1 Extreme Gradient Boosting

Extreme Gradient Boosting is a gradient boosting decision tree algorithm. Boosting refers to an additive training technique where starting with some base classifiers $\hat{y}^{(0)} = f_0(x_0)$, an objective function

$$Objective\ function = \sum_{i=1}^{N} l(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k)$$

that consists of a loss function $l(y_i, \hat{y}_i)$ and a regularization term $\Omega(f_k)$ gets optimized by adding at each iterative stage a term to the classifier function that compensates for the errors of the previous model.

$$
\begin{aligned}
\hat{y}^{(0)} &= 0 \\
\hat{y}^{(1)} &= \hat{y}^{(0)} + f_1(x_1) = f_1(x_1) \\
\hat{y}^{(2)} &= \hat{y}^{(1)} + f_2(x_2) = f_1(x_1) + f_2(x_2) \\
&\vdots \\
\hat{y}^{(t)} &= \sum_{k=1}^{K} f_k(x_k) = \hat{y}^{(t-1)} + f_t(x_t)
\end{aligned}
$$

This sequential procedure stops when no further improvements can be made, e.g. when the reduction of the loss function is smaller then the change in the regularization term.

The main reason why to implement Xgboost is because of its execution speed and model performance (proofed in various Kaggle competitions) [**?**].

... bla bla bla