

树集成的代码实现

Python、GBDT、模型联结

Python

目录

ONE GBDT代码实现

learning rate

TWO GBDT的模型联结

GBDT + LR

GBDT代码实现

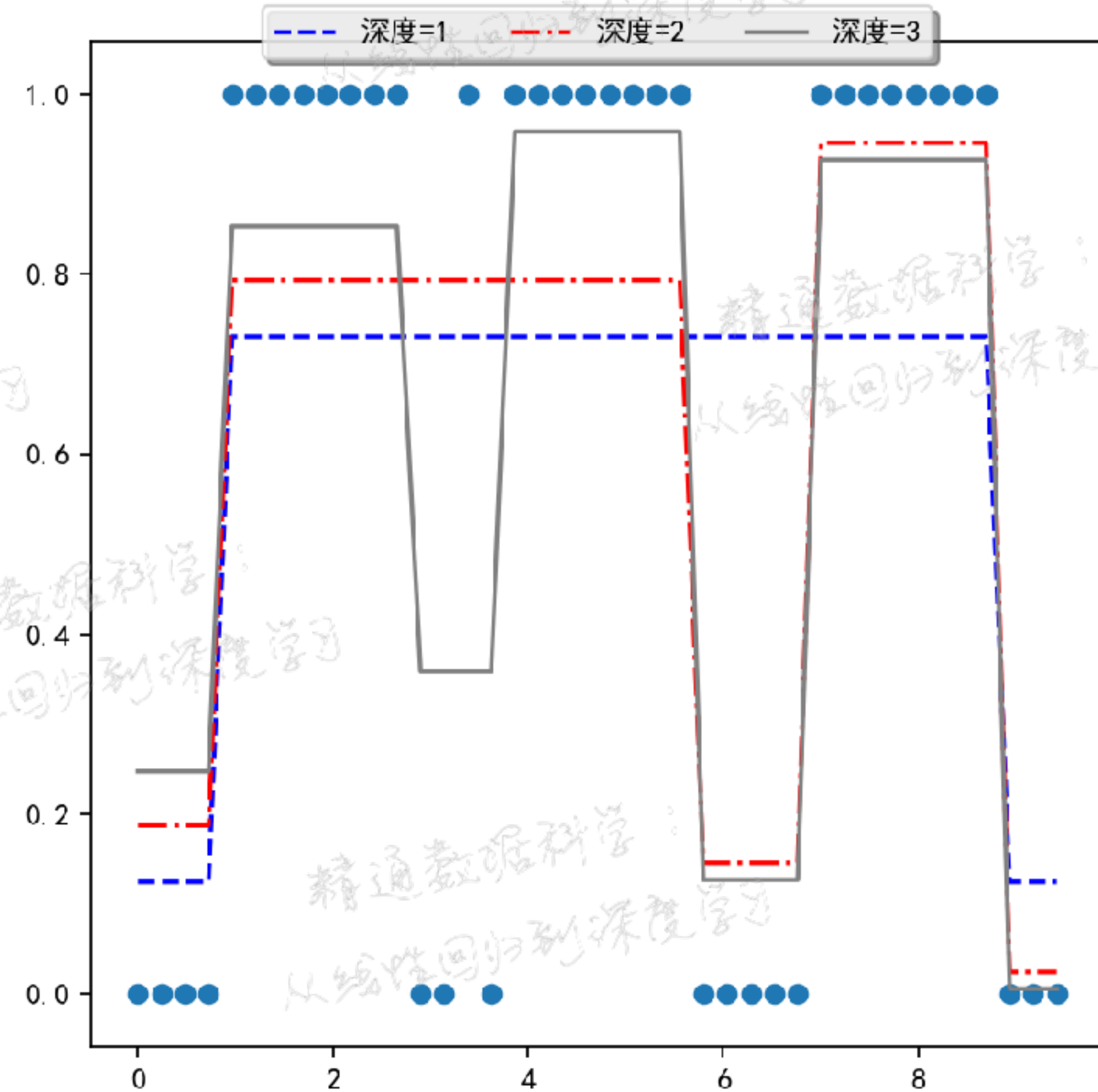
learning rate

scikit learn在GBDT模型的工程实现上，
引入了一个模型超参数：

- learning rate，由这个参数来调节模型
的更新速率：

$$F_k(\mathbf{X}_i) = F_{k-1}(\mathbf{X}_i) + v\gamma_k h_k$$

learning rate



目录

ONE

GBDT代码实现

learning rate

TWO

GBDT的模型联结

GBDT + LR

GBDT的模型联结

GBDT + LR

将数据分为训练集和测试集

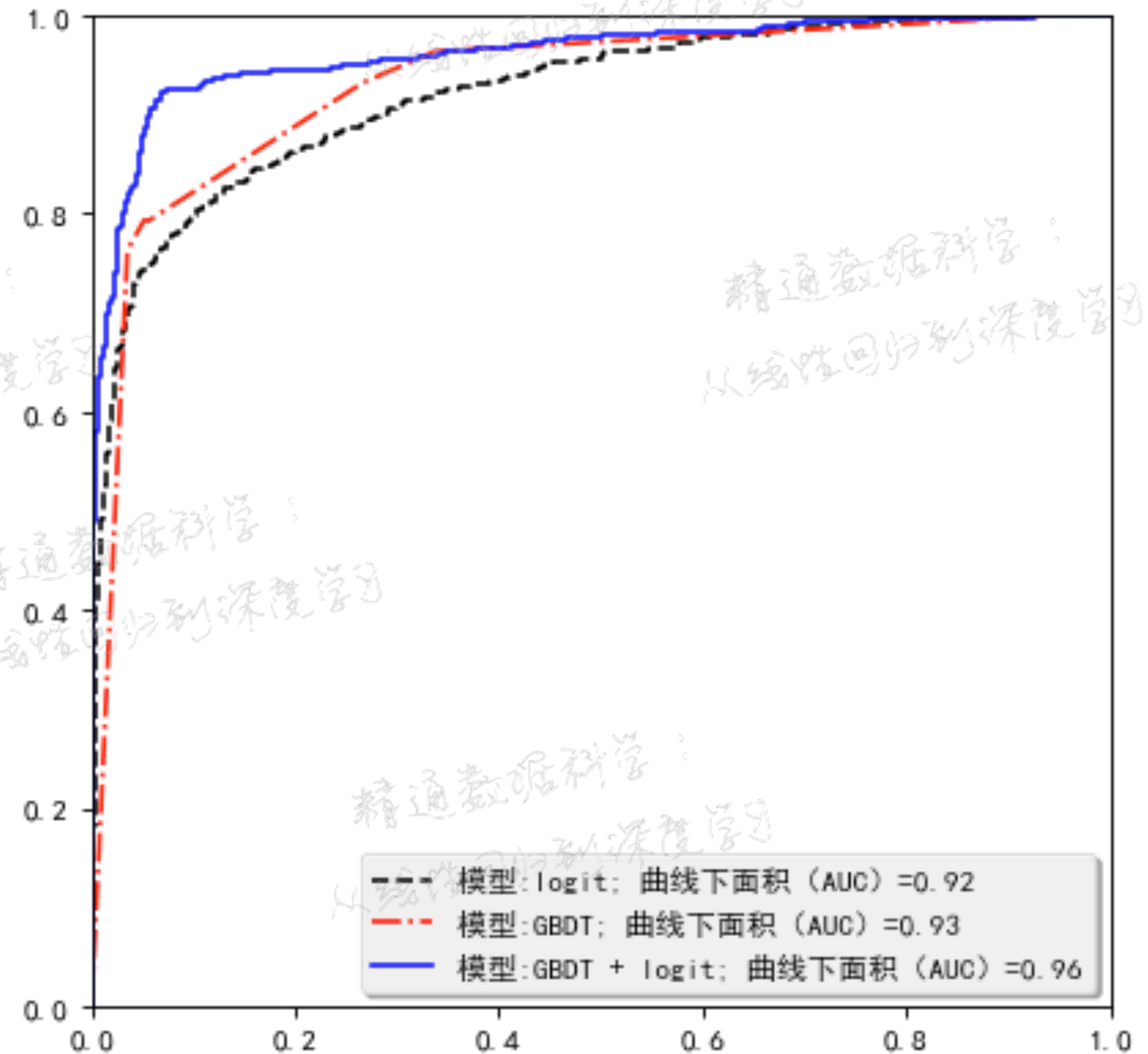
将训练集进一步划分为GBDT训练集和逻辑回归训练集

使用训练好的GBDT对数据进行转换, 得到新特征

使用新特征和旧特征一起搭建逻辑回归模型

注意虚拟变量陷阱

3种模型效果比较



THANK YOU

精通数据挖掘科学：
从线性回归到深度学习