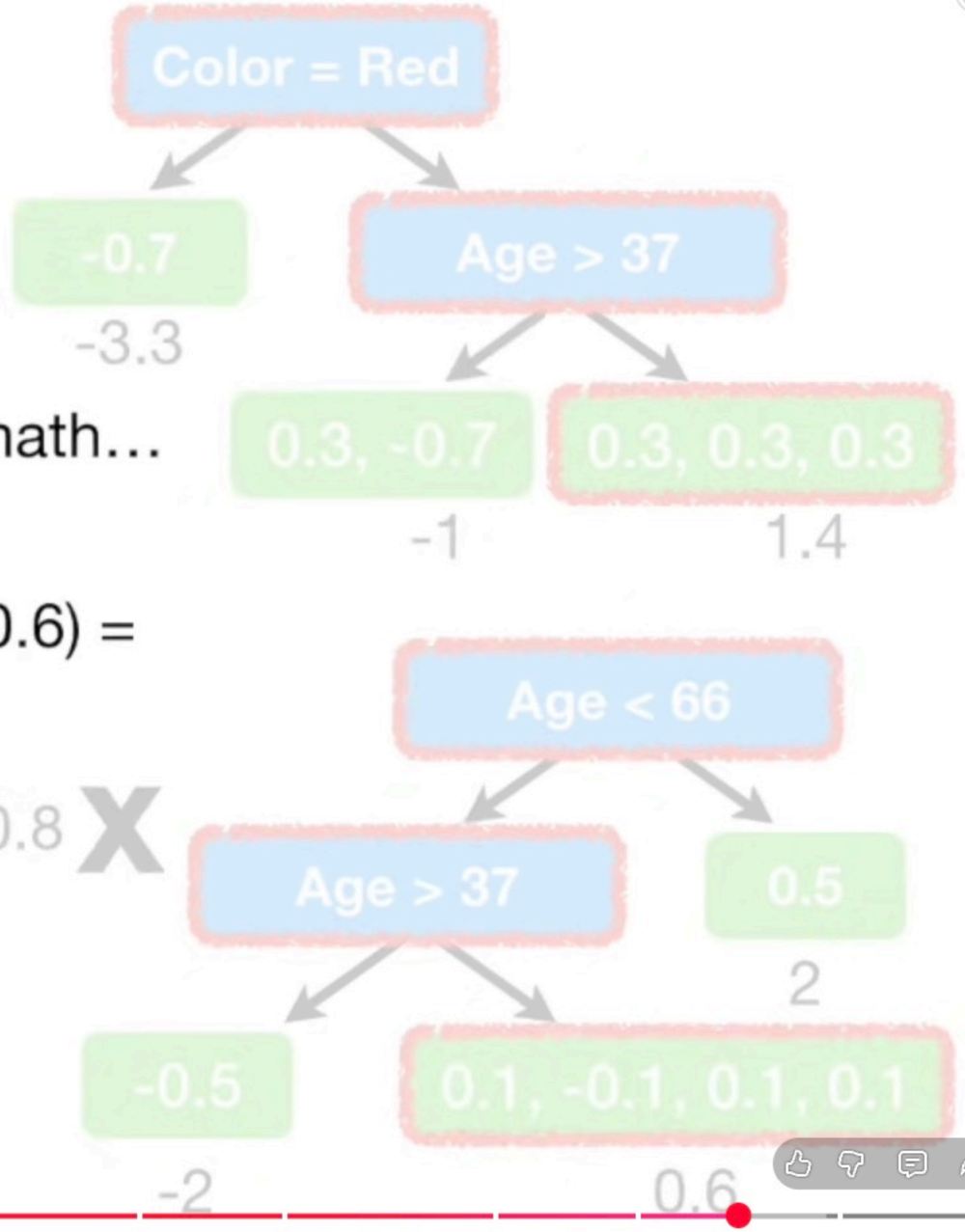
Then we built another tree based on the new **Residuals**, the difference between the **Observed** values and the values **Predicted** by the leaf ***and*** the first tree…





This process repeats until we have made the maximum number of trees specified, or the residuals get super small.

log(4/2) = 0.7 + 0.8 X



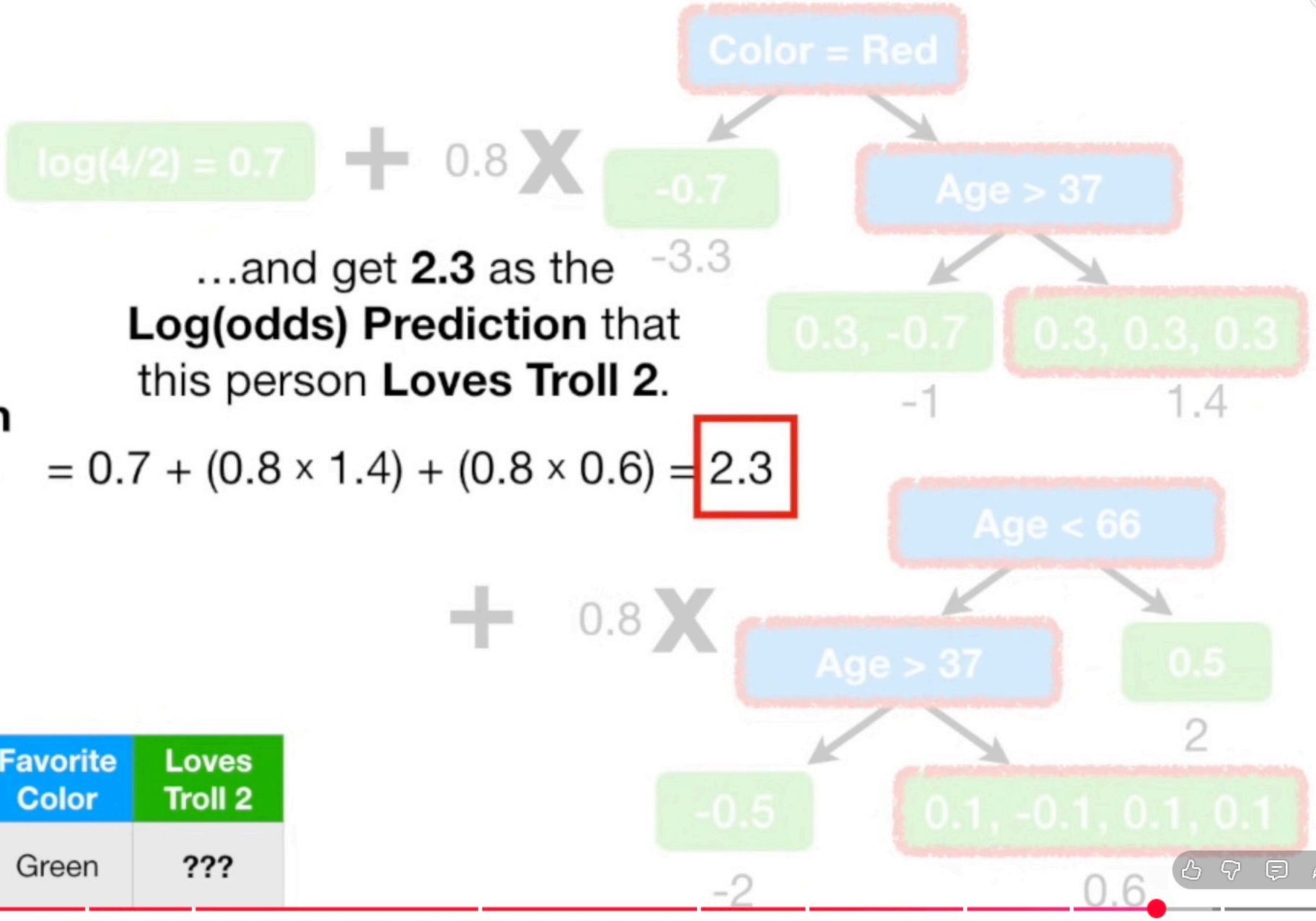
Now we just do the math…

Log(odds) Prediction that someone **Loves Troll 2**:

$= 0.7 + (0.8 \times 1.4) + (0.8 \times 0.6) =$

+ 0.8 X

Likes Popcorn	Age	Favorite Color	Loves Troll 2
Yes	25	Green	**???**



…and get **2.3** as the **Log(odds) Prediction** that this person **Loves Troll 2**.

Log(odds) Prediction that someone **Loves Troll 2**:

$$= 0.7 + (0.8 \times 1.4) + (0.8 \times 0.6) = \boxed{2.3}$$

Likes Popcorn	Age	Favorite Color	Loves Troll 2
Yes	25	Green	???

…and the **Predicted Probability** that this individual will **Love Troll 2** is **0.9**.

Log(odds) Prediction that someone **Loves Troll 2**:

$$= 0.7 + (0.8 \times 1.4) + (0.8 \times 0.6) = 2.3$$

$$\text{Probability} = \frac{e^{2.3}}{1 + e^{2.3}} = \boxed{0.9}$$

Likes Popcorn	Age	Favorite Color	Loves Troll 2
Yes	25	Green	**???**