LightGBM参数介绍

Xgboost和LightGBM部分参数对照：

Xgboots	LightGbm
booster(default=gbtree)	boosting(**default=**gbdt)
eta(default=0.3)	learning_rate(default=0.1)
max_depth(default=6)	num_leaves(default=31)
min_child_weight(default=1)	min_sum_hessian_in_leaf(1e-3)
gamma(default=0)	min_gain_to_split(default=20)
subsample(default=1)	bagging_fraction(default=1.0)
colsample_bytree(default=1)	feature_fraction(default=1.0)
alpha(default=0)	lambda_l1(default=0)
lambda(default=1)	lambda_l2(default=0)
objective(default=reg:linear)	application(default=regression)
eval_metric(default according to objective)	metric
nthread	num_threads

1. 使用num_leaves

因为LightGBM使用的是leaf-wise的算法，因此在调节树的复杂程度时，使用的是num_leaves而不是max_depth。
大致换算关系：num_leaves = 2^(max_depth)。它的值的设置应该小于2^(max_depth)，否则可能会导致过拟合。

2.对于非平衡数据集：

可以param['is_unbalance']='true'

3. 3.Bagging参数：

bagging_fraction+bagging_freq（必须同时设置）、feature_fraction。bagging_fraction可以使bagging的更快的运行出结果，feature_fraction设置在每次迭代中使用特征的比例。

4.min_data_in_leaf：

这也是一个比较重要的参数，调大它的值可以防止过拟合，它的值通常设置的比较大。

5.max_bin:

调小\max_bin的值可以提高模型训练速度，调大它的值和调大num_leaves起到的效果类似。

参数速查：

xgb	lgb	xgb.sklearn	lgb.sklearn
booster='gbtree'	boosting='gbdt'	booster='gbtree'	boosting_type='gbdt'
objective='binary:logistic'	application='binary'	objective='binary:logistic'	objective='binary'
max_depth=7	num_leaves=2**7	max_depth=7	num_leaves=2**7
eta=0.1	learning_rate=0.1	learning_rate=0.1	learning_rate=0.1
num_boost_round=10	num_boost_round=10	n_estimators=10	n_estimators=10
gamma=0	min_split_gain=0.0	gamma=0	min_split_gain=0.0
min_child_weight=5	min_child_weight=5	min_child_weight=5	min_child_weight=5
subsample=1	bagging_fraction=1	subsample=1.0	subsample=1.0
colsample_bytree=1.0	feature_fraction=1	colsample_bytree=1.0	colsample_bytree=1.0
alpha=0	lambda_l1=0	reg_alpha=0.0	reg_alpha=0.0
lambda=1	lambda_l2=0	reg_lambda=1	reg_lambda=0.0
scale_pos_weight=1	scale_pos_weight=1	scale_pos_weight=1	scale_pos_weight=1
seed	bagging_seed feature_fraction_seed	random_state=888	random_state=888
nthread	num_threads	n_jobs=4	n_jobs=4
evals	valid_sets	eval_set	eval_set
eval_metric	metric	eval_metric	eval_metric
early_stopping_rounds	early_stopping_rounds	early_stopping_rounds	early_stopping_rounds
verbose_eval	verbose_eval	verbose	verbose

参数速查：