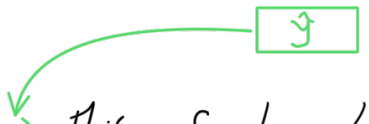
Gradient Boost

Grad. Boost

▽ Boost  STONKS

- Start by making a Single Leaf

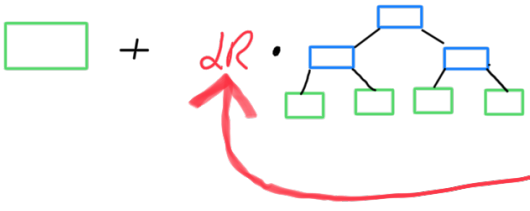


⇒ This Single leaf is a guess about the prediction of the samples

↳ Trivially this would be the Average $\left(\frac{\sum_i^N y_i}{N}\right)$

- Then build a Tree Next based on the errors from the last leafs

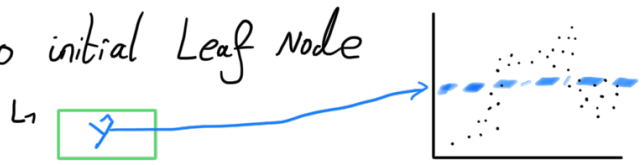
↳ Larger than a Stump. Still restricts the Number of Leaf Nodes



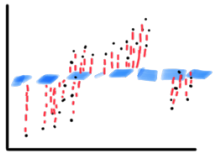
- Scales Trees All together (Learning rate)
- Continues Making Trees until it has reached a predefined Max or adding More Trees to the forest fails to improve fit

Actual Process:

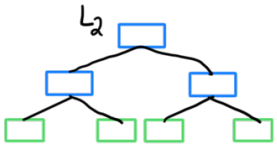
- Do initial Leaf Node



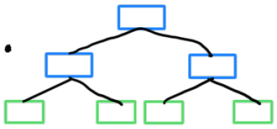
- Calculate psuedo-Residuals $(y - \hat{y})$



- Now build Tree To predict the psuedo residuals



- Calculate New psuedo-Residuals

↳ $y - \left(\square + \left(LR \cdot \right) \right)$ 

Repeat This Tree building and addition until you Reach a Maximum Tree amount or you fail To increas Accuracy

LR = Constant

|Leaf Nodes| = Constant