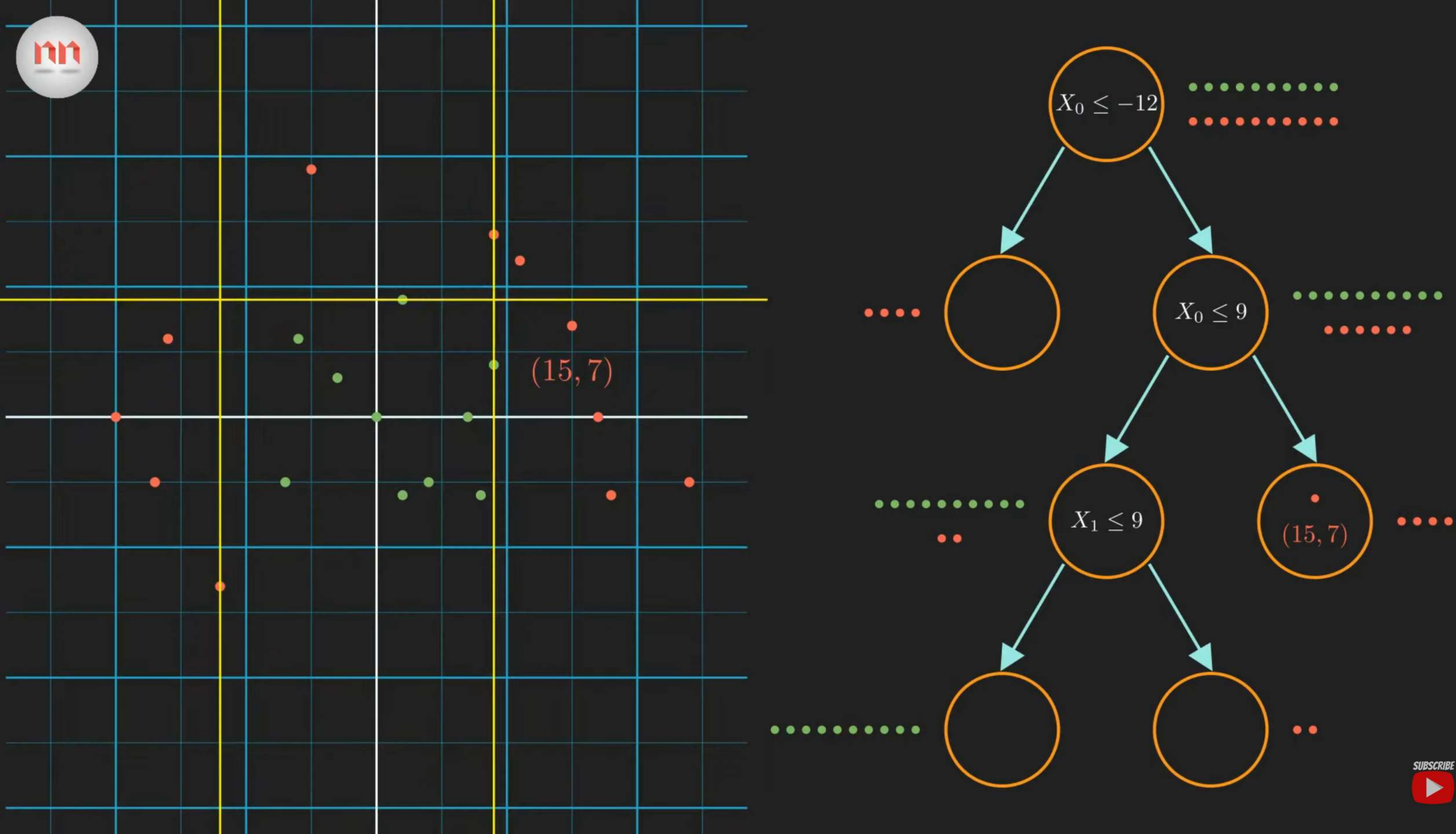
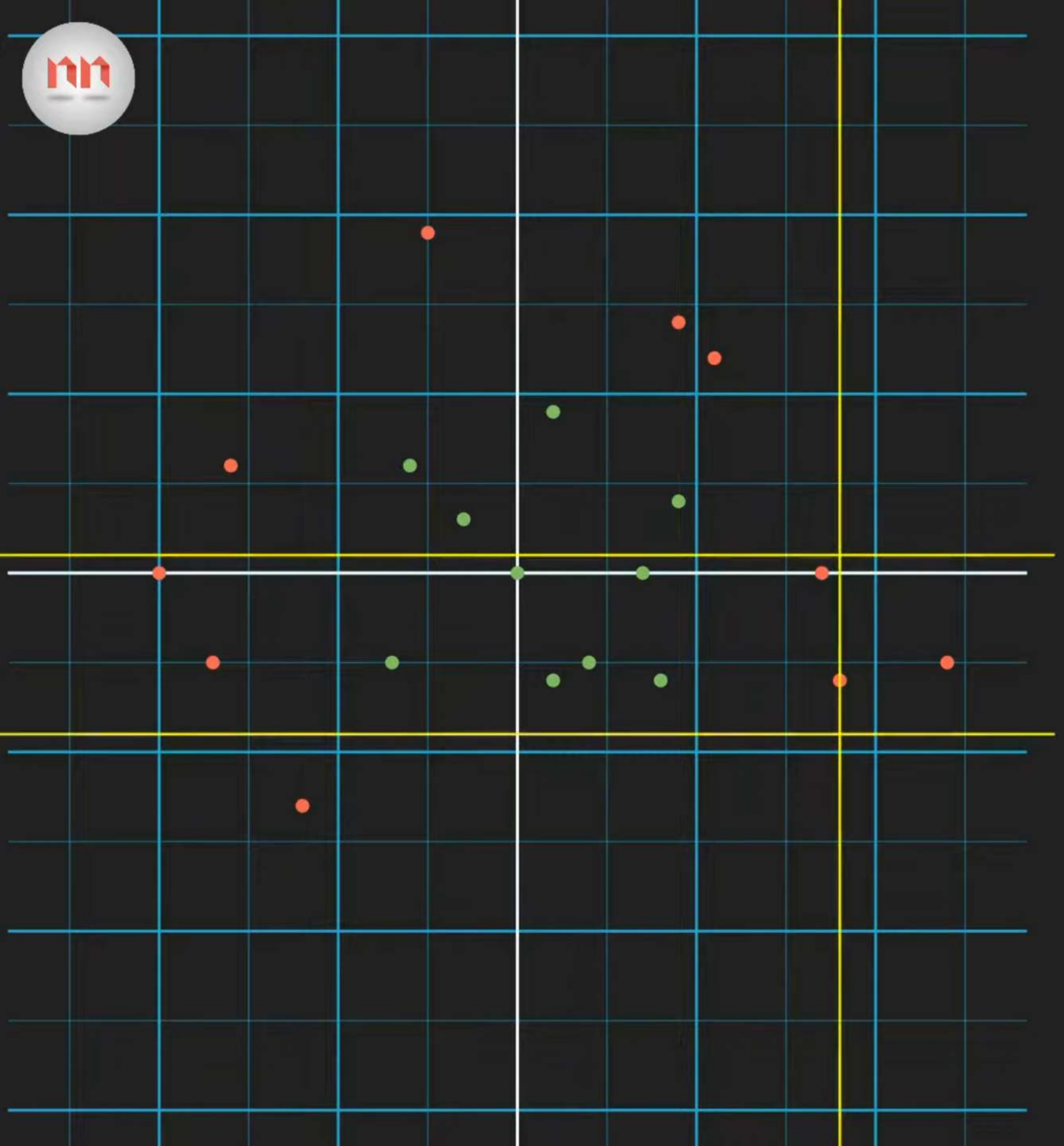
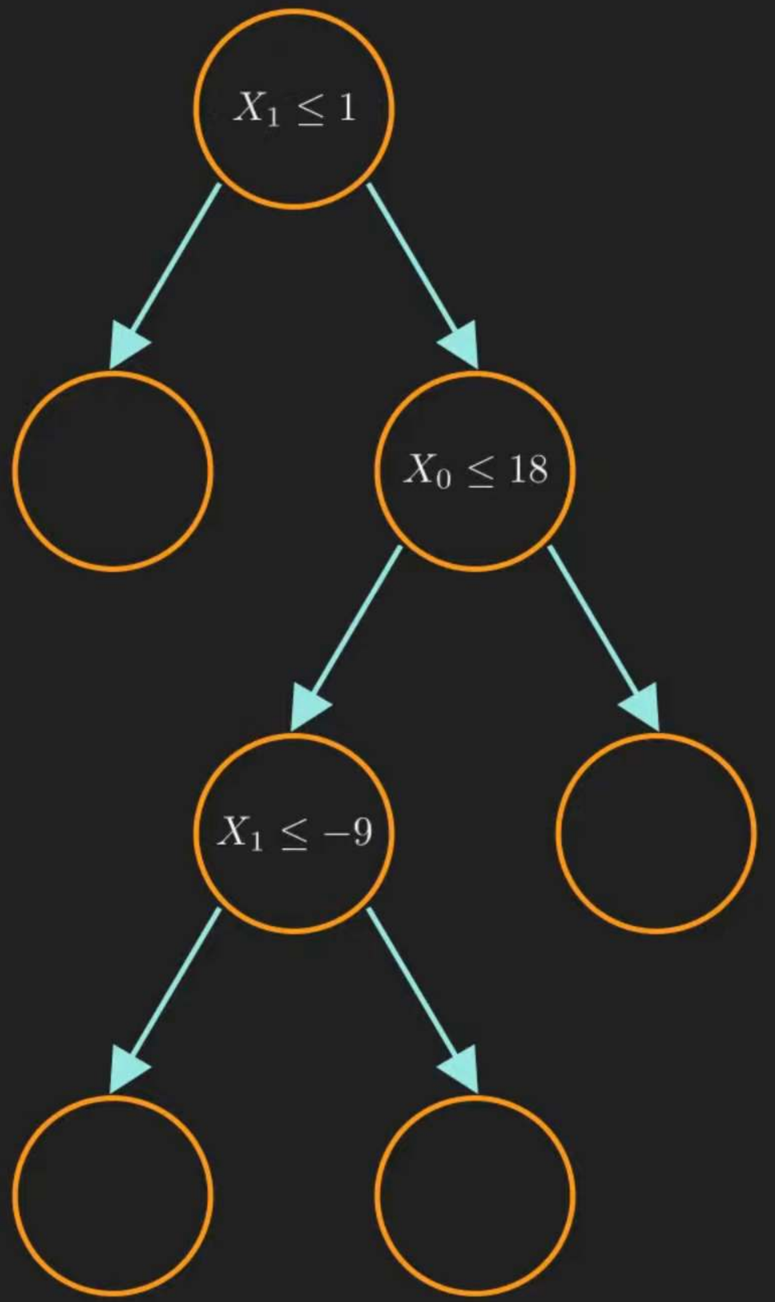
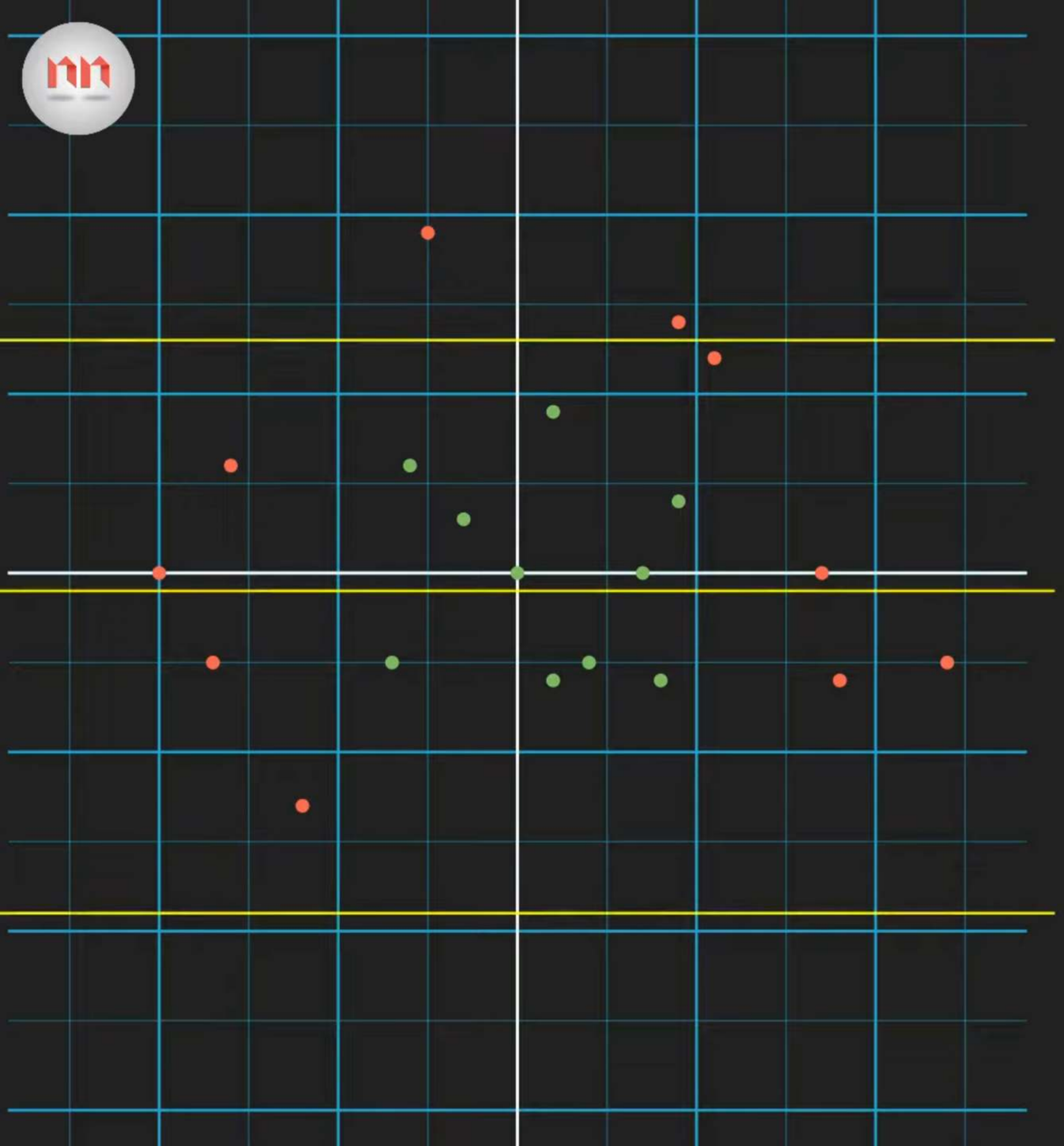
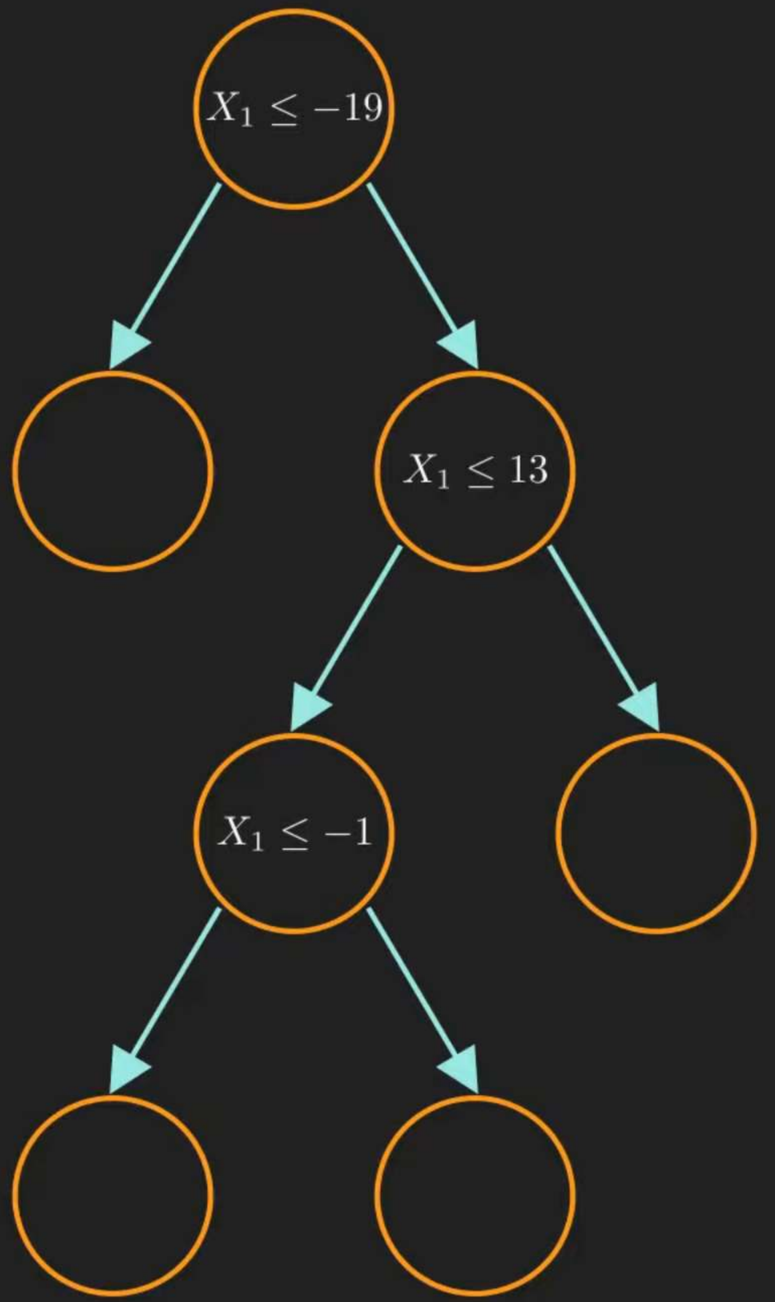
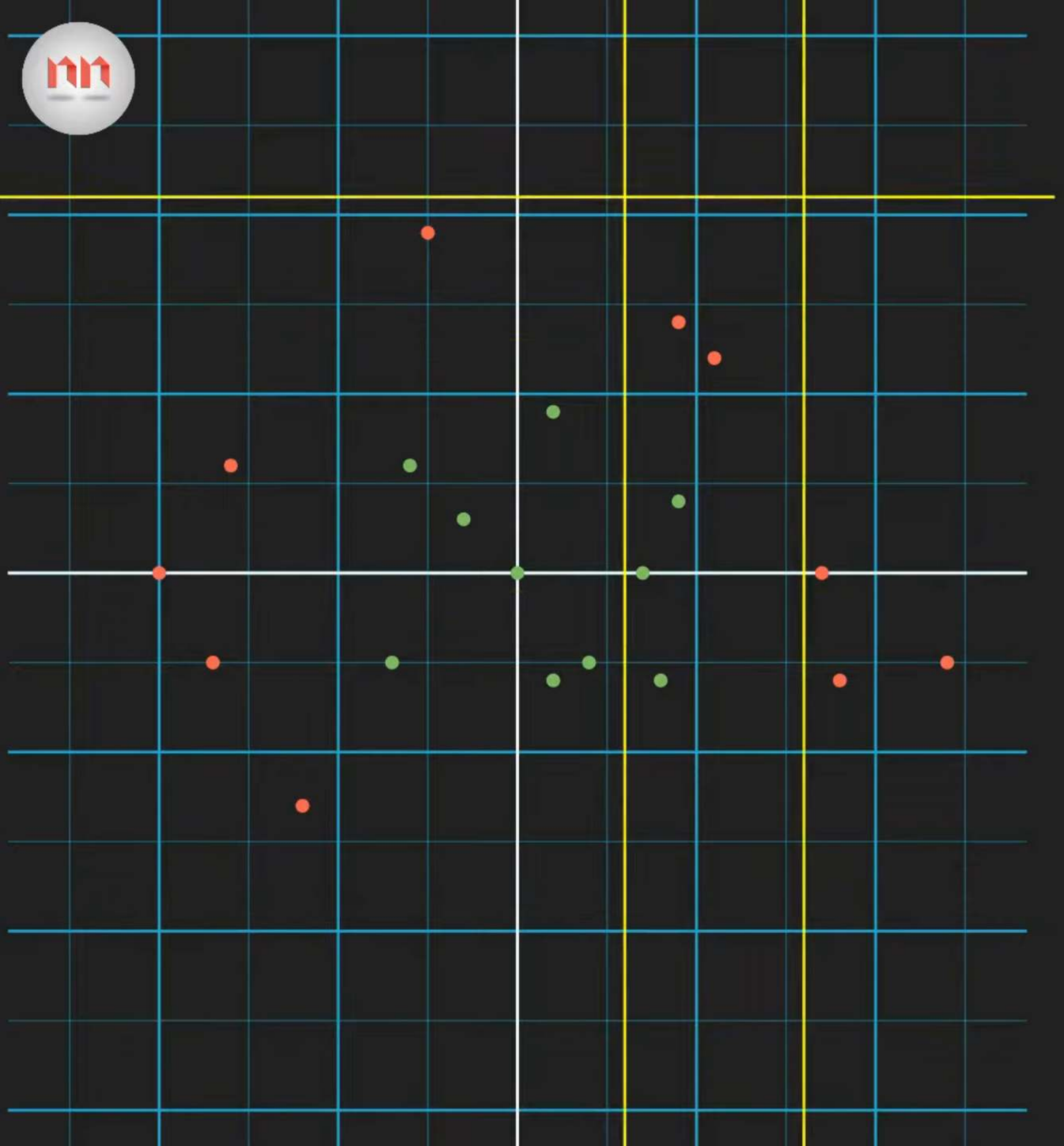

(15, 7)
$X_0 \leq -12$
$X_0 \leq 9$
$X_1 \leq 9$
(15, 7)
SUBSCRIBE

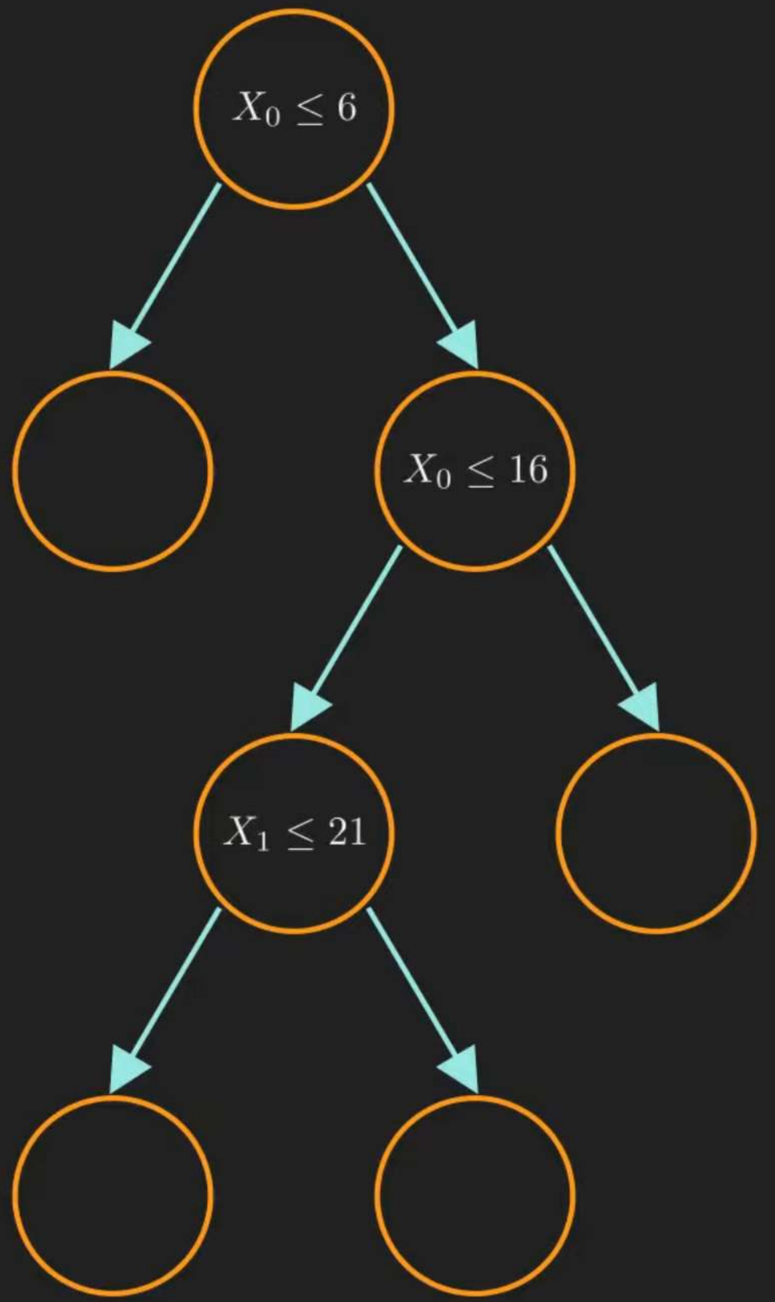


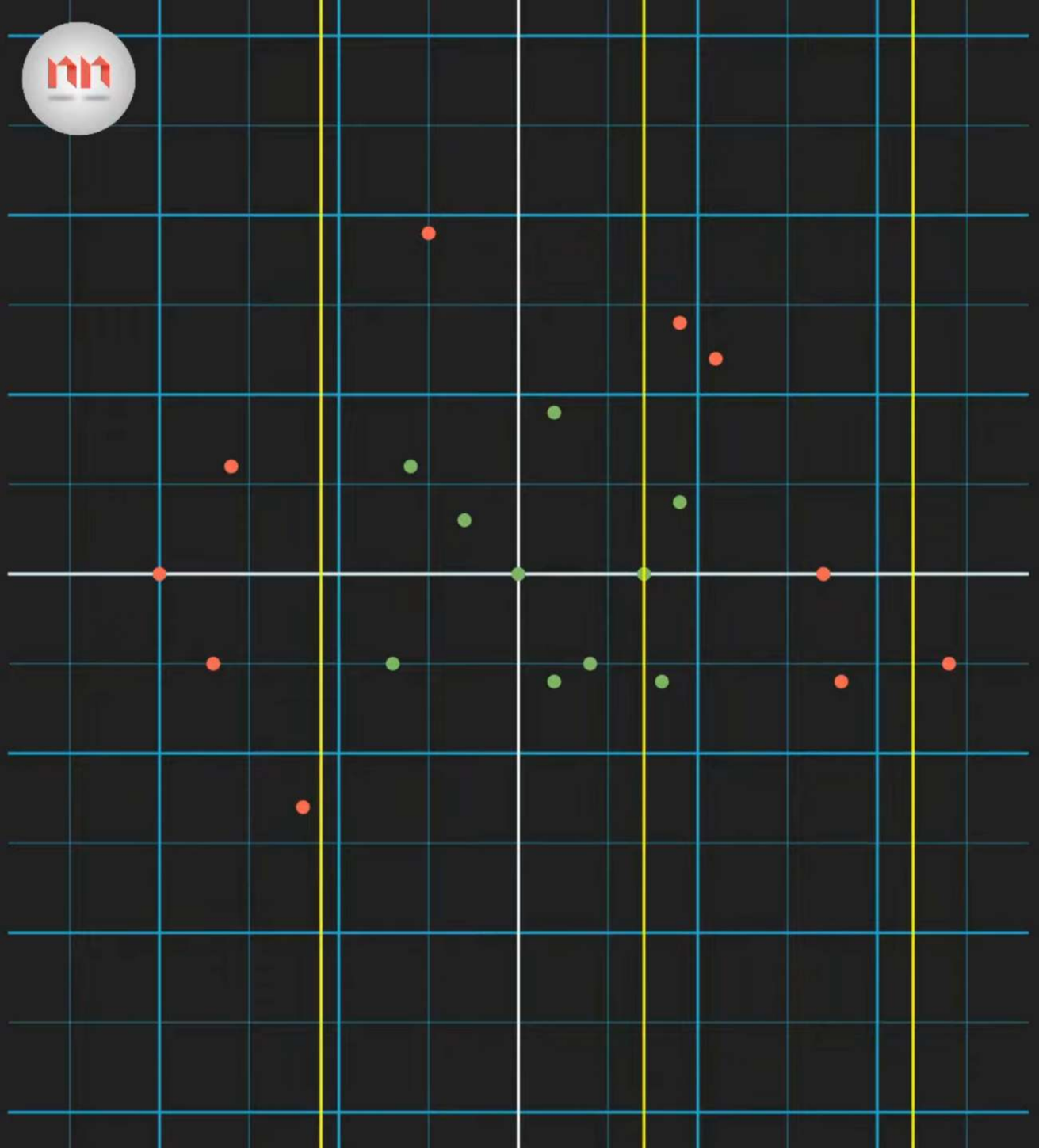


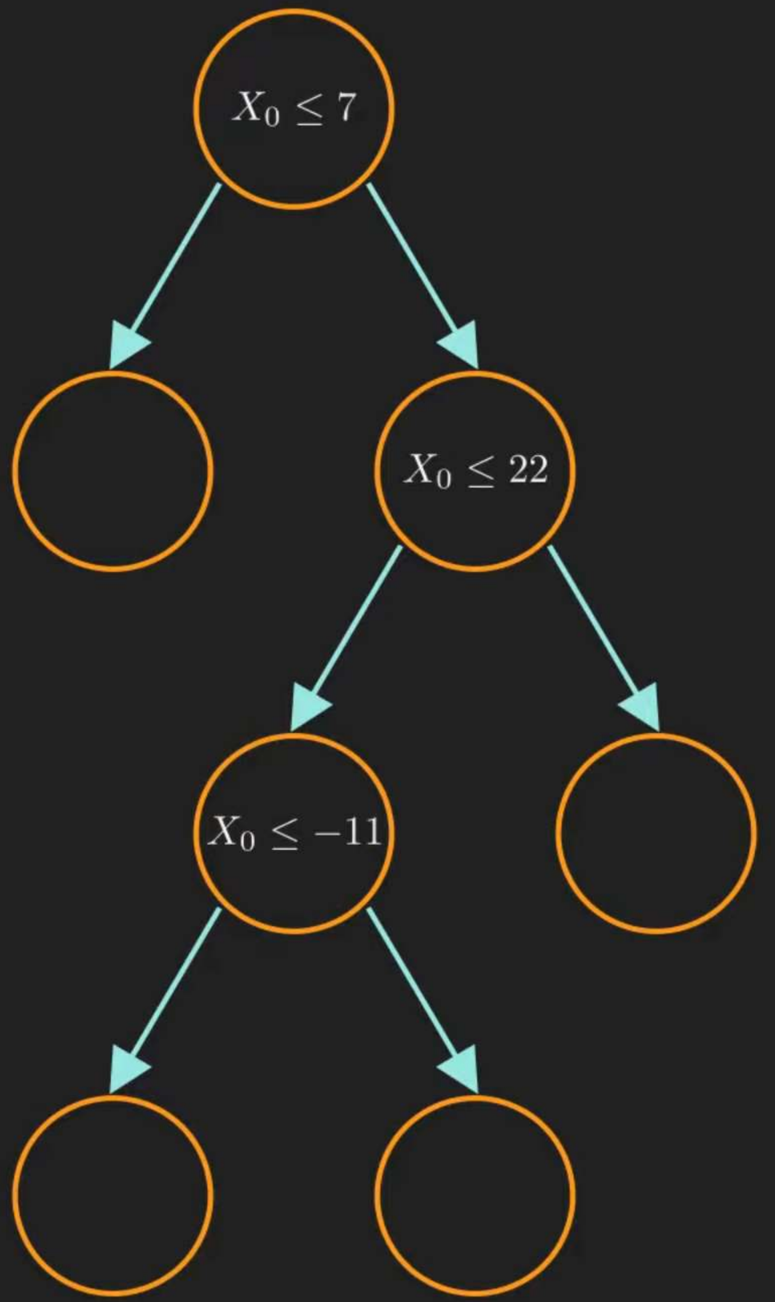


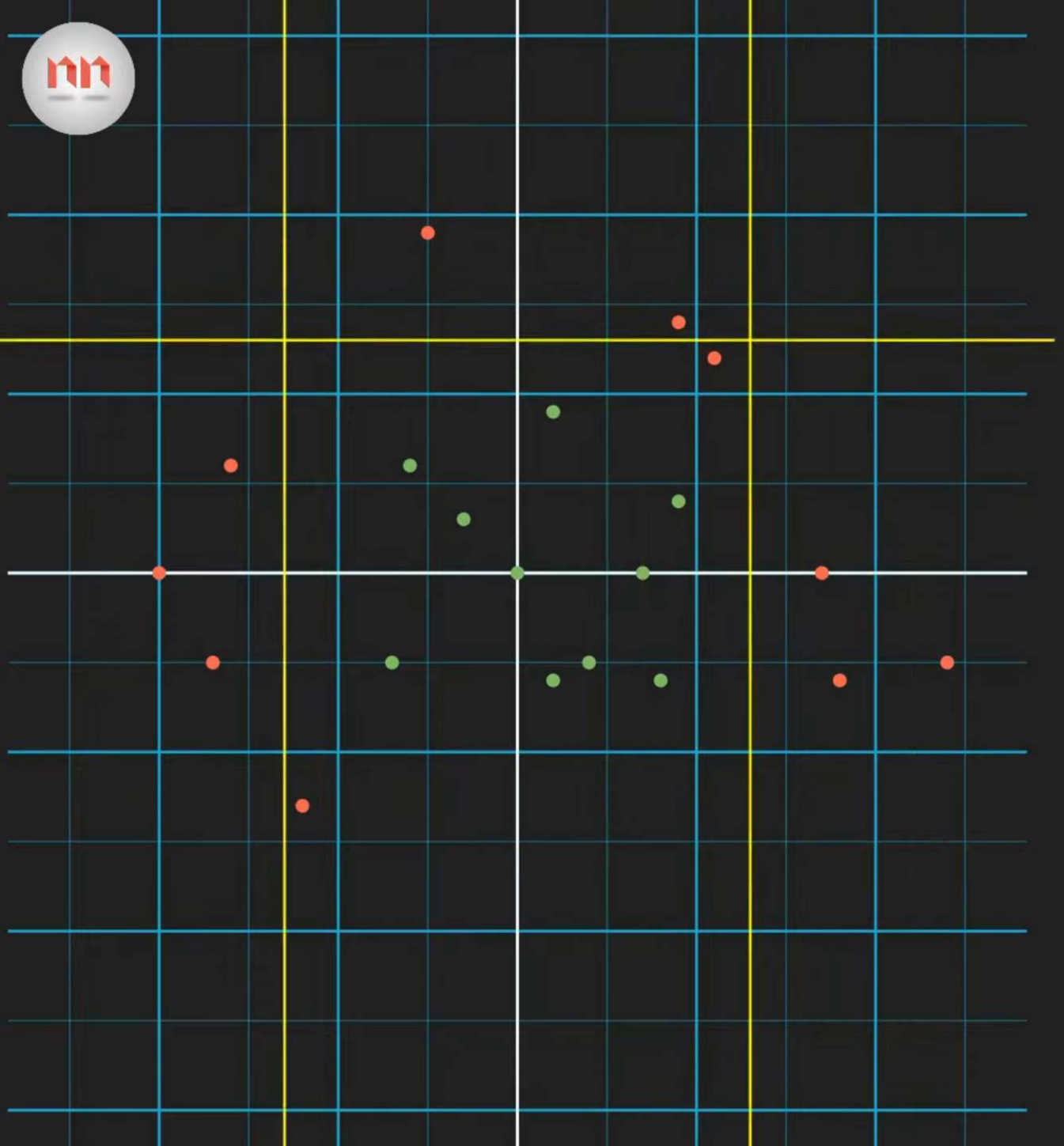


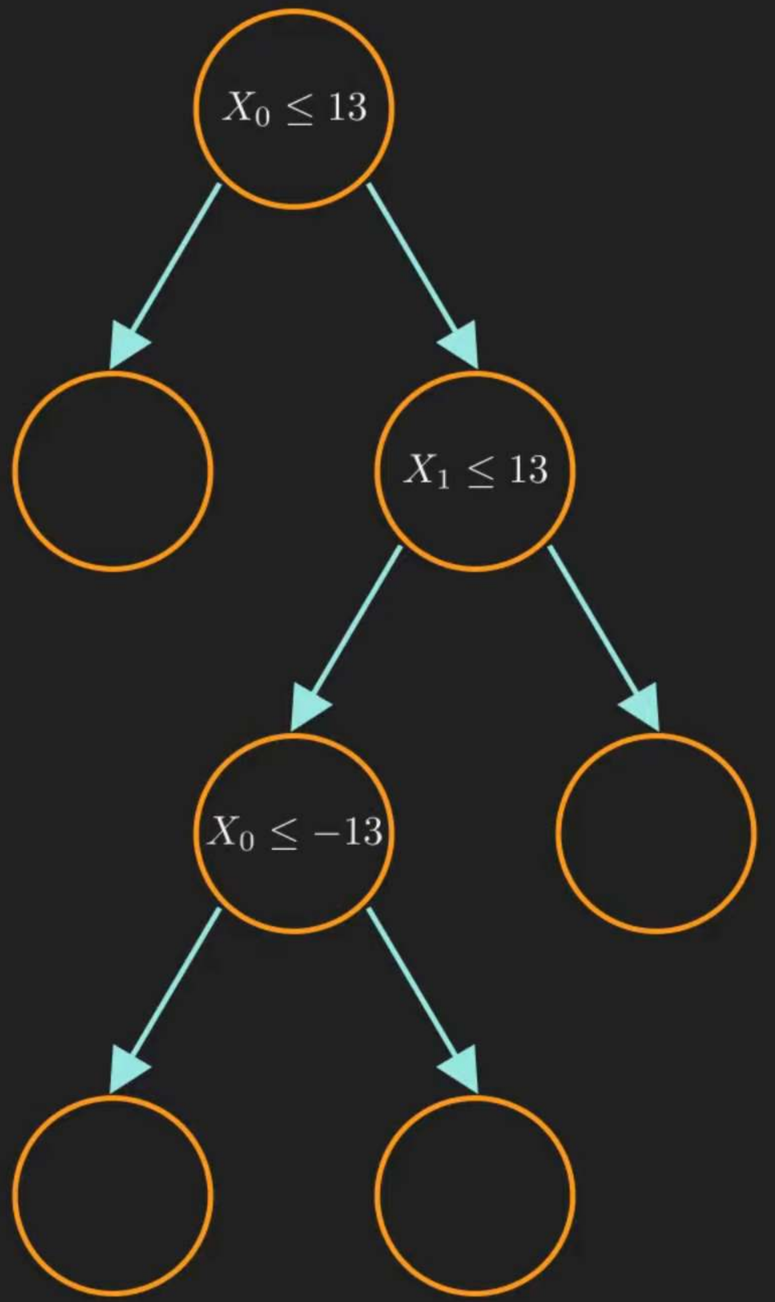


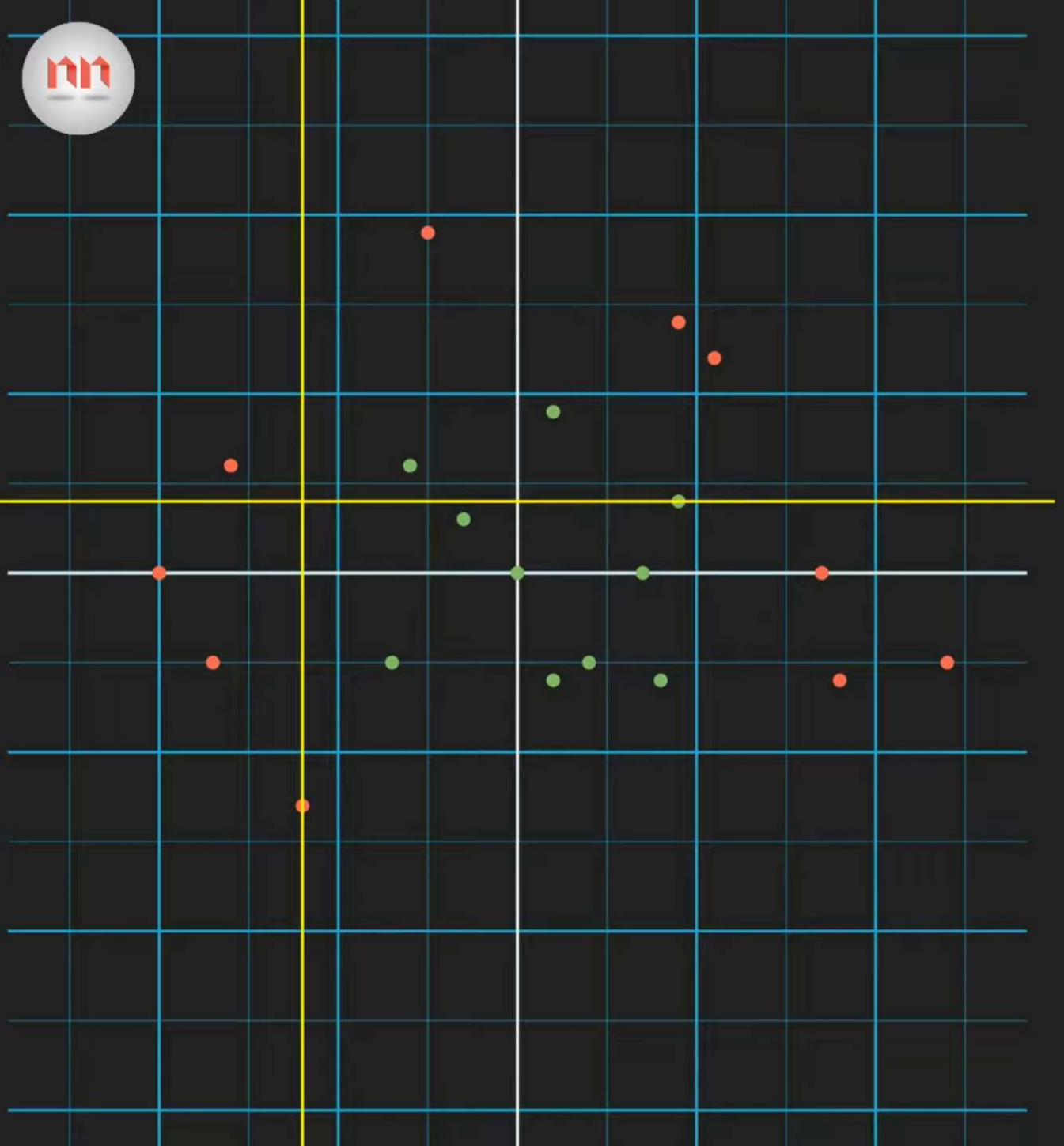


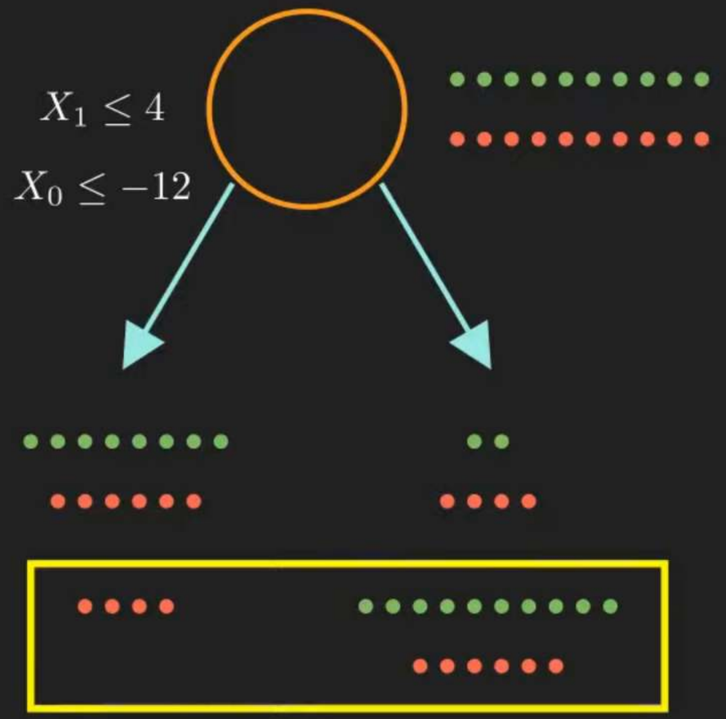




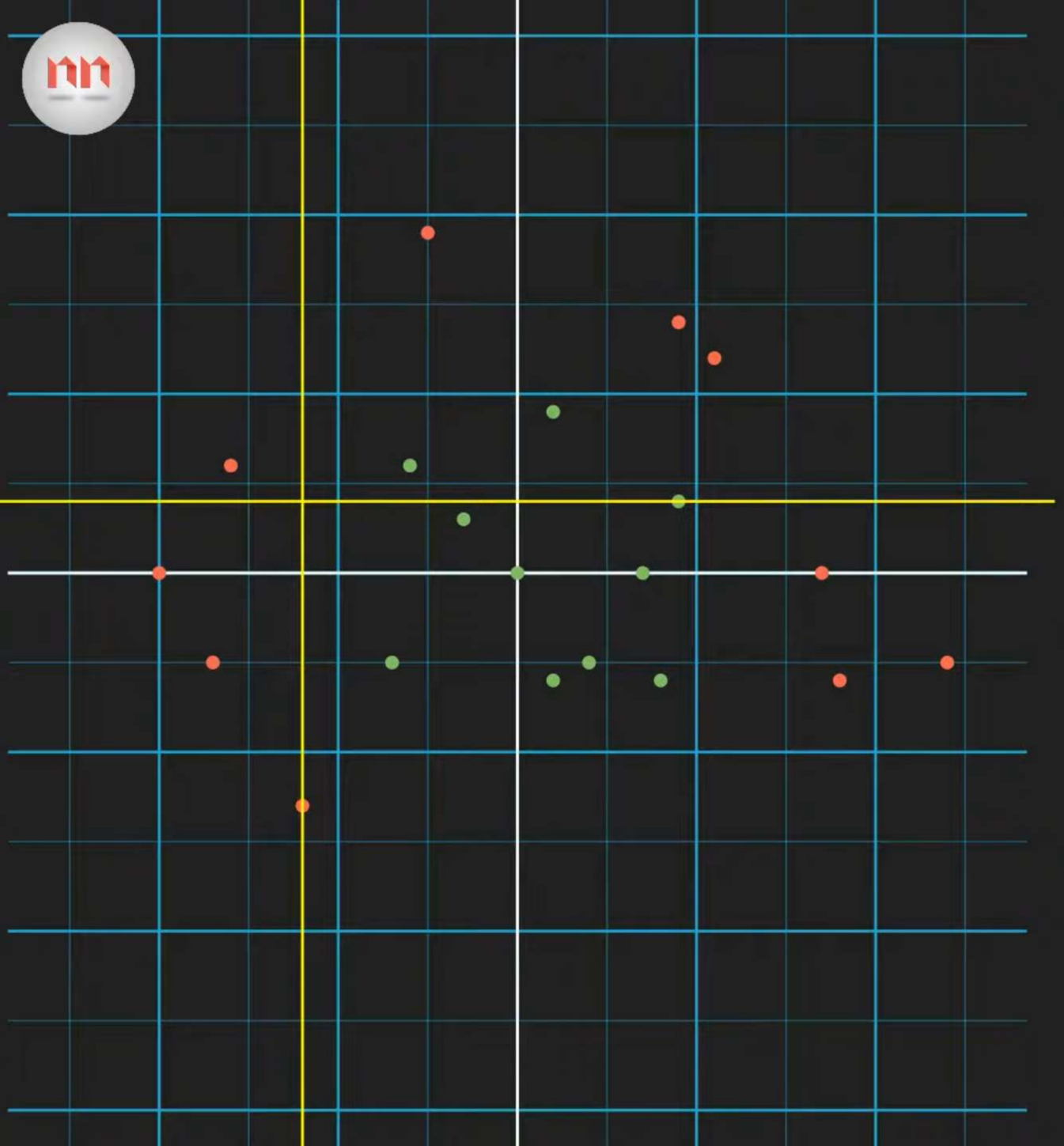


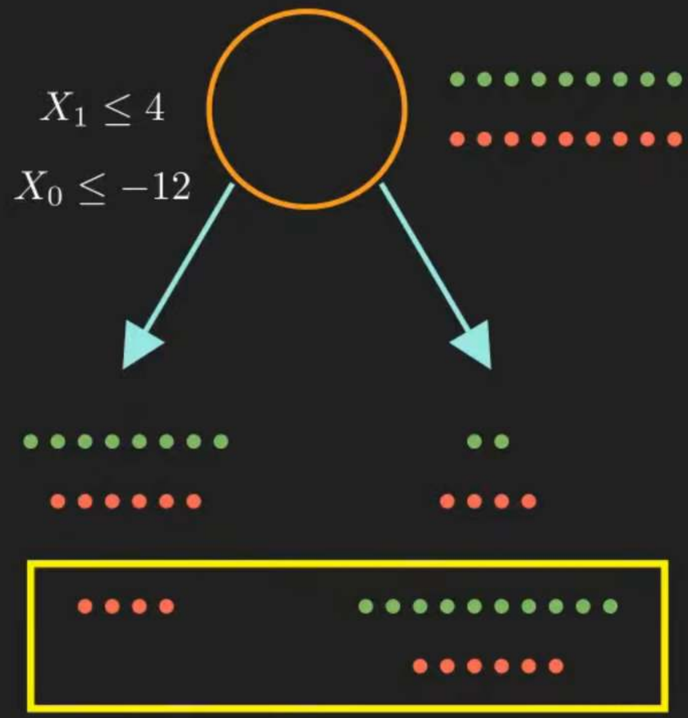






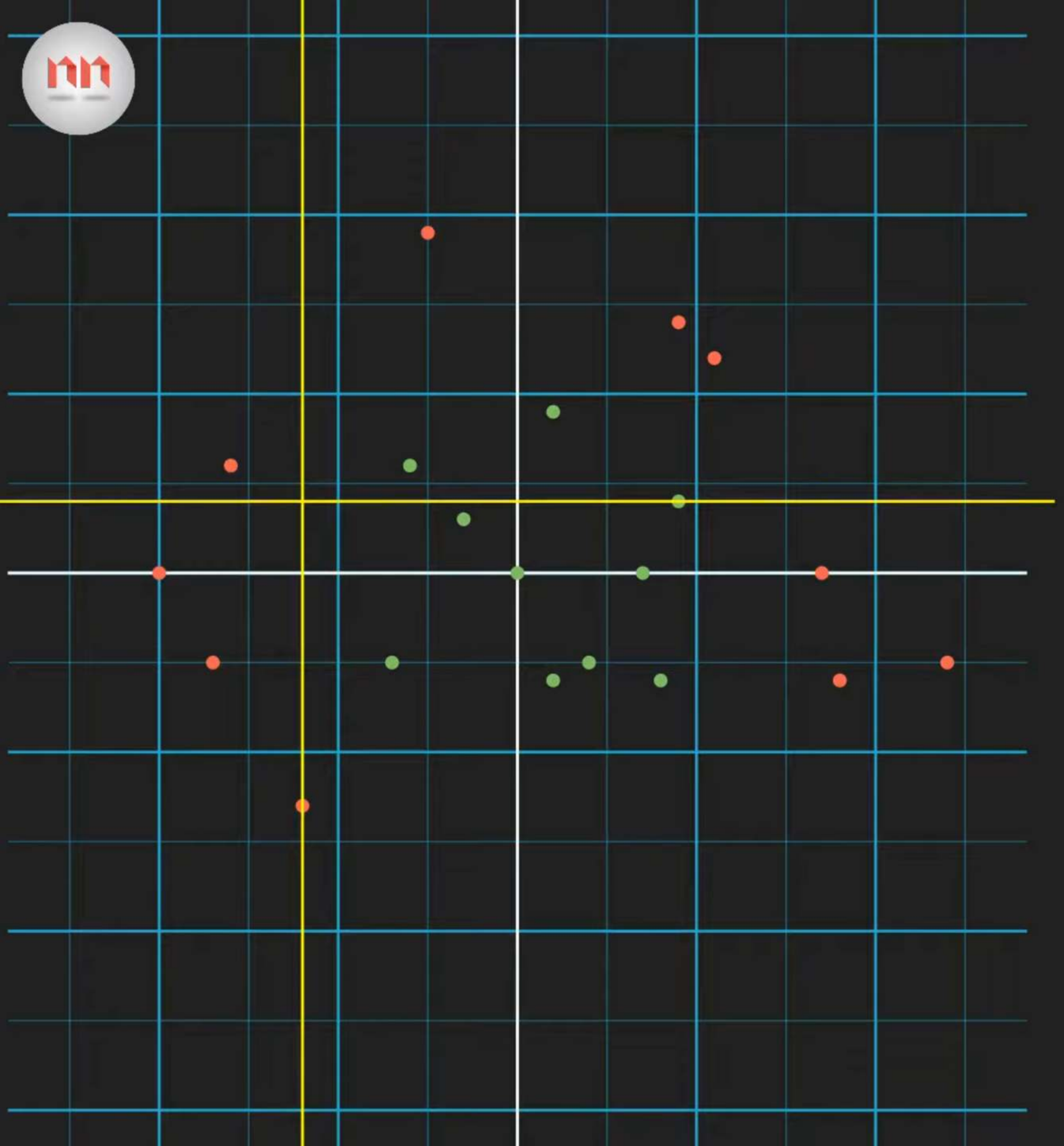
Which split is better?

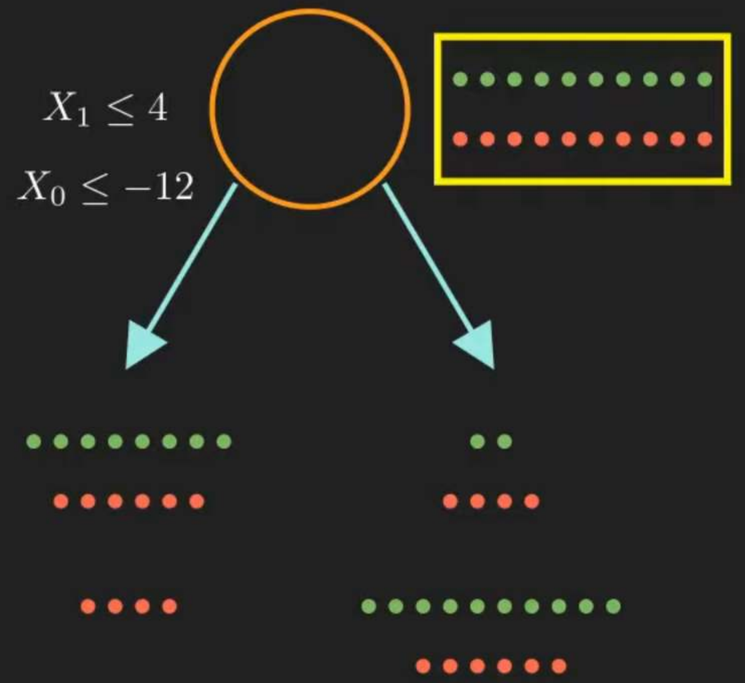




Information Gain

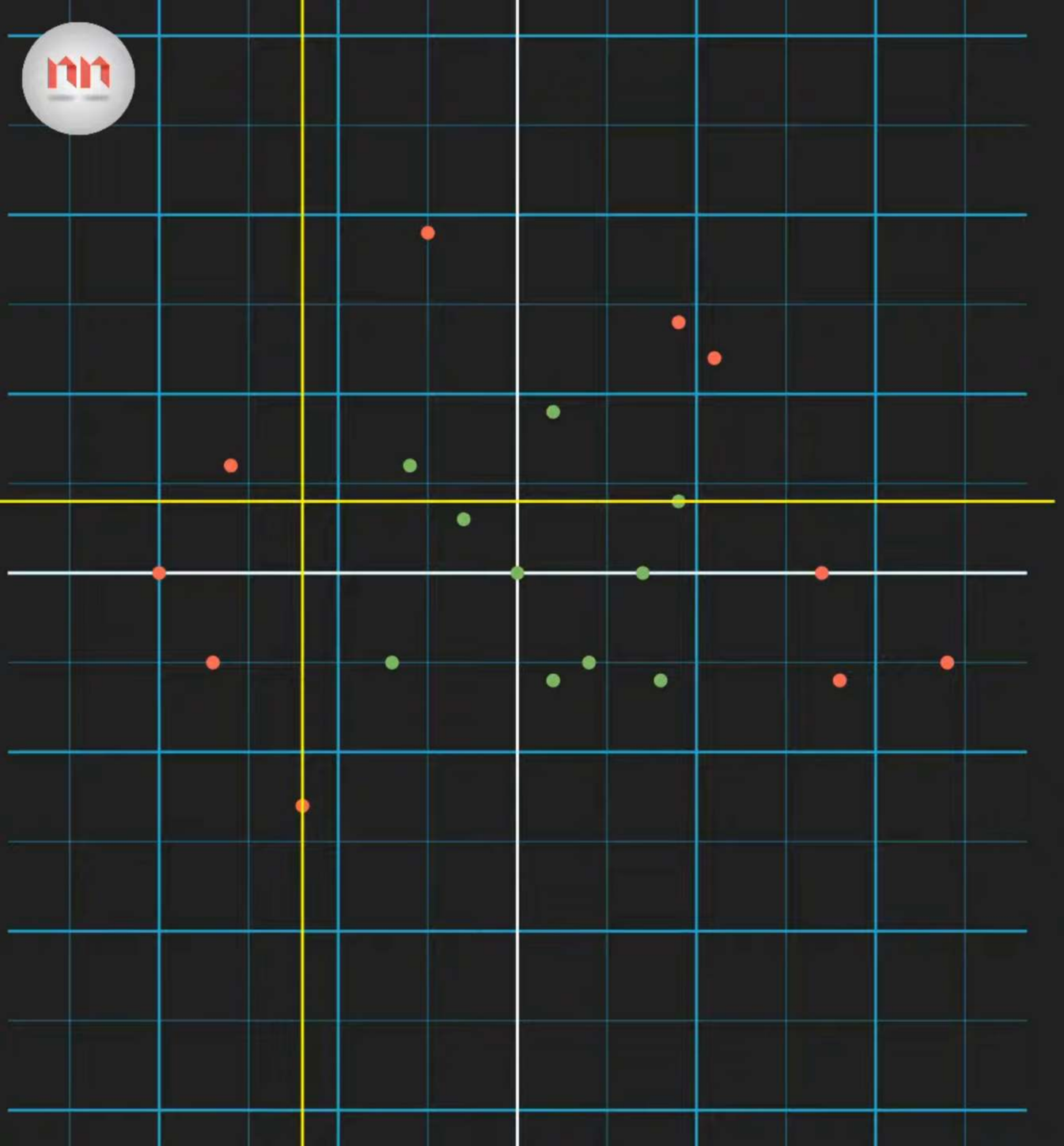
SUBSCRIBE

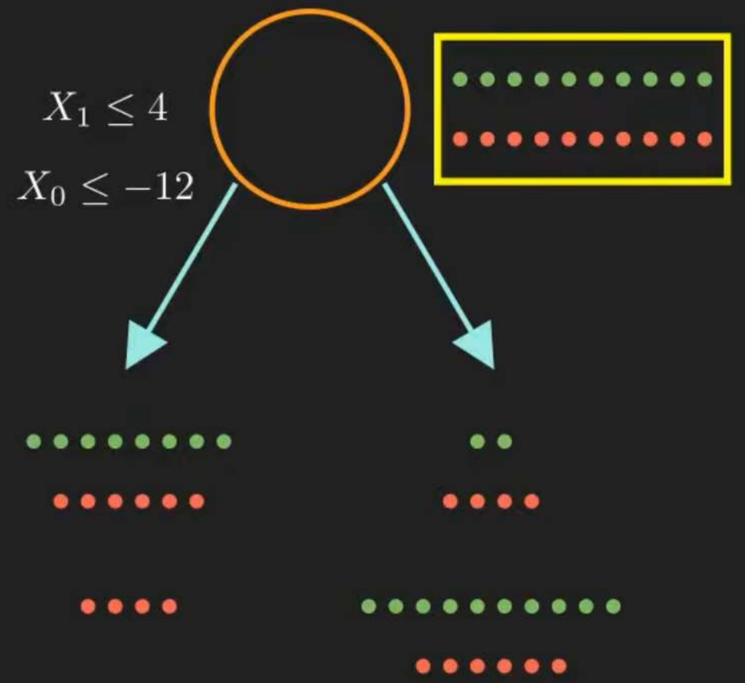




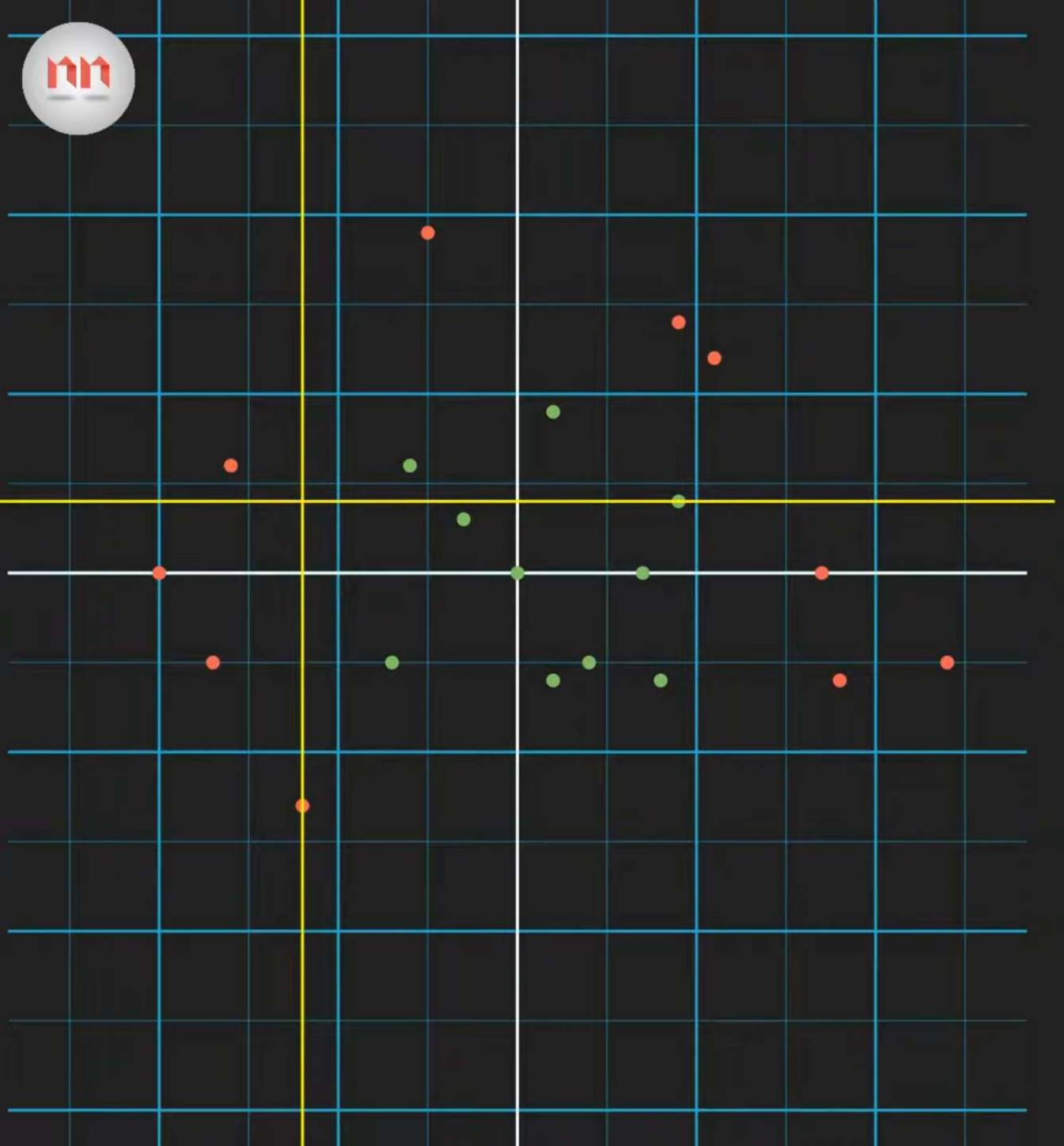
$$Entropy = \sum - \ p_i \ \log(p_i)$$

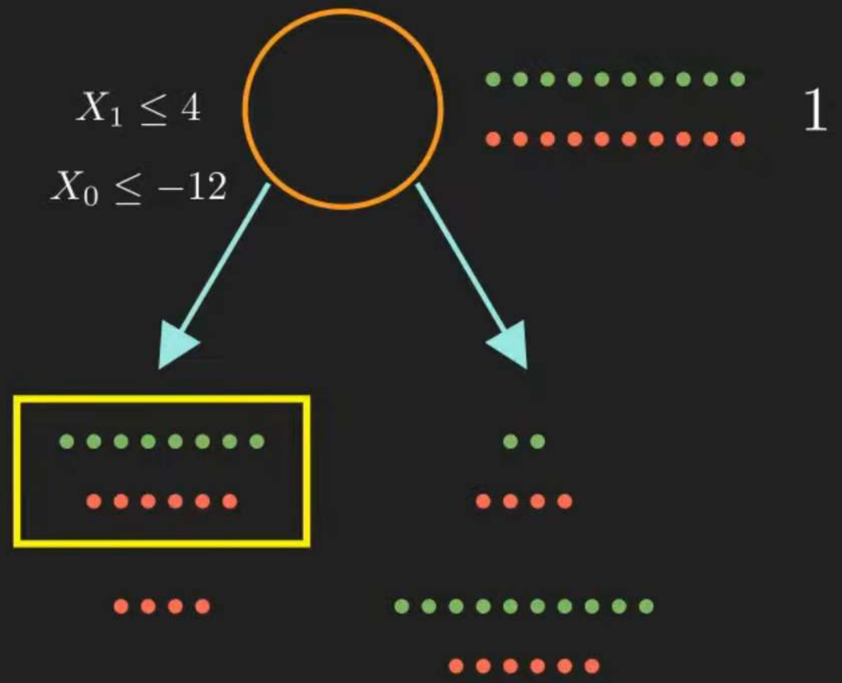
p_i = probability of class i



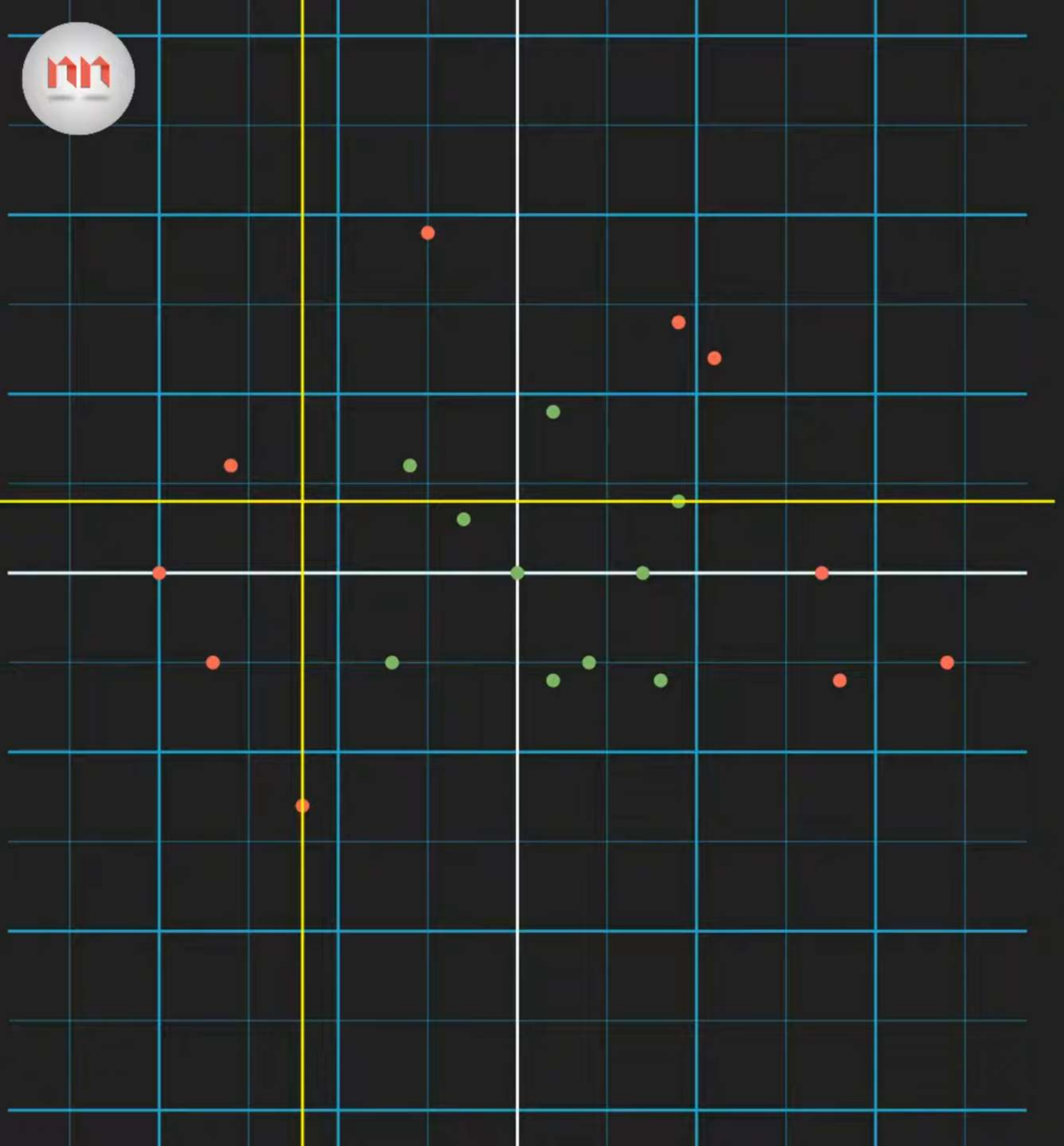


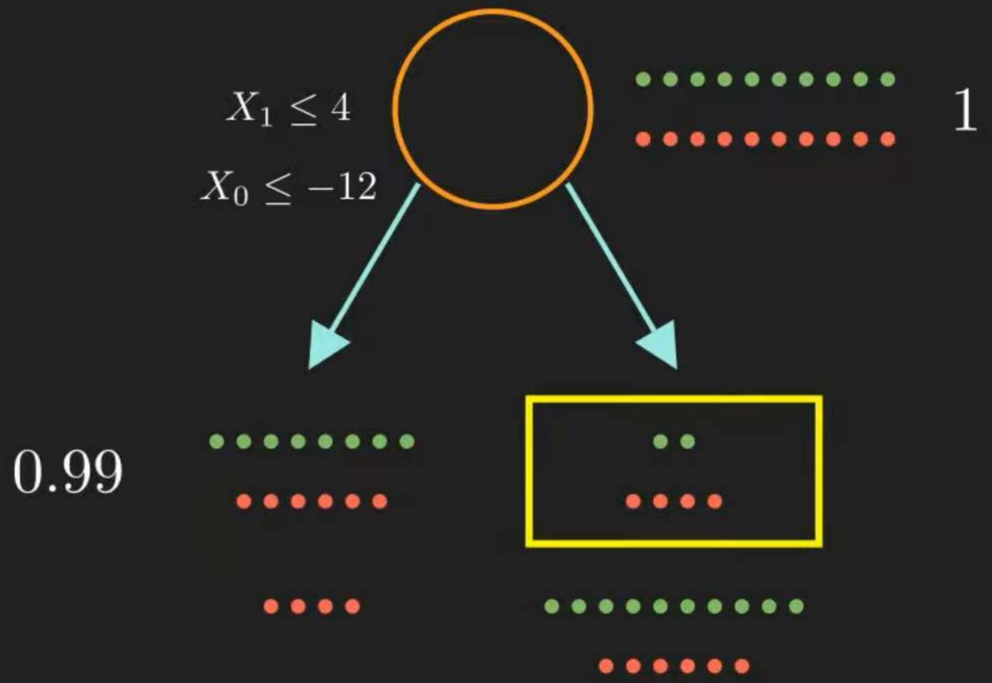
$$-0.5\log(0.5) - 0.5\log(0.5)$$



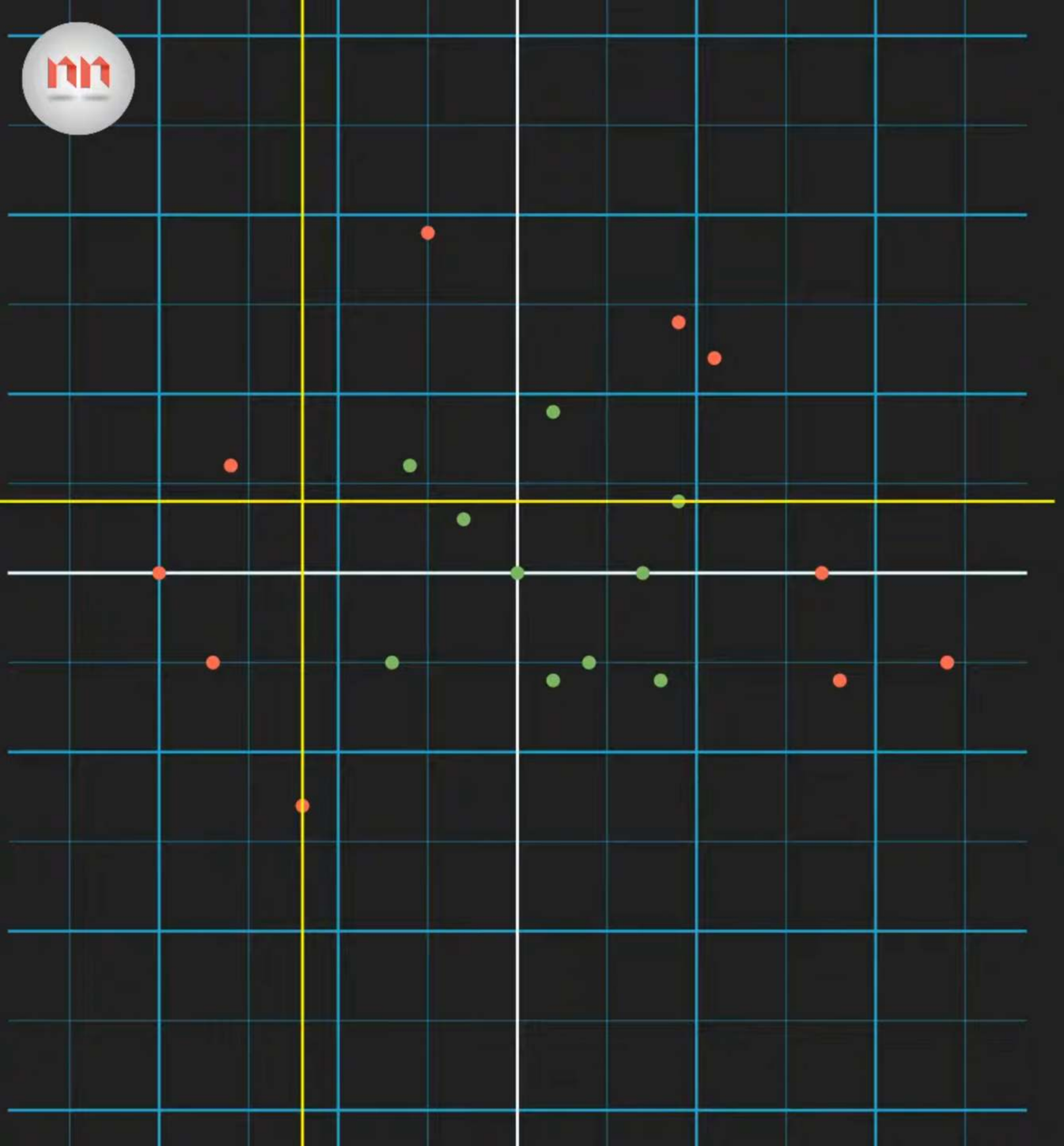


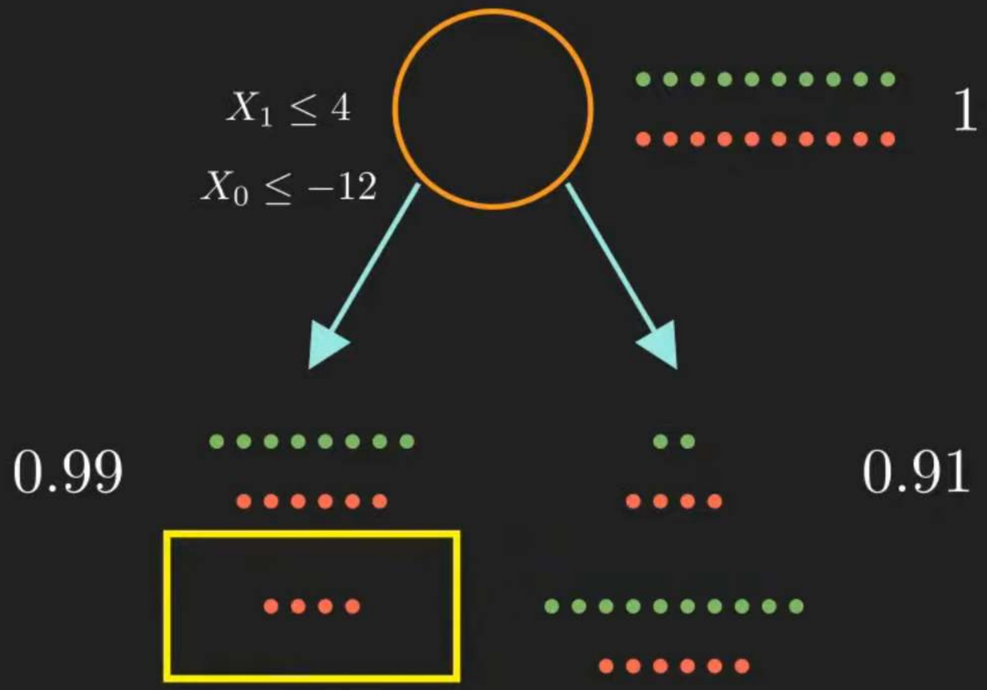
$$-0.57\log(0.57) - 0.43\log(0.43)$$



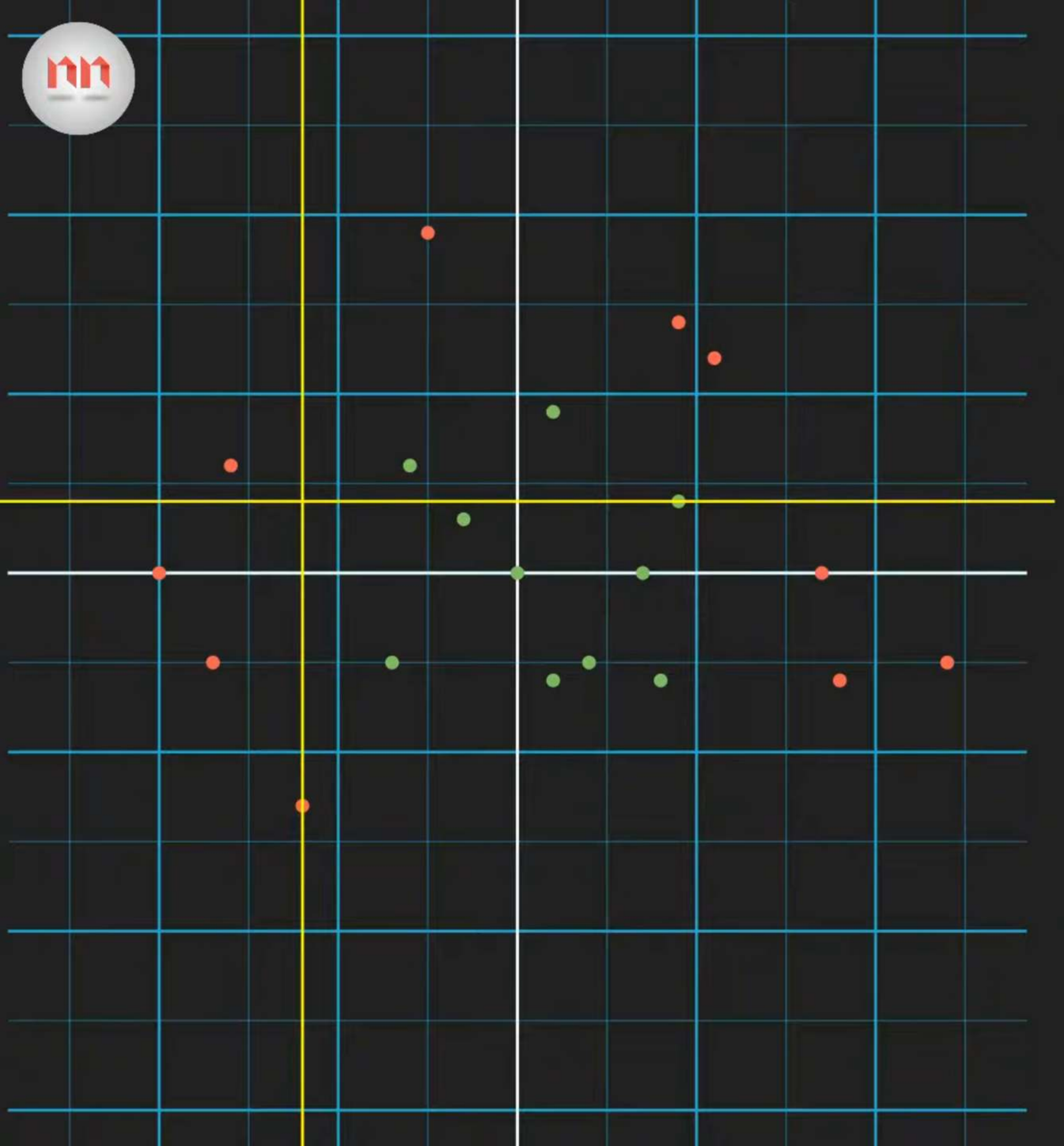


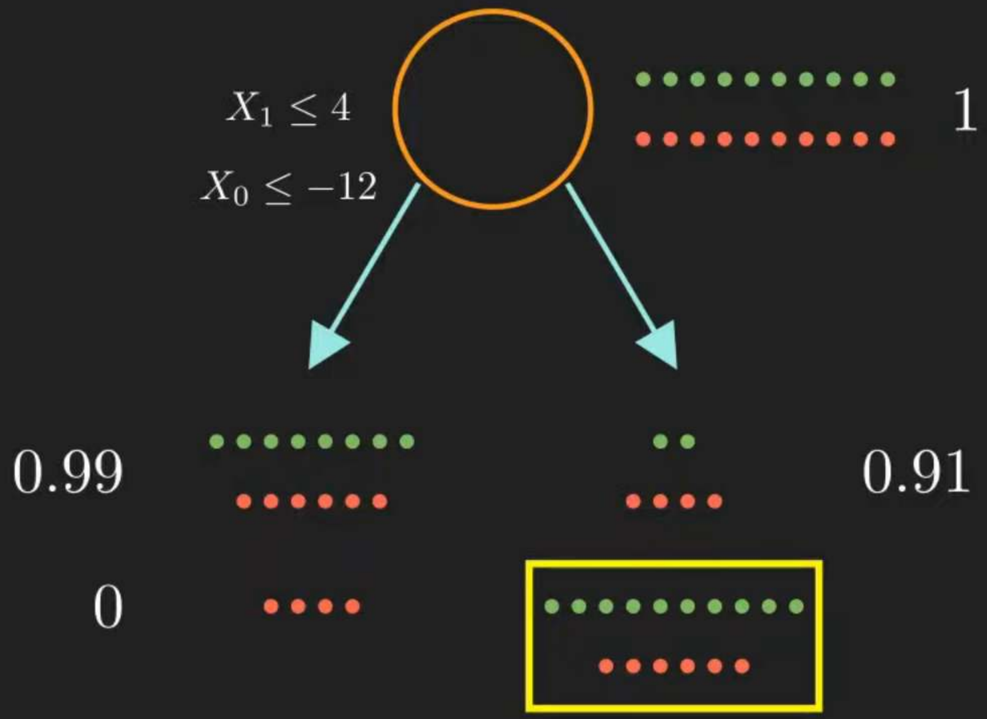
$$-0.33\log(0.33) - 0.67\log(0.67)$$



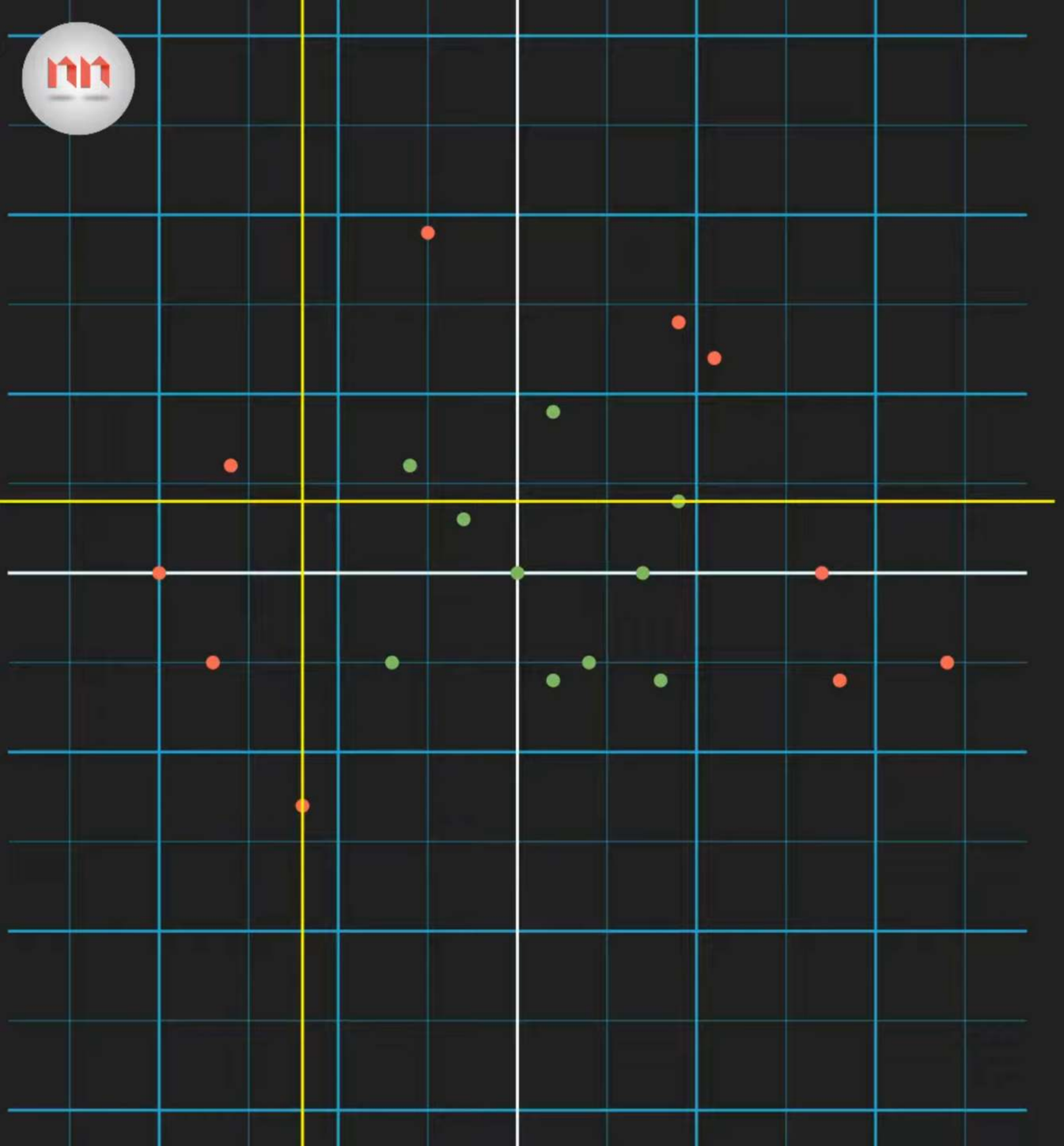


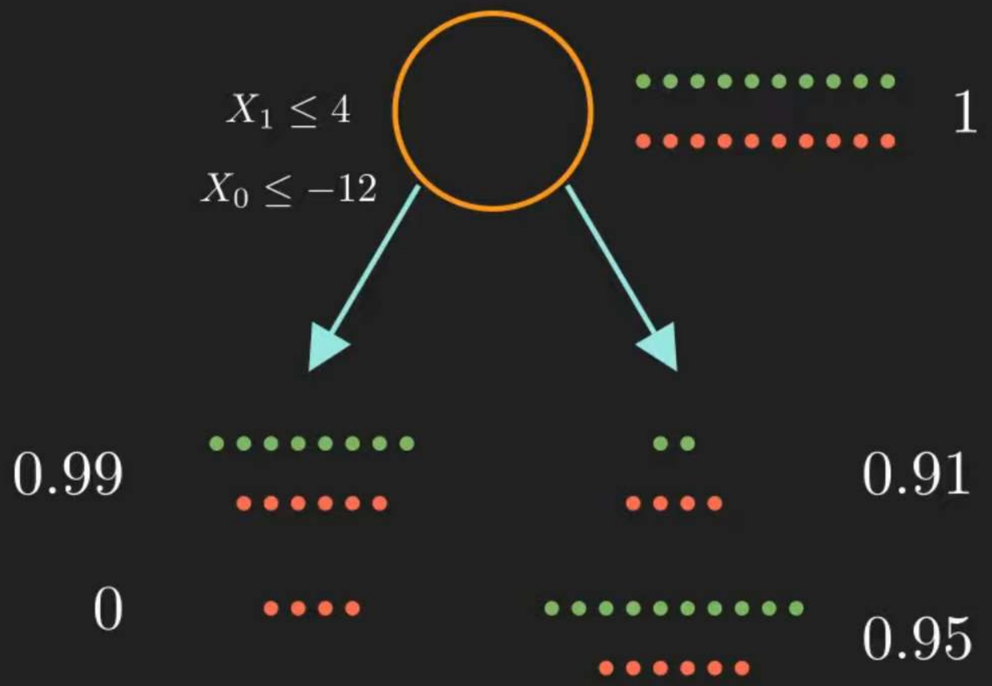
$$-1\log(1) - 0\log(0)$$

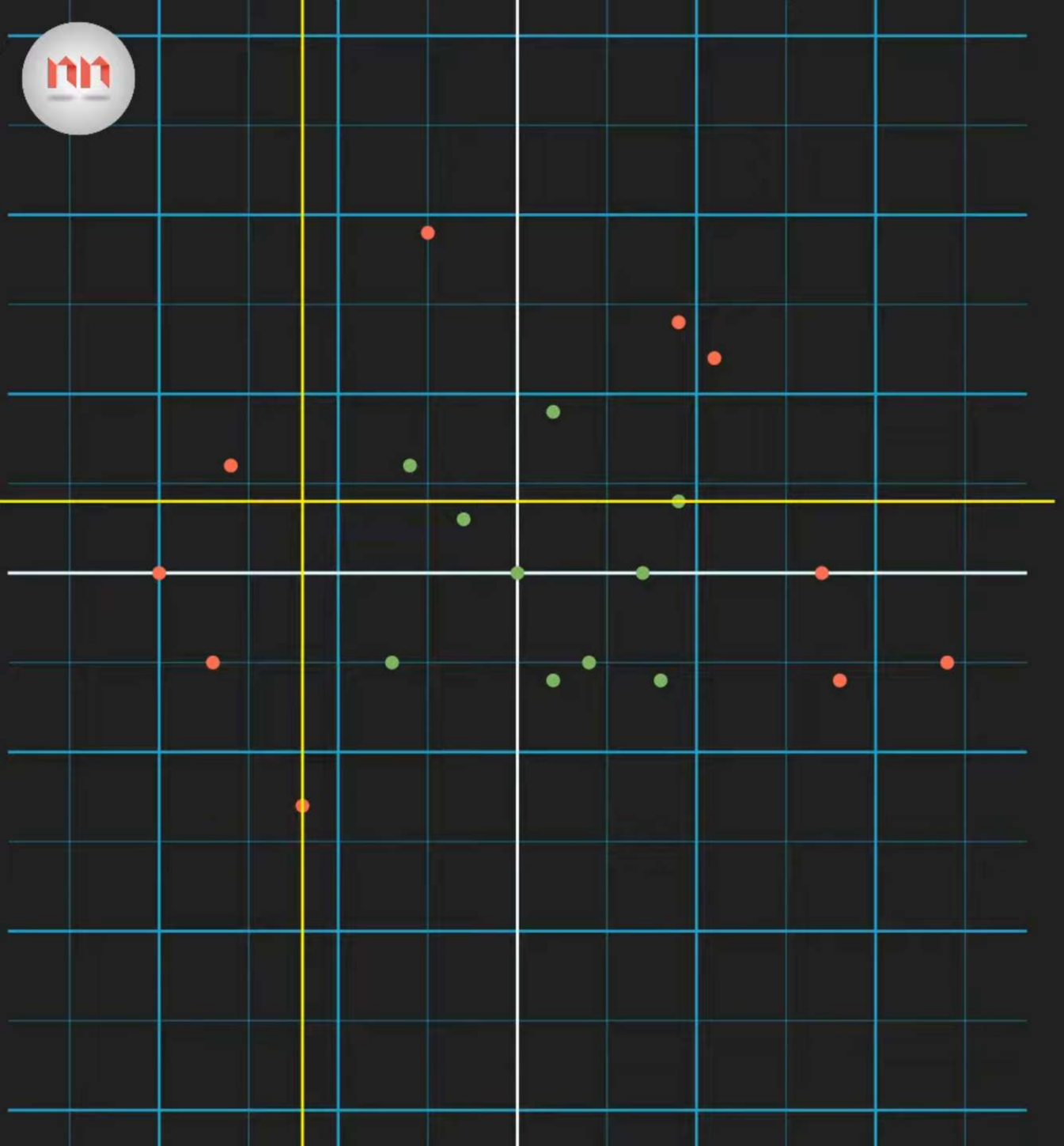


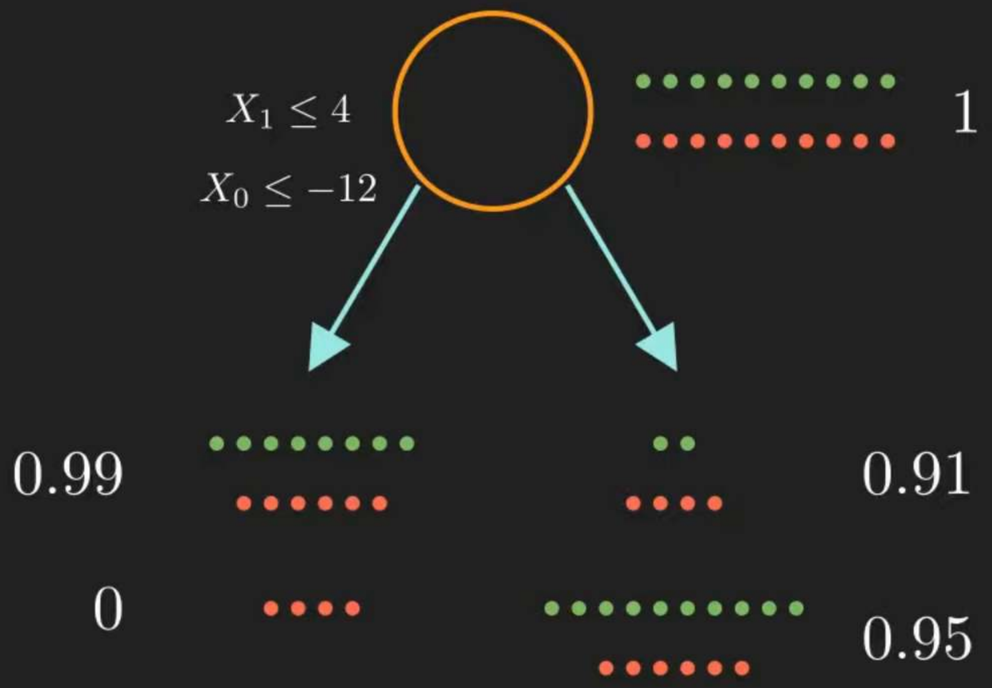


$$-0.63 \log(0.63) - 0.37 \log(0.37)$$





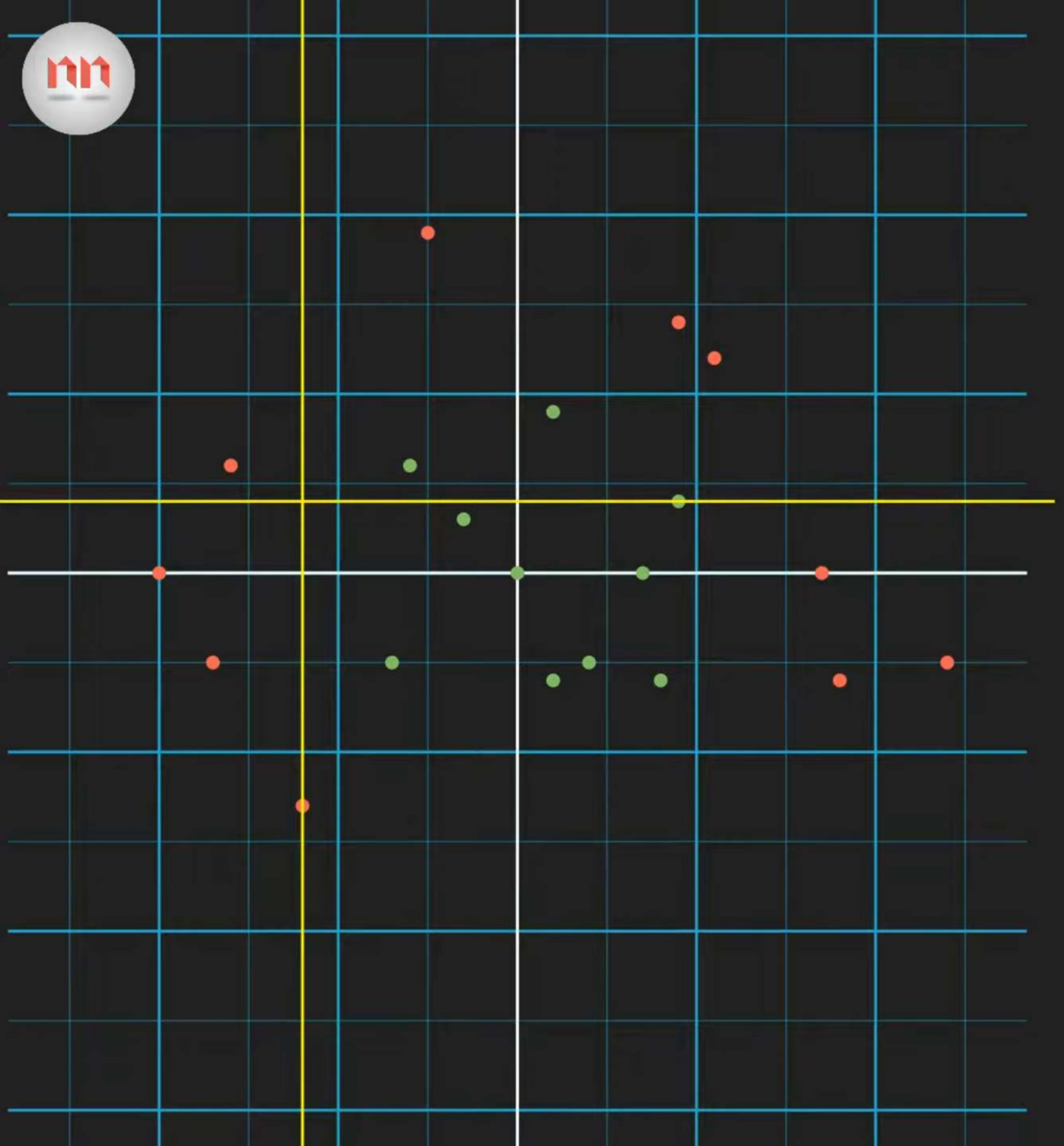


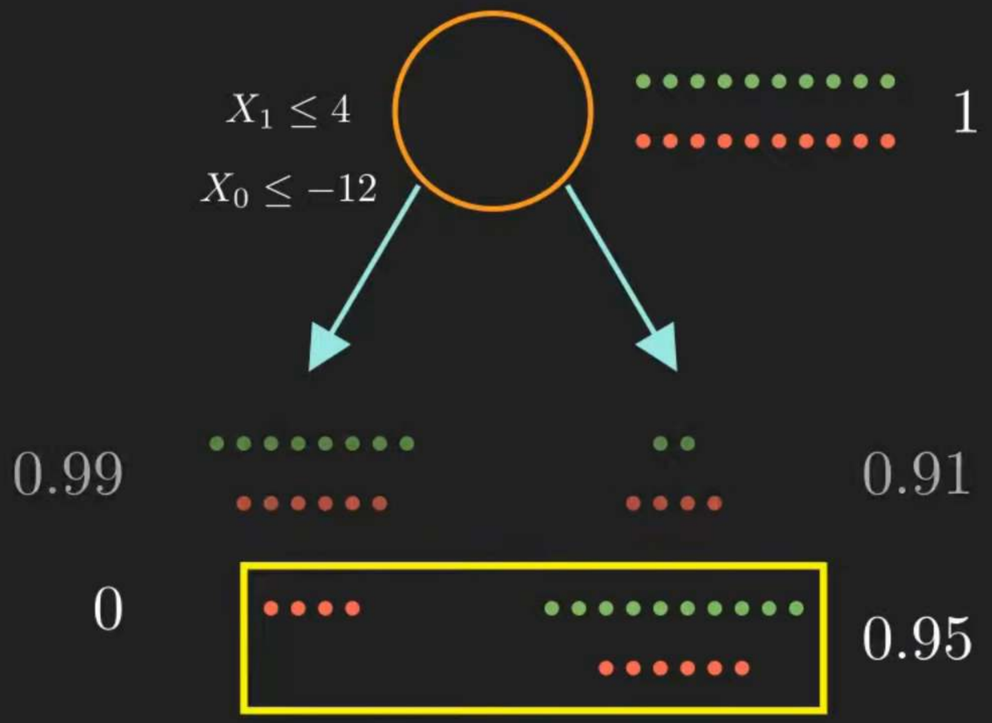


$$\text{IG} = E(parent) - \sum w_i \; E(child_i)$$

$$\text{IG}_1 = 1 - \frac{14}{20} \times .99 - \frac{6}{20} \times .91 = 0.034$$

$$\text{IG}_2 = 1 - \frac{4}{20} \times 0 - \frac{16}{20} \times .95 = 0.24$$





$$\mathrm{IG}_2 > \mathrm{IG}_1$$