



实验五

集成学习与特征选择



实验五 集成学习与特征选择

- 实验目的
 - 使用集成学习系统练习，体会集成学习系统对分类性能的影响。
 - 学习**scikit-learn**模块的用法，练习使用模块中的基本分类器解决问题。
 - 了解对有监督学习，特征选择对分类结果的影响。



实验数据

- 数据集为Microarray数据，具有高维度小样本的特点。
- 所有数据需转置，转换成样本*特征的方式。
- 转置之后，第一列为类别标签。



实验内容

- 使用多种不同分类器，比较在基分类器参数设定一致的情况下，使用不同基分类器，不同集成学习方法，获得的结果的不同，以及与单个分类器获得结果的区别，并进行分析；
- 请用**soft**、**hard**方法分别集成多个不同的基本分类器，比较结果的区别；
- 通过结合特征选择方法，分析特征选择对实验结果的影响。