现有的文档编写和调试：

多写一个stacking：

第一层：XGB+Linear+RF+GBDT+SVM

第二层：XGB

需要有对输出集的评估标准

调试一下

**Common parameters**

* nan\_mode (string): 处理输入数据中缺失值的方法，包括Forbidden(禁止存在缺失)，Min(用最小值补)，Max(用最大值补)。默认Min。
* calc\_feature\_importance (bool): 是否计算特征重要性。默认True。
* fold\_permutation\_block\_size (int): 在数据随机排列前分块，值越小越慢。默认1。
* ignored\_features (list): 忽略数据集中的某些特征。默认None。
* use\_best\_model (bool): 设置此参数时，需要提供测试数据，树的个数通过训练参数和优化loss function获得。默认False。
* **loss\_function (string/ object)**: 支持的有RMSE, Logloss, MAE, CrossEntropy, Quantile, LogLinQuantile, Multiclass, MultiClassOneVsAll, MAPE, Poisson。默认Logloss。
* custom\_loss (object): 训练过程中损失函数的值。默认None。
* eval\_metric (string): 用于过拟合检验（设置True）和最佳模型选择（设置True）的loss function，用于优化。
* **iterations (int)**: 最大树数。默认500。
* border (float): 用于二分类／使用Logloss function中，大于border认为是正样本。默认0.5。
* gradient\_iterations (int): 梯度下降的步数。默认1。
* **depth (int)**: 树深，最大16，建议在1到10之间。默认6。
* **learning\_rate (float)**: 学习率。默认0.03。
* rsm (float [0; 1]): 随机子空间（Random subspace method）。默认1。
* **partition\_random\_seed (int)**: 随机种子。默认None，每次训练随机选择。
* leaf\_estimation\_method (string): 计算叶子值的方法，Newton/ Gradient。默认Gradient。
* **l2\_leaf\_reg (int)**: l2正则参数。默认3
* has\_time (bool): 在将categorical features转化为numerical features和选择树结构时，顺序选择输入数据。默认False（随机）。
* priors (string): 训练过程中指定先验。默认None。
* feature\_priors (list): 在将categorical features转化为numerical features时，指定先验。
* name (string): 在可视化工具中的实验名称。默认experiment。
* fold\_len\_multiplier (float): folds长度系数。设置大于1的参数，在参数较小时获得最佳结果。默认2。
* approx\_on\_full\_history (bool): 计算近似值，False：使用1／fold\_len\_multiplier计算；True：使用fold中前面所有行计算。默认False。
* class\_weights (list): 类别的权重。默认None。
* classes\_count (int): 类别label的上限。默认：类别label最大值＋1。
* one\_hot\_max\_size (bool): 如果feature包含的不同值的数目超过了指定值，将feature转化为float。默认False
* random\_strength (float): 分数标准差乘数。默认1。
* bagging\_temperature (float): 贝叶斯套袋控制强度，区间[0, 1]。默认1。

**Overfitting detection settings**   
- od\_type (string): 过拟合检测类型：IncToDec/ Iter。默认IncToDec。   
- od\_pval (float): 使用IncToDec时的阈值，值越大越早检测出过拟合。默认0（不使用过拟合检测）。   
- od\_wait (int): 在最小化损失函数后的迭代次数。使用InctoDec时，表示当达到阈值后，忽略过拟合检测，继续训练。使用Iter时，表示达到指定次数后，停止训练。默认20。

**CTR settings**   
- ctr\_description (string): categorical features的二值化设置。默认None。包括CTR类型（Borders, Buckets, BinarizedTargetMeanValue，Counter），边界数（只对回归，范围1－255，默认1），二值化类型（只对回归，Median, Uniform, UniformAndQuantiles, MaxSumLog, MinEntropy, GreedyLogSum，默认MinEntropy）。默认None。   
- counter\_calc\_method (string): 计算点击率类型的方法，PrefixTest考虑测试集中当前对象，FullTest考虑测试集中所有对象，SkipTest不考虑测试集中的对象，Full考虑训练和测试集中的全部对象。默认None（PrefixTest）。   
- ctr\_border\_count (int): categorical features的分割数，范围1－255。默认16。   
- max\_ctr\_complexity (int): 组合categorical features的最大数目。默认4。   
- ctr\_leaf\_count\_limit (int): categorical features最大叶子数，如果超过设置值则部分叶子被丢弃。叶子按值的频率排序，选择前n个（n为设置值），之后的叶子全丢弃。默认None。   
- store\_all\_simple\_ctr (bool): 忽略不使用的categorical features。与ctr\_leaf\_count\_limit一起使用。默认False。

**Binarization settings**   
- border\_count (int): numerical features的分割数，范围1－255。默认128。   
- feature\_border\_type (string): numerical features的二值化模式，Median, Uniform, UniformAndQuantiles, MaxSumLog, MinEntropy, GreedyLogSum。默认MiniEntropy。

**Performance settings**   
- thread\_count (int): 训练模型时使用的thread，不影响结果。默认None。

**Output settings**   
- verbose (bool): 显示详细信息。默认False。   
- train\_dir (string): 储存训练期间的文件目录。默认当前目录。   
- allow\_writing\_files (bool): 允许在训练期间写analytical和snapshot文件。如果设置为False，snapshot和可视化工具不能使用。默认True。   
- save\_snapshot (bool): 启用snapshot在中断后存储训练进度。默认None。   
- snapshot\_file (string): 存储的文件名称。默认experiment.cbsnapshot。   
- plot (bool): 训练期间输出以下信息：损失函数值，自定损失值，已训练时间，距训练结束时间。在jupyter notebook中可以使用。默认False。