

# GBM

Sunday, September 16, 2018 8:21 PM

<https://www.analyticsvidhya.com/blog/2016/02/complete-guide-parameter-tuning-gradient-boosting-gbm-python/>

1. **min\_samples\_split = 500** : This should be ~0.5-1% of total values. Since this is imbalanced class problem, we'll take a small value from the range.
2. **min\_samples\_leaf = 50** : Can be selected based on intuition. This is just used for preventing overfitting and again a small value because of imbalanced classes.
3. **max\_depth = 8** : Should be chosen (5-8) based on the number of observations and predictors. This has 87K rows and 49 columns so let's take 8 here.
4. **max\_features = 'sqrt'** : It's a general thumb-rule to start with square root.
5. **subsample = 0.8** : This is a commonly used start value

From <<https://www.analyticsvidhya.com/blog/2016/02/complete-guide-parameter-tuning-gradient-boosting-gbm-python/>>