

Ensemble method (2)

Boosting

Ensemble method review

Problem: Trees are weak learner and trees overfit

Idea 1: Let's ensemble them

Idea 2: Let's decorrelate trees



Random Forest

Bagging and Random Forest

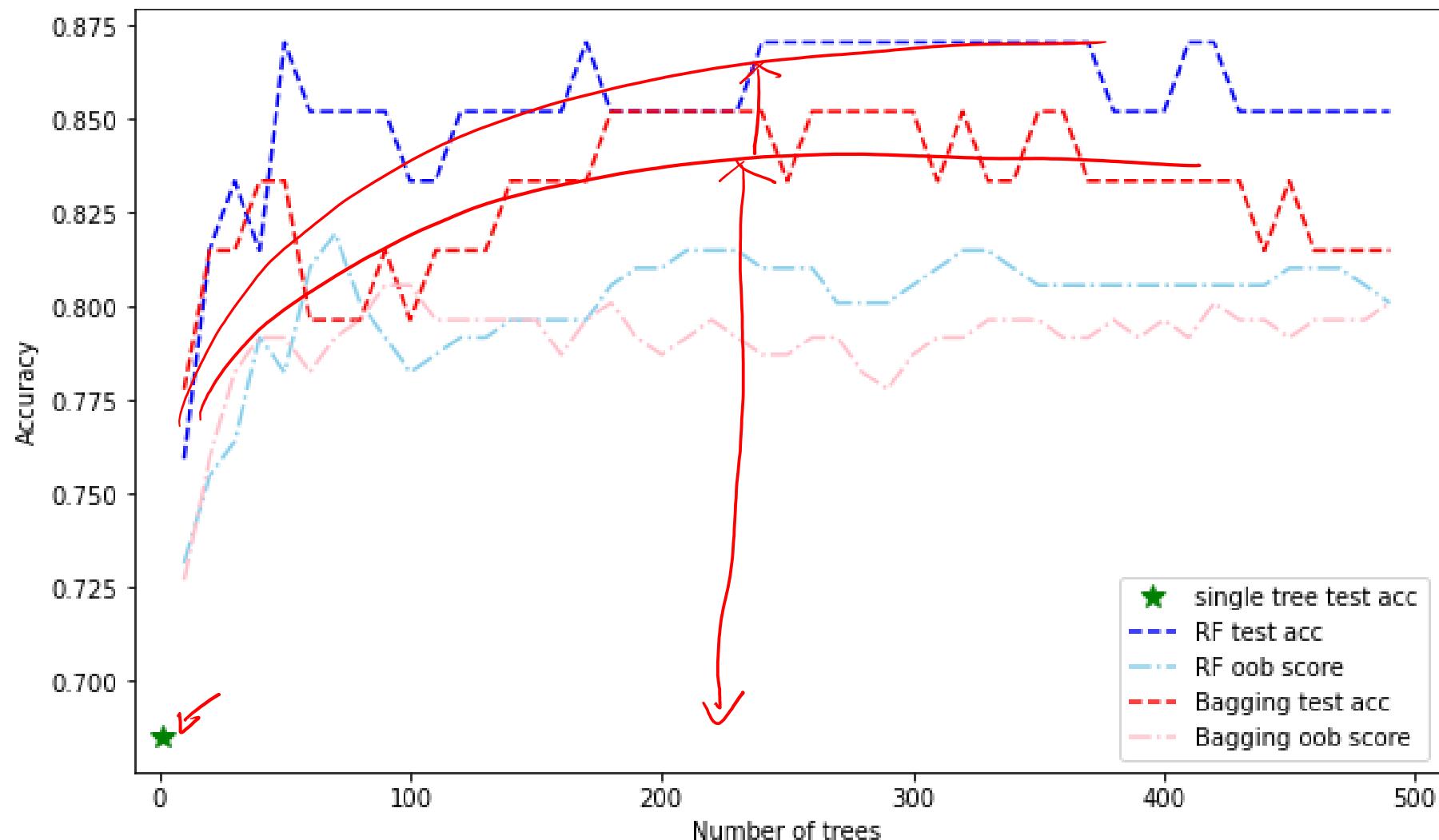
Bagging

- Random sample on data (row)
- Parallel ensembling

Random Forest

- Also on features (col)
- Further decorrelates the trees
- Parallel ensembling

Power of an ensemble of trees

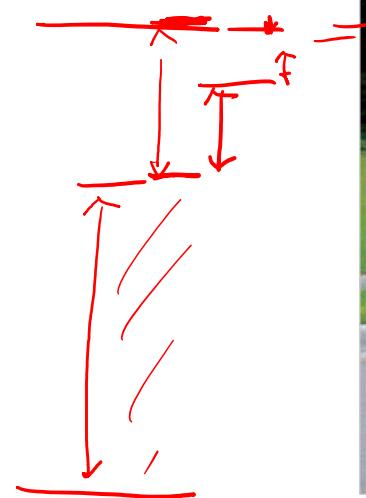
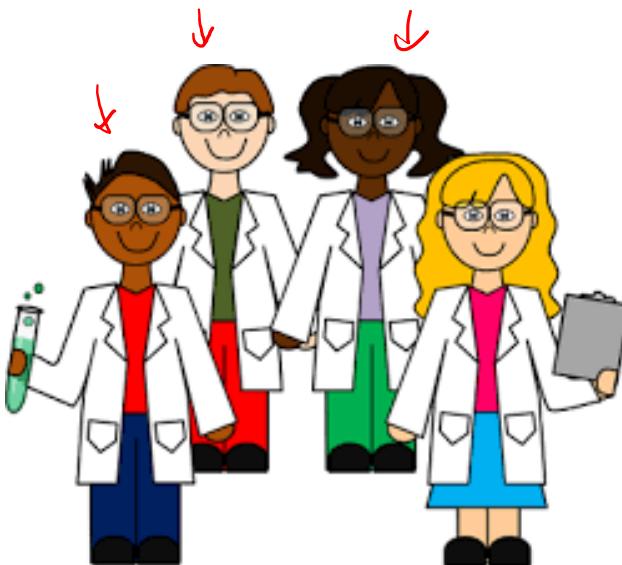


Boosting

Problem: Trees are weak learner and trees overfit

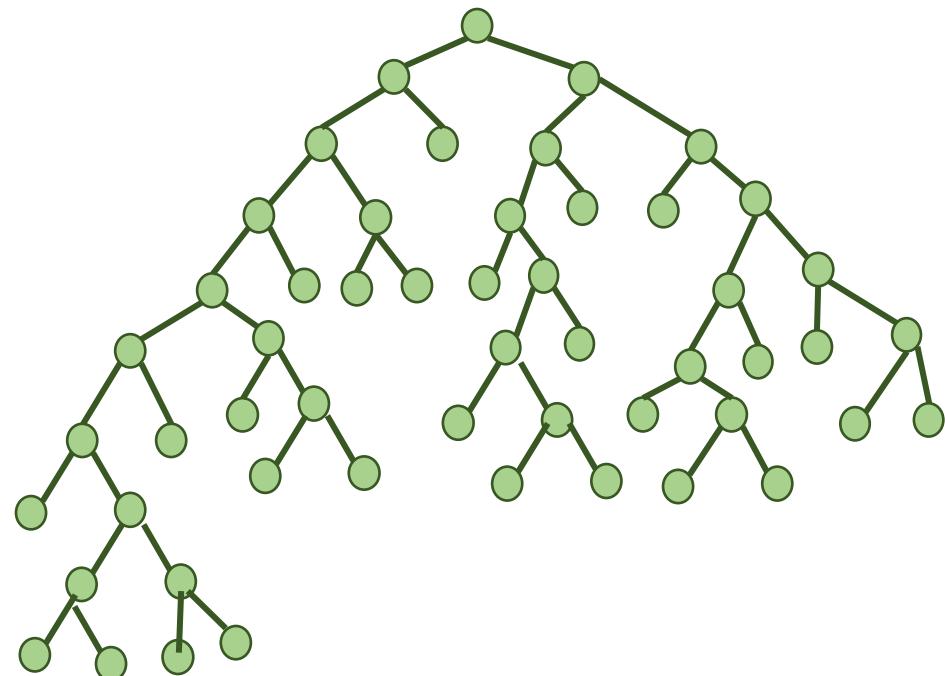
Idea 3: Let's make the trees a strong learner

How: Grow a small tree (stump) to fit residual

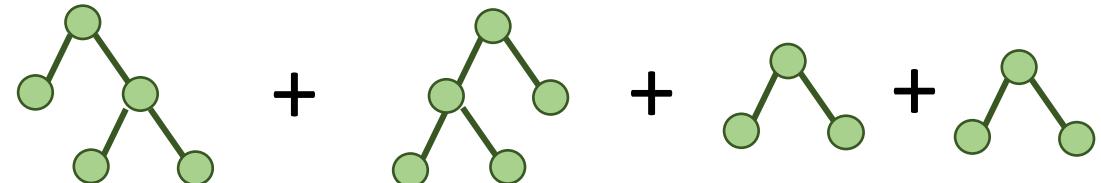


Boosting

Decision Tree



Boosting



Boosting algorithm

1. Initialize $f(x) = 0, r = y$

2. For $b = 1, 2, \dots, B$, repeat

a) Fit a tree $f_b(x)$ to the training data (X, r)

b) $\underline{f(x)} \leftarrow f(x) + \overset{\circ}{\lambda} \underline{f_b(x)}$

c) $r \leftarrow \underline{r} - \lambda \underline{f_b(x)}$

3. output $\sum_{b=1}^B \lambda f_b(x)$

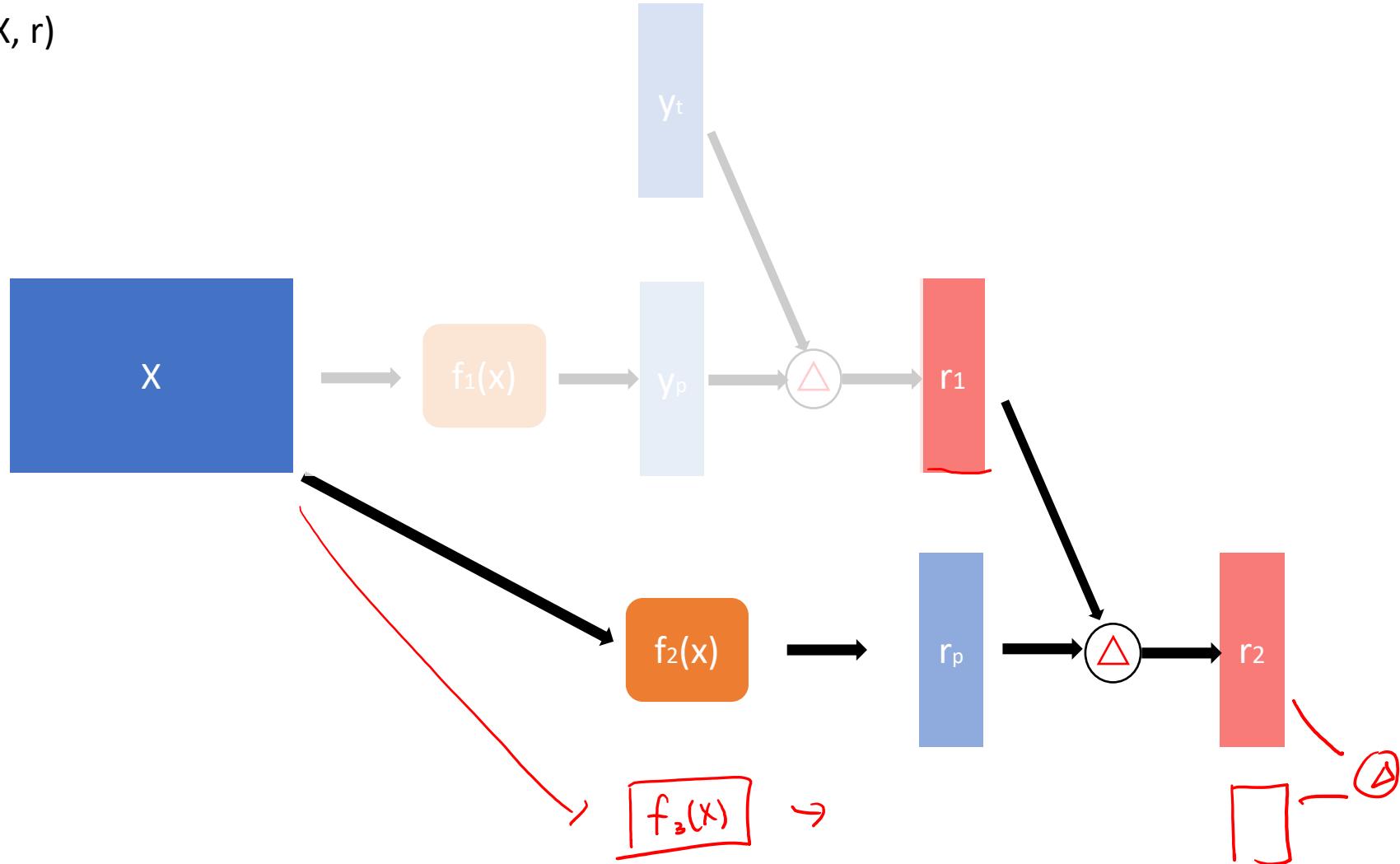
Boosting algorithm

For $b = 1, 2, \dots, B$, repeat

Fit a tree $f_b(x)$ to the training data (X, r)

$$f(x) \leftarrow f(x) + \lambda f_b(x)$$

$$\underline{r \leftarrow r - \lambda f_b(x)}$$



Boosting Methods

- AdaBoost
- Gradient Boost