Cheat sheet

XGBoost

XGBoost

A state-of-the-art machine learning model built for great accuracy. However, needs to be adjusted for Time Series problems.

Ensemble Learning

A Machine Learning framework that consists in combining multiple models, usually poor/weak. The combination leads to a stronger model than each one alone.

Boosting

Mechanism to improve accuracy. Consists in only looking at part of the observations at once for each model. Incorrectly predicted observation gain weight for future iterations of the model.

Feature Sampling

A mechanism that only includes a fraction of the predictors. The features in models with higher accuracy gain importance.

Minimum Child Weight

A parameter that states how easily the trend curve should change.Relates to the sum of the weights of each observation. Low values can mean that maybe not a lot of observations are in the round.

ETA

Learning Rate. How fast do youwant the model to learn?

Max Depth

How big should the tree be? Bigger trees go into more detail

Gamma

How fast should the tree be split?

Subsample

Share of observations in each tree

Colsample by tree

How much of the tree should be analysed per round?

Number of rounds

How many times do wewant the analysis to be run?

Diogo Resende