Since the response variable only have 0 and 1, y is the true value and is the predicted value, the loss function is:

，so,

Generate *m*=1, 2…, *M* trees, and in each tree iterate *i*=1, 2…, N (Sample size), fit the tree with training data and calculate the residual by the following gradient formula:

Use to fit another new tree and minimize the Loss function:

The new model

In case of the overfitting, we add a shrinkage parameter, so the new model is )，and the final tree can be interpreted as below: