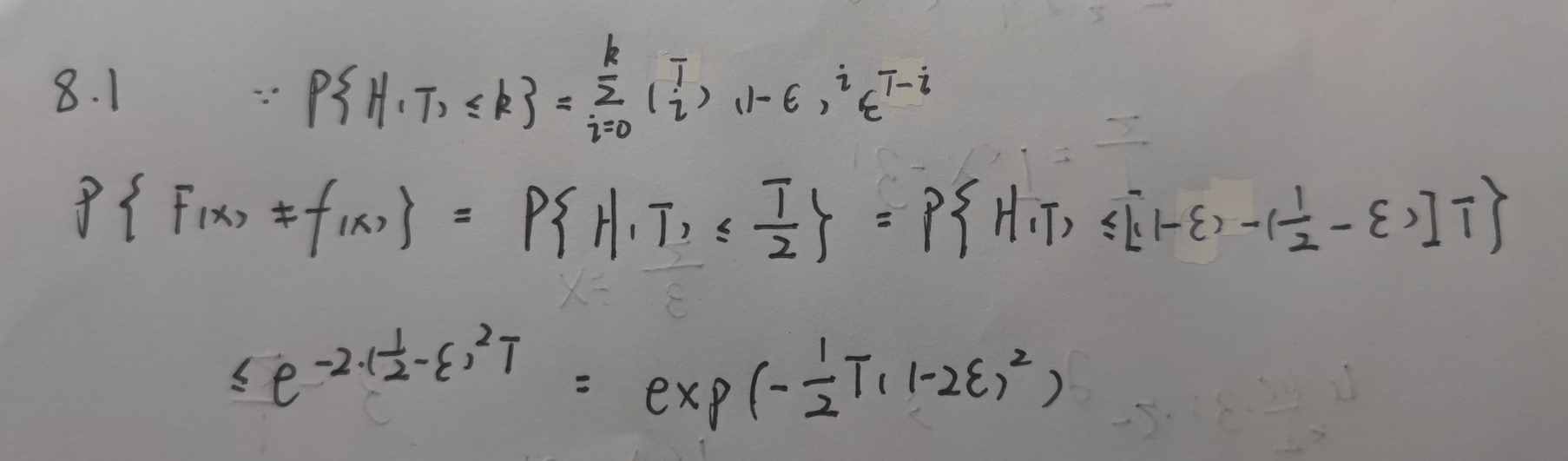
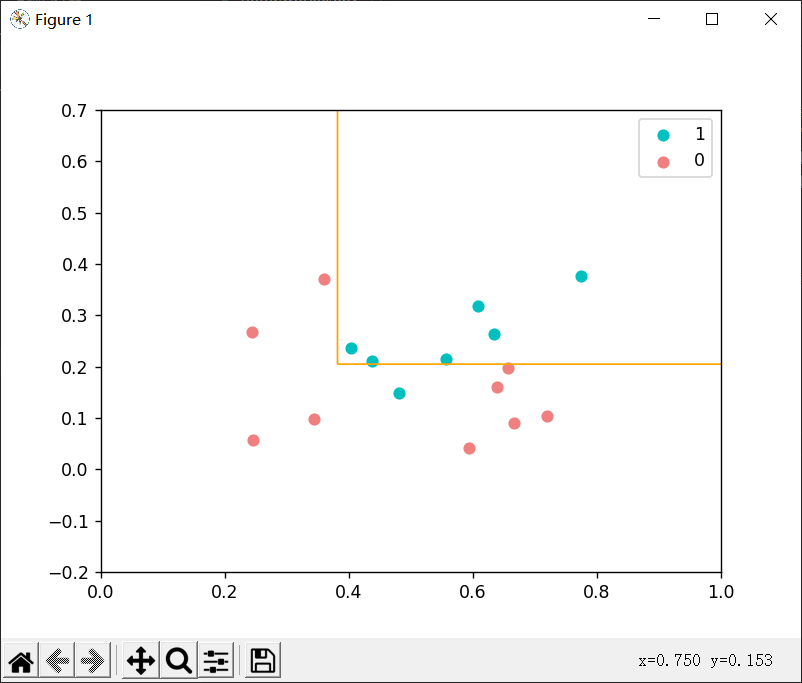
**8.1**

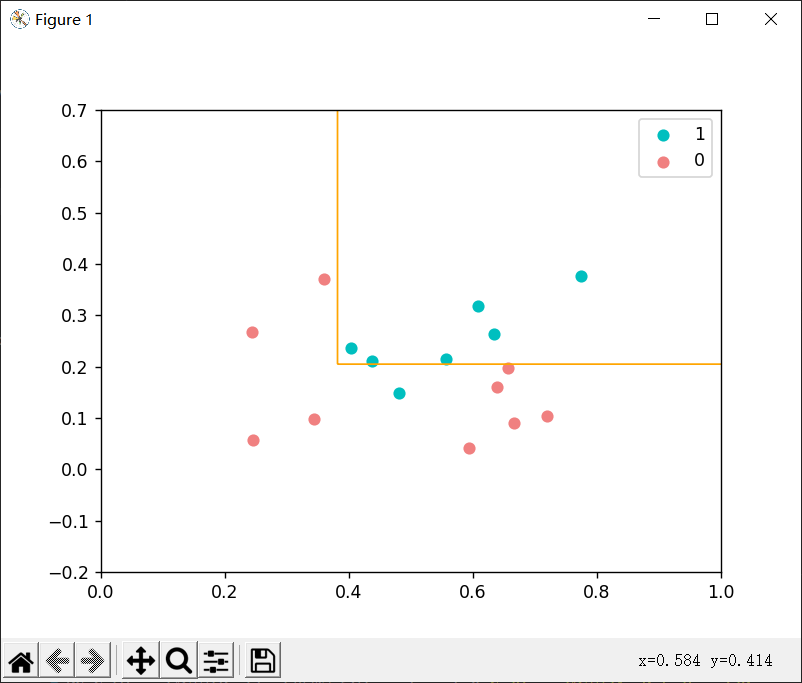


**8.3**

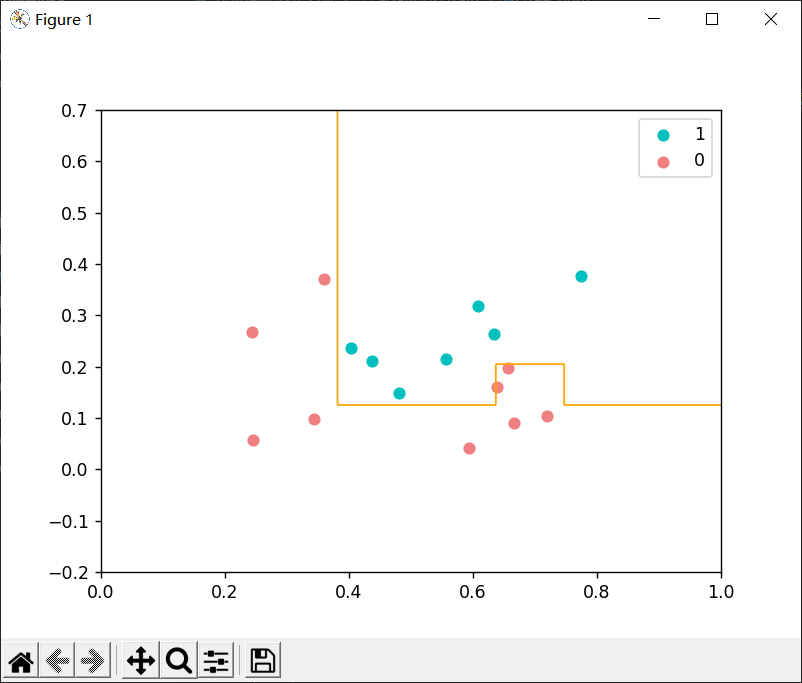
**训练结果如下图所示：**



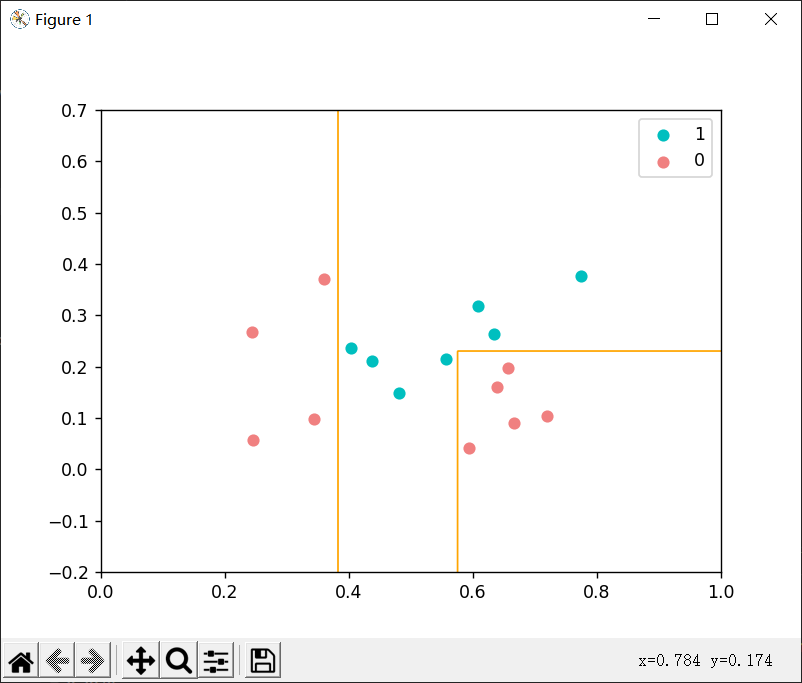
**图1 基学习器数=1**



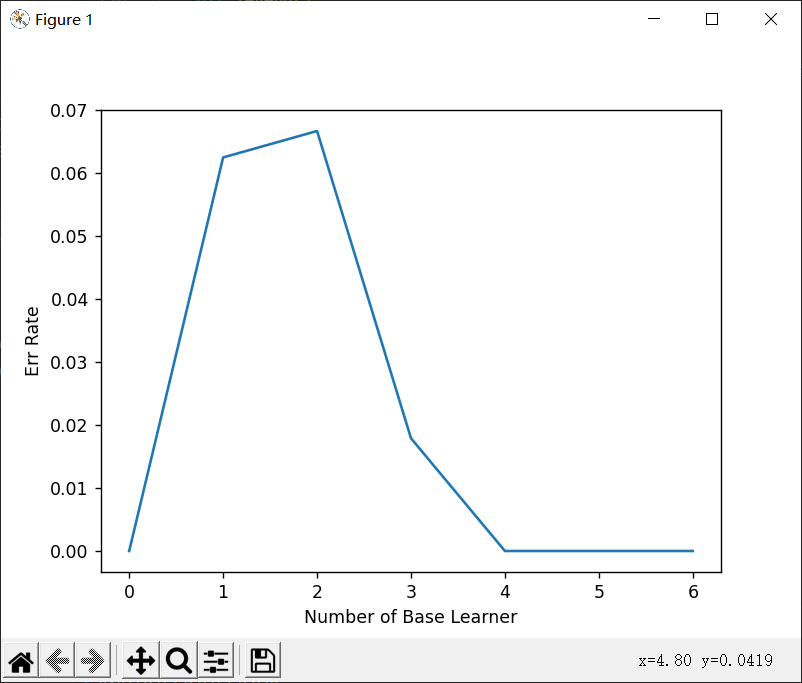
**图2 基学习器数=2**



**图3 基学习器数=3**



**图4 基学习器数=4**



**图5 基学习器数量与分类错误率的关系图**

**可以看出，分类错误率与基学习器数量成反比。当基学习器的数量达到4个时，分类的错误率即可到0。与图8.4相比，该实现可以使用更少的基学习器达到更好的分类效果。**

**8.7**

**1)随机森林采用了一定的随机性，对选择的特征进行随机抽样，这样可以降低过拟合的风险，提高模型泛化能力，并且减少不同决策树之间的相似性。**

**2)随机森林中每个决策树的节点在分裂时都是随机选择特征进行分裂，使多棵树可以并行地构建。**

**3)随机森林能够提前使用随机样本特征，在计算信息熵、基尼指数等时，可以直接判断出最佳分割点，节省计算时间。**