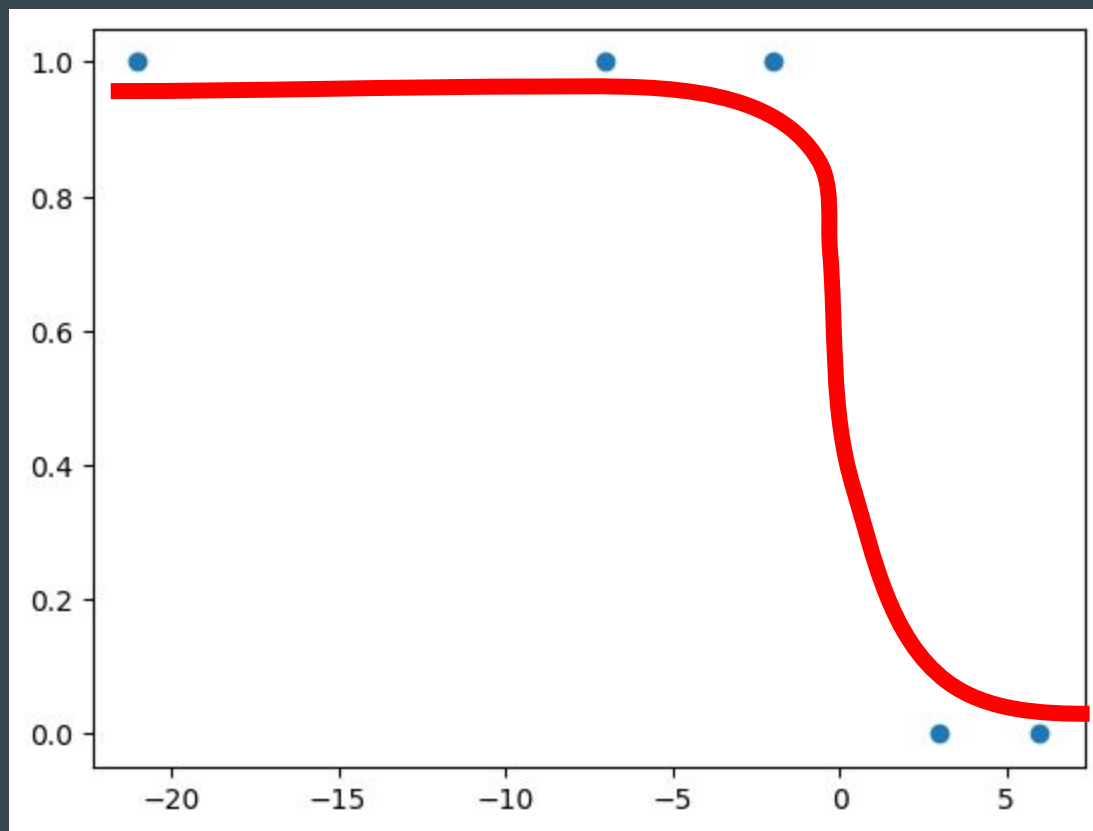


XGBoost Presentation

...

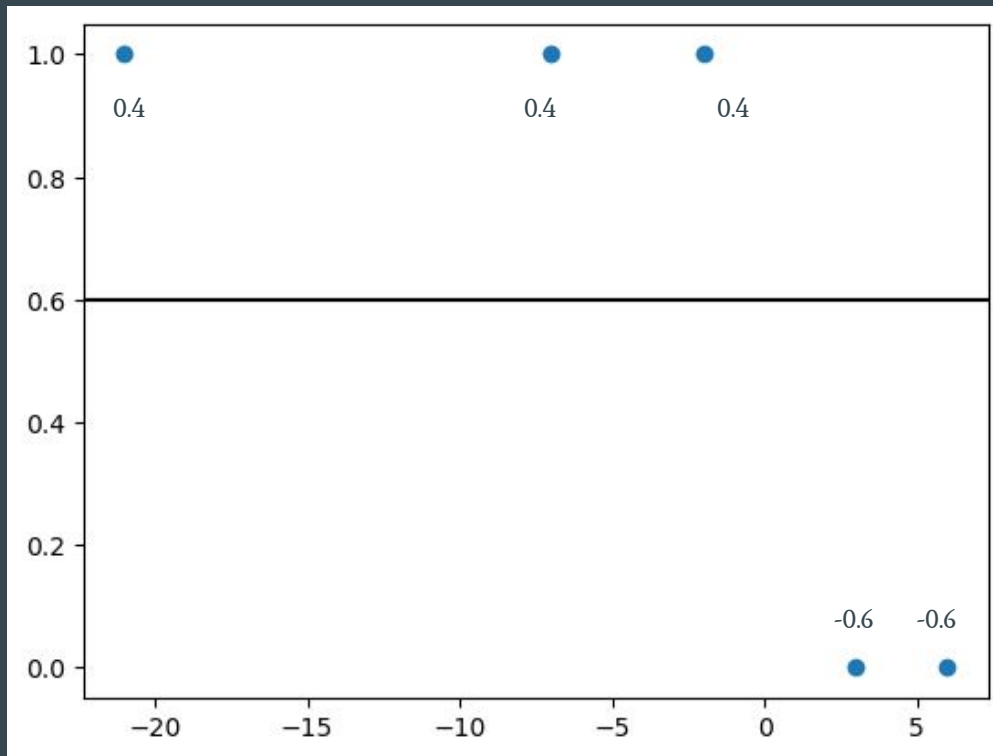
By: Your Boy

Matchup	Spread	Home Winner
LAC vs. CIN	-7	1
SF vs. SEA	3	0
PIT vs. BAL	-2	1
NE vs. LAR	6	0
NO vs. CLE	-21	1



Initial Prediction

$\frac{3}{5} = 60\%$



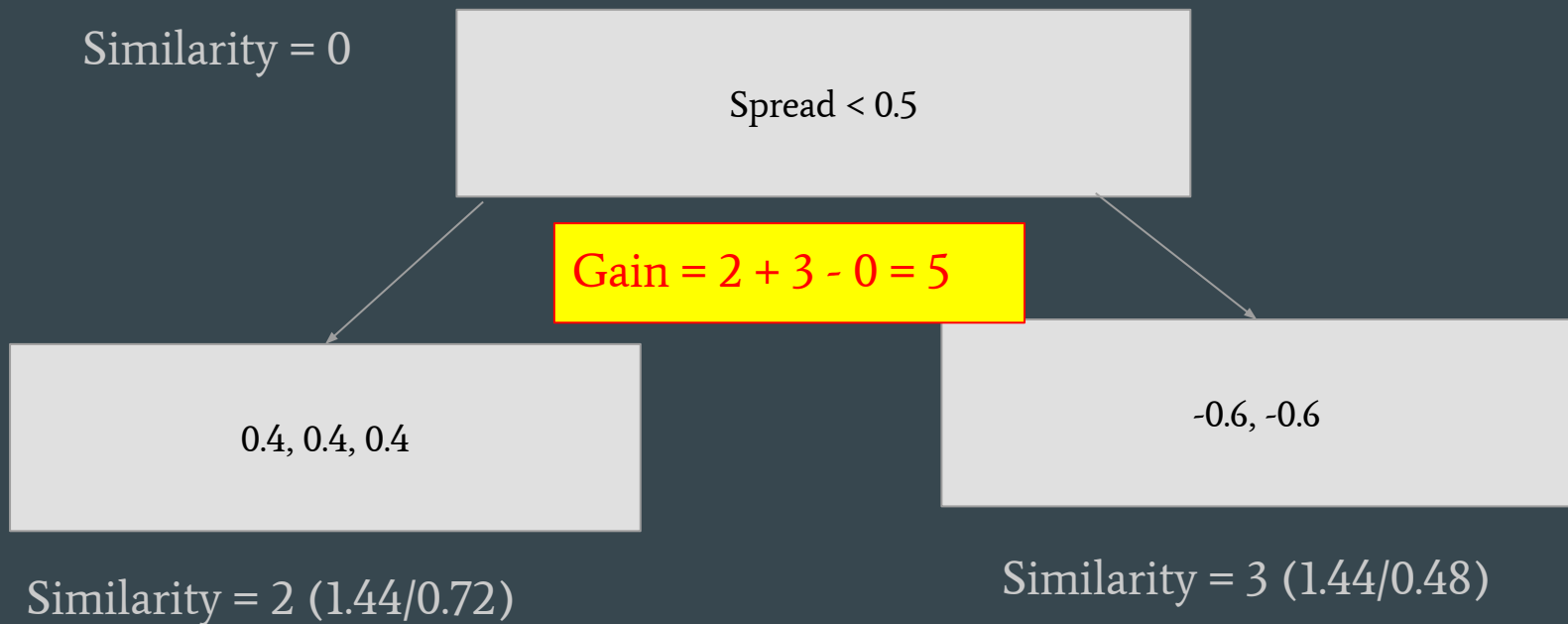
Building the Trees

0.4, 0.4, 0.4, -0.6, -0.6

Similarity score: $(\text{Sum of the residuals})^2 / (\text{Sum of the prev prob} * 1 - \text{of prev prob}) + \text{lambda}$

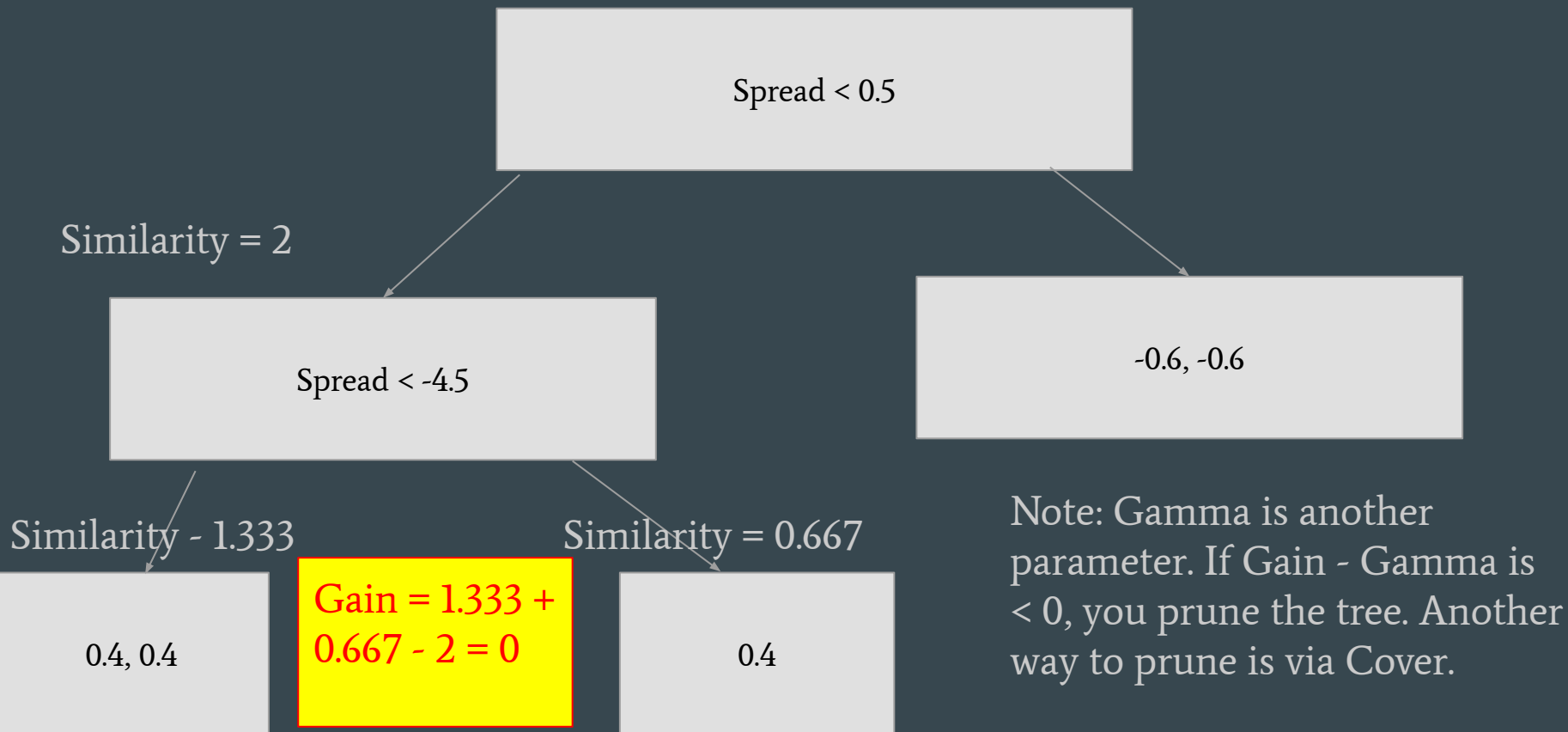
$$\text{Similarity Score} = 0 / (0.4 * 0.6) + (0.4 * 0.6) + (0.4 * 0.6) + (0.4 * 0.6) + (0.4 * 0.6) + 0 = 0$$

Building Trees Cont.

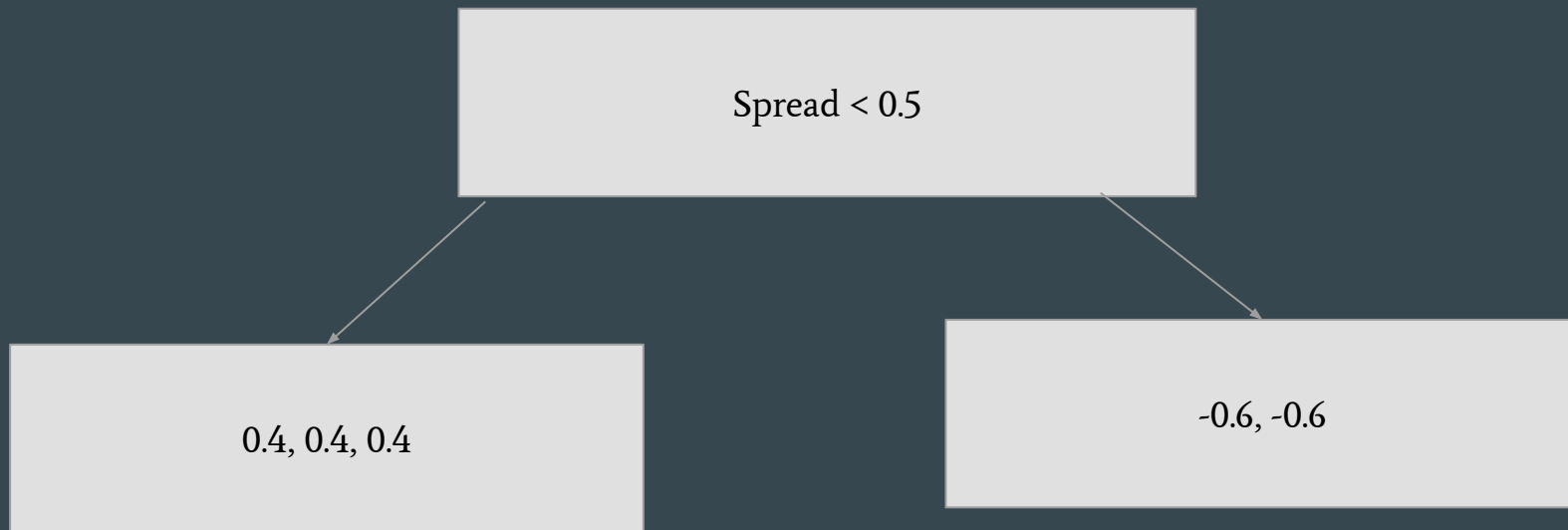


Note: If I set lambda to 1, the similarity scores decrease to 0.837 and 0.973 respectively.

Building Trees Cont.



Final Tree



Making Predictions

Log(odds): $\log(p/1-p) + \text{learning rate} * \text{observed residual}$

Where $p = 0.6$ and learning rate = 0.3 (default).

Ex) Spread = -21: $0.405 + 0.3 * -21 = -5.895$

To convert log of odds to logistic function: $e^{\log(\text{odds})} / 1 + e^{\log(\text{odds})}$

Ex) Spread = -21: $e^{5.895} / 1 + e^{5.895} = 0.997$

Matchup	Spread	Home Winner	Initial Prediction	Predictions
LAC vs. CIN	-7	1	0.6	0.845 = 84.5%
SF vs. SEA	3	0	0.6	0.213 = 21.3%
PIT vs. BAL	-2	1	0.6	0.587 = 58.7%
NE vs. LAR	6	0	0.6	0.100 = 10%
NO vs. CLE	-21	1	0.6	0.997 = 99.7%

XGBoost Parameters

Objective: “Binary:logistic”

N_estimators: Maximum # Of trees to make

Math Depth

Learning Rate: alias eta

Gamma

Min Child Weight: alias cover

Reg Lambda

Reg Alpha: mainly for regression

XGBoost Benefits

- Fast
- Handles NA Values
- Really good with tabular data
- Very accurate

Note: Only use with complex datasets. Can overfit to simple ones.