横向安全提升树算法参数介绍文档

|  |  |  |  |
| --- | --- | --- | --- |
| 参数名 | 参数含义 | 默认值 | 可选值 |
| penalty | 正则化，防止模型过拟合。通过在损失函数后加L1或者L2范式的倍数实现。这个增加的范式，被称为“正则项”，也被称为"惩罚项"。 | L2 | L1,L2 |
| task\_type | 任务类型 | classification | 'classification', 'regression' |
| learning\_rate | 学习率 | 0.3 | float、int或long |
| num\_tress | 最大boost轮数 | 5 | 大于零的int值 |
| n\_iter\_no\_change | 残余误差小于tol，建树过程将停止 | True | True/False |
| bin\_num | 在quantile中使用的bin号码 | 32 | int, 大于1的正整数 |
| zero\_as\_missing | 0作为缺失值 | False | True/False |
| subsample\_random\_seed | 加密随机种子 | None | 随机整数 |
| subsample\_feature\_rate | 一个浮点数，不知何用 | 0.8 | [0,1] |
| tree\_param | 单颗决策树的参数 |  |  |
| tree\_param.max\_depth | 每一颗决策树的最大深度 | 3 | 任何一个int值 |
| tree\_param.min\_sample\_split | 分割节点最小的数量，也即是说一个节点上最少有几个样本才可以分裂 | 2 | 任意一个int值 |
| tree\_param.min\_sample\_split | 单个节点分裂所要达到的最小值，也就是打分的最小值小于这个最小值不会进行分裂 | 1e-3 | 大于0任意实数 |
| tree\_param.min\_child\_weight | 子节点所需的二阶导数矩阵的值 | 0 | 大于零的任意一个实数 |
| tree\_param.max\_split\_nodes | 出于对内存容量的考虑，将使用不超过这个max\_split\_nodes的并行节点数量来寻找切分点 | 65536 | 大于零的任意int |
| tree\_param.feature\_importance\_type | 如果是 "split"，feature\_importance按特征分割次数计算。  如果是'gain'，则特征重要性按特征分裂的增益计算 | 'split' | 'split', 'gain' |
| tree\_param.using\_missing | 是否使用缺失值来训练 | False | True/False |
| tree\_param.zero\_as\_missing | 是否将0值认为是缺失值 | False | True/False |
| tree\_param.deterministic | 保证直方图的稳定性，设置为 "true "可以确保在使用相同的数据和相同的参数时有稳定的结果。  相同的数据和相同的参数。但它可能会减慢计算速度 | False | True/False |