

Gradient Boosting



Machine Learning



Sunit

Astonishing!!!!

**Kaggle Grandmasters
win competitions**

**Outperforms
deep learning
models**

**Wins the \$1 million
Netflix Prize in 2009**

**Cause for rapid
advancements in
recent years**

**Achieves state-of-the-
art performance**



Sunit

Implementations

XGBoost

Scalable and fast gradient boosting

LightGBM

High-performance gradient boosting

CatBoost

Handles categorical features efficiently

GradientBoosting Regressor

Built-in sklearn package

H2O

Scalable and distributed



Sunit

vs Random Forest

Feature	GB	RF
Type	Sequential	Parallel
Optimization	Gradient	Bootstrap aggregation
Sample size	No replacement	Replacement
Feature selection	High correlation	Randomness
Interpretability	Low	High



Gradient

Refers to the gradient of the loss function that is being optimized

The loss function measures the difference between the predicted values of the model and the actual values of the target variable

The goal is to iteratively improve the model's predictions by minimizing the loss function

Gradient of the loss function used to "boost" the performance of a weak learner



Boosting

Class of algorithms that combine weak learners into a strong learner by...

Iteratively adding new models that focus on the misclassified samples from previous models

Results in a sequence of models that gradually improve the overall prediction accuracy by reducing the bias and variance of the final ensemble

Coined by Jerome Friedman in a 1999 paper titled "Greedy Function Approximation: A Gradient Boosting Machine"



Gradient Boosting

Powerful algorithm, type of ensemble learning method

A series of weak models, usually decision trees, are trained sequentially, with each subsequent model built to correct the errors made by the previous model

Iteratively optimizes the loss function by following the gradient of the loss function to minimize the error of the model's predictions

The goal is to iteratively improve the model's predictions by minimizing the loss function.



Advantages

High Accuracy

Accurate predictions

Handles Mixed Data

Categorical &
continuous

Feature Importance

Identifies important
features

Missing data

no need for imputation

Robust to Outliers

Due to iterative approach



Disadvantages

**Computationally
expensive**
long training time

Overfitting
if not tuned

Prone to bias
dependent on data

Black box model
hard to interpret

Large memory usage
space requirements



Follow

Sunit Ghosh

**to get #interesting
and latest #titbits
on #java, #AiMl, #cloud
technologies**