纵向联邦随机森林

B: X Y C: X

损失函数：

IMG_256

主要流程：

1. B方生成一对密钥，对标签进行同态加密。表示对数据进行同态加密。初始化构造树的终止条件。
2. 对B，C方数据样本进行样本ID对齐，并由B方将加密后的标签发送到C方。
3. C方计算每个特征P个数据箱上的统计数据来掩盖原始数据信息，P指定了用于数据混淆的箱子的数量。在C方对按特征值进行排序，代表C方部分特征的第k个特征。计算第k个特征的百分位数，这里的表示的第k百分位数。
4. 经过（3），将C方的第k个特征划分为了P个数据箱。获得一个数据箱的数据样本。
5. 获得这个数据箱的大小，计算同态加密的统计数据。
6. 重复执行（5）-（6）获得每个数据箱的统计数据。
7. 重复执行（4）-（7），计算出C方数据同态加密的统计数据。
8. B方获得C方传输的加密统计数据，初始化损失阈值。
9. B方解密C方传输过来的赋值给，B方部分包含的标签为，其中k表示各方所包含数据的特征个数，v表示某个特征所包含的特征值个数。
10. B方在当前节点实例空间利用函数计算损失值。
11. 重复执行（9）-（10），根据每个特征和每个特征值来计算损失。选择使损失值最小的特征和特征值并更新，，。
12. 如果当前的大于阈值，则根据，进行分裂。
13. 判断构造停止条件，如果满足，则树构造完成。如果不满足则根据当前节点实例空间和被选择的特征，特征值分裂成左节点空间和右节点空间。
14. B方对此次分裂进行记录并通知C方。
15. 对左，右节点空间重复执行（3）-（15），直到在B，C方完成一棵树的构造。