

Machine Learning - Ensemble

ImpactDeal 2022

A dark blue diagonal gradient bar that starts from the bottom left and extends towards the top right, covering the lower half of the slide.

Boosting

1. **Boosting**
 - a. History
 - b. How Boosting Works
 - c. Difference with Random Forest
 - d. Popular Implementations
 - i. XGBoost
 - ii. LightGBM

History

Randomized Trees

Ho proposes the idea of aggregating predictions of trees trained on randomly selected features



1995

History

Randomized Trees

Ho proposes the idea of aggregating predictions of trees trained on randomly selected features

1995

1997

AdaBoost

First successful realization of boosting by Freund and Schapire

History

Randomized Trees

Ho proposes the idea of aggregating predictions of trees trained on randomly selected features

1995

Gradient Boosting Machines

Friedman reformulates and generalizes boosting algorithms.

1999

1997

AdaBoost

First successful realization of boosting by Freund and Schapire



History

Randomized Trees

Ho proposes the idea of aggregating predictions of trees trained on randomly selected features

1995

1997

AdaBoost

First successful realization of boosting by Freund and Schapire

Gradient Boosting Machines

Friedman reformulates and generalizes boosting algorithms.

1999

2001

Random Forests

Breiman combines bagging and random feature selection

History

Randomized Trees

Ho proposes the idea of aggregating predictions of trees trained on randomly selected features

Gradient Boosting Machines

Friedman reformulates and generalizes boosting algorithms.

XGBoost

Researcher Tianqi Chen wins the Higgs Challenge competition on Kaggle with his version of gradient boosting and open source it

1995

1997

1999

2001

2014

2016

AdaBoost

First successful realization of boosting by Freund and Schapire

Random Forests

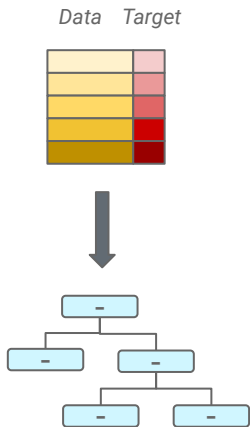
Breiman combines bagging and random feature selection

LightGBM

Microsoft publishes its open source version of gradient boosting machines

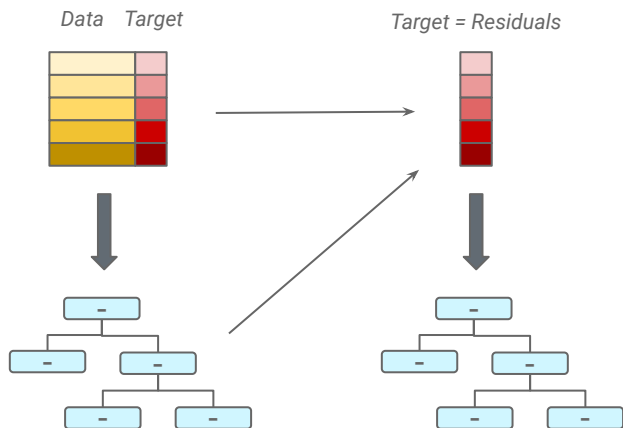
How Boosting Works

Boosting creates an ensemble model, **sequentially** combining a series of **weak learners**.



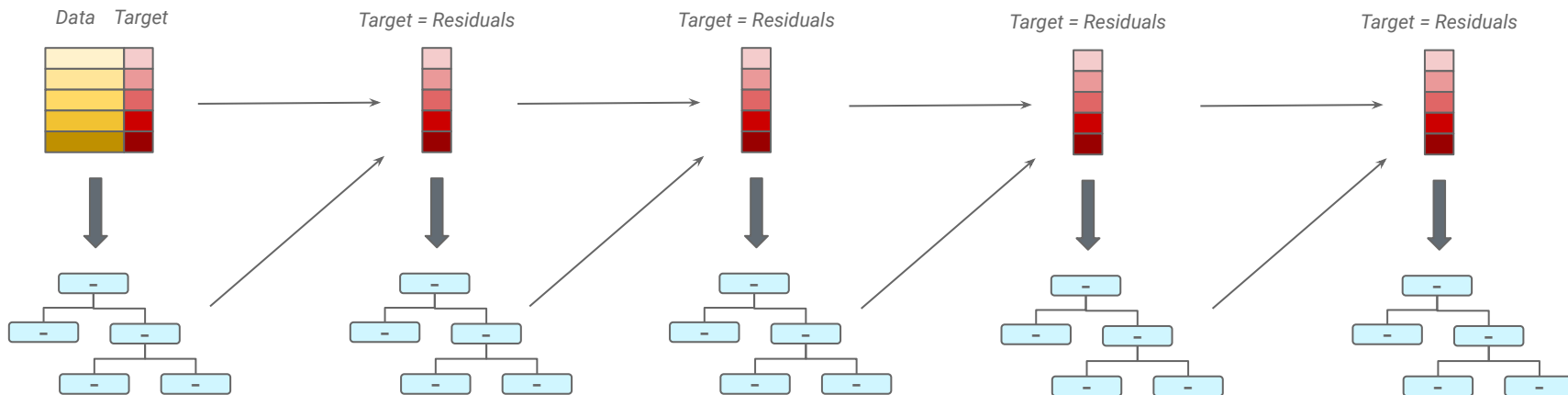
How Boosting Works

Boosting creates an ensemble model, **sequentially** combining a series of **weak learners**.



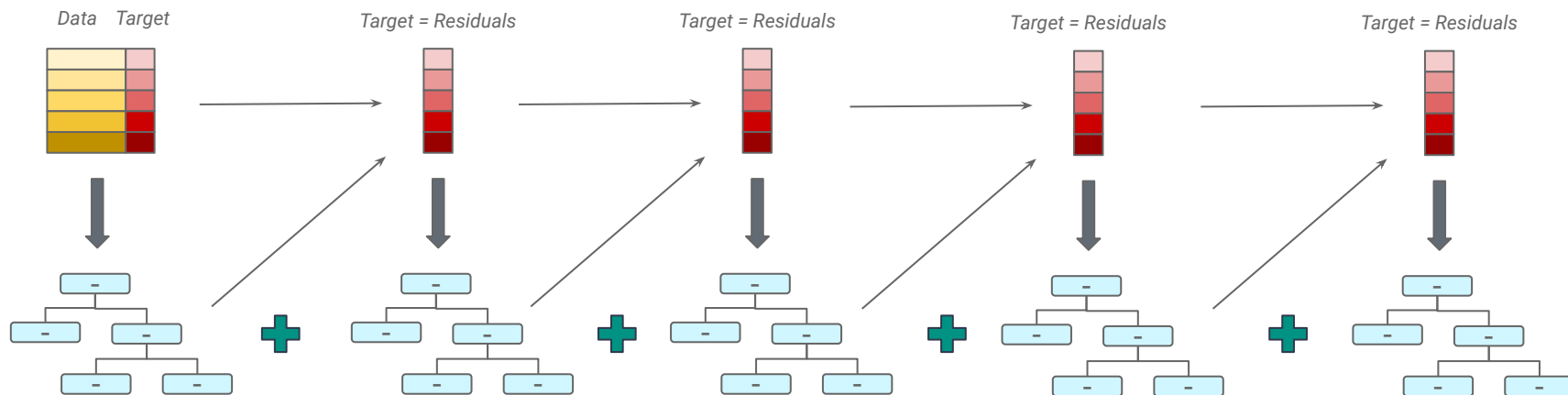
How Boosting Works

Boosting creates an ensemble model, **sequentially** combining a series of **weak learners**.

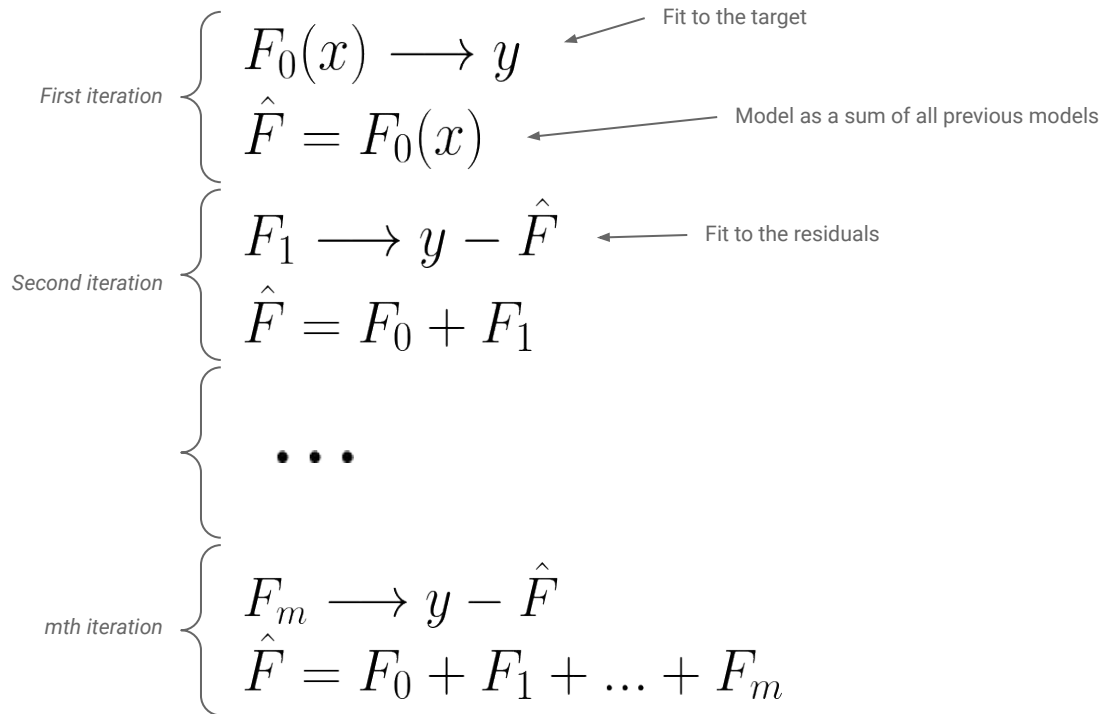


How Boosting Works

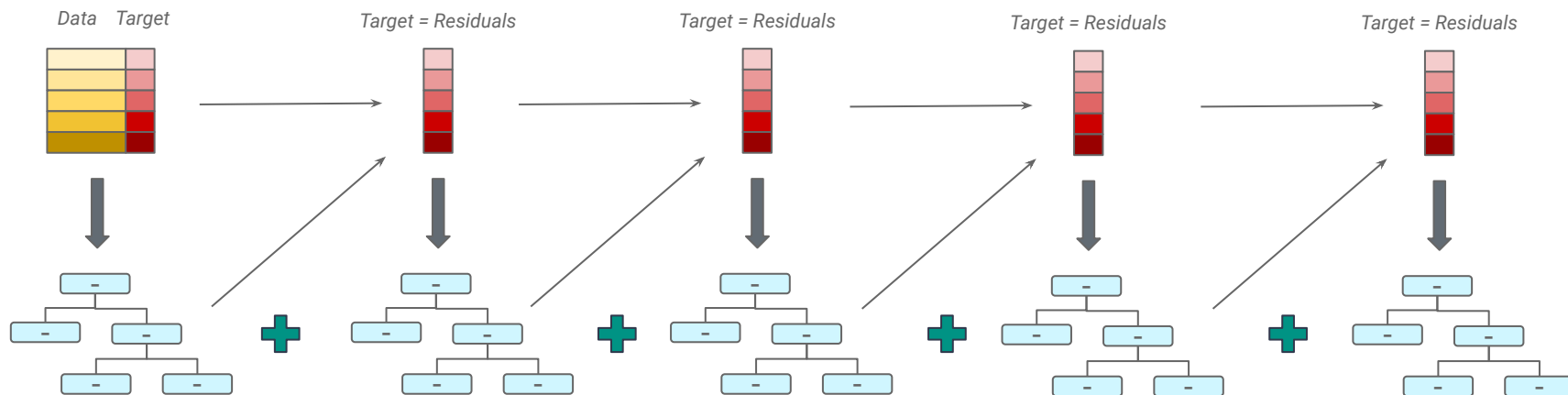
Boosting creates an ensemble model, **sequentially** combining a series of **weak learners**.



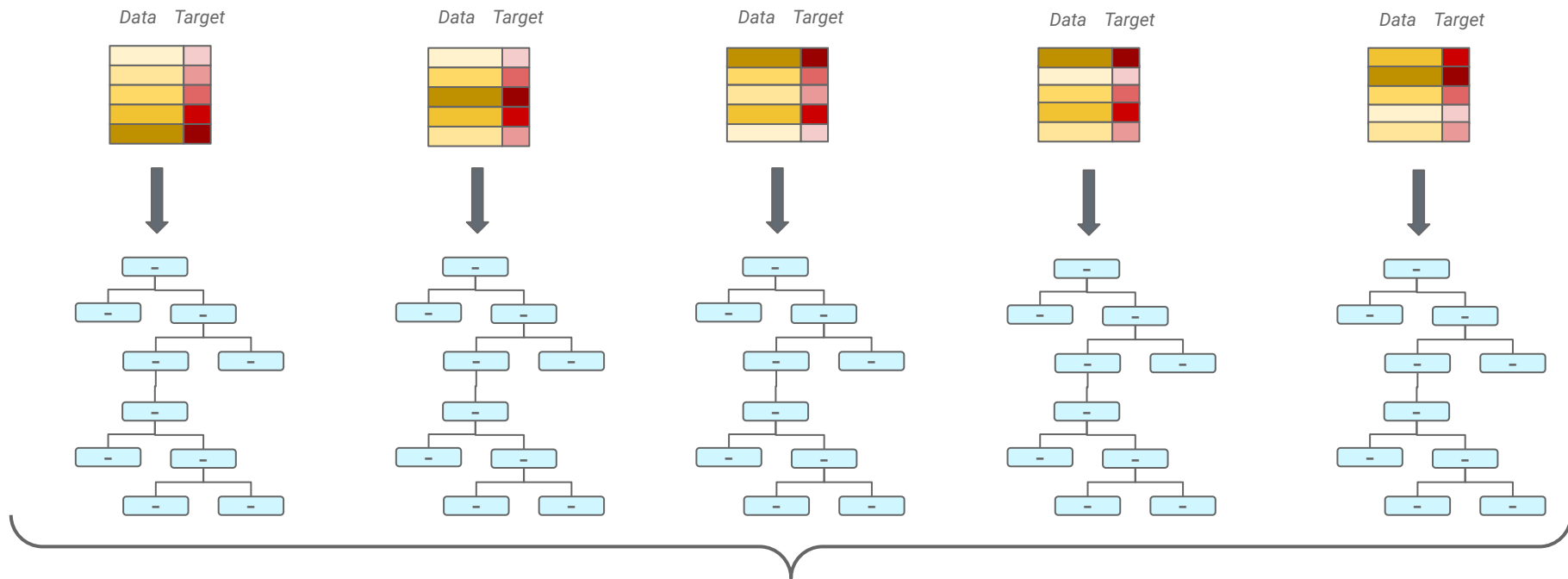
How Boosting Works



Difference with Random Forest



Difference with Random Forest



Difference with Random Forest

Gradient Boosting Trees

- Trees are grown sequentially
- Additive model made of weak learners
- Trees are fitted on a re-weighted version of the dataset

Random Forests

- Trees are grown in parallel
- Predictions of full-grown trees are averaged
- Trees are fitted on bootstrap samples of the dataset

Popular Implementations

XGBoost

- Further randomization and regularization
- Parallel processing and system optimisations
- Several loss functions available
- Handling of missing values
- Native support of categorical features
- GPU and distributed computations supported

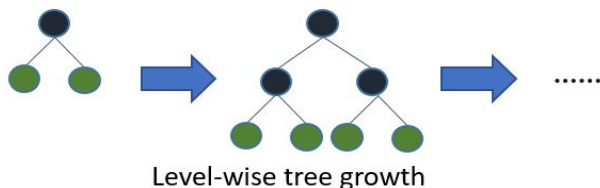
LightGBM

- Further randomization and regularization
- Parallel processing and system optimisations
- Several loss functions available
- Handling of missing values
- Native support of categorical features
- GPU and distributed computations supported

Popular Implementations

XGBoost

- Further randomization and regularization
- Parallel processing and system optimisations
- Several loss functions available
- Handling of missing values
- Native support of categorical features
- GPU and distributed computations supported



LightGBM

- Further randomization and regularization
- Parallel processing and system optimisations
- Several loss functions available
- Handling of missing values
- Native support of categorical features
- GPU and distributed computations supported

