

机器学习

关注者

10

被浏览

1,209

为什么 Xgboost中树的叶节点得分可以用来衡量树的复杂度？

为什么 Xgboost中树的叶节点得分可以用来衡量树的复杂度？它不只是树的输出结果吗？求助

已关注

写回答

邀请回答

添加评论

分享

举报

...

你可以邀请下面用户，快速获得回答

搜索你想邀请的人

**委托刘看山邀请**

开启后为你智能邀请潜在的回答者

开启邀请

**塞罕宝贝**

在「机器学习」话题下有 7 个回答

邀请回答

**郭晓辉**

最近回答过该领域问题

邀请回答

**Tang AI**

在「机器学习」话题下有 32 个回答

邀请回答

**polarlion**

在「机器学习」话题下获得过 26 个赞

邀请回答

更多推荐结果

2 个回答

默认排序

**风间**

follow your heart

4 人赞同了该回答

$$\Omega(f) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T w_j^2$$

xgb模型复杂度两部分构成：叶子节点数目，叶子越多，空间划分越细，局部描述越准确，模型越复杂；叶子节点权重l2norm，l2norm越大，表示随机选择一个点，从一个叶子节点区域移动到相邻叶子节点区域时，目标函数（loss）值波动越大，大部分情况下，相邻区间目标函数不应该相差很大，想象一下一个连续回归问题

参考：[Definition of model complexity in XGBoost](#)

编辑于 2017-11-10

赞同 4

4 条评论

分享

收藏

喜欢

...

**仰望sky**

机械自动化 -> 控制工程 -> ECE

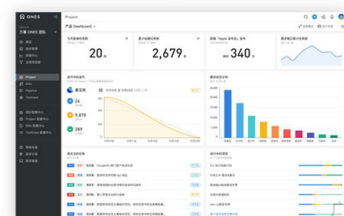
2 人赞同了该回答

个人觉得，叶节点得分并不能衡量树的复杂度吧，至少不是一个好的衡量标准。

在XGBoost: A Scalable Tree Boosting System这篇论文中，增加一个新的booster (tree)的优化过程，利用了泰勒二阶展开。"泰勒展近似相等"本身就只是局部区域有效，所以如果 $f_t(x_i)$ 也就

ONES

追踪项目进度，量化团队表现



相关问题

adaboost算法中训练弱分类器应当使用什么算法？ 6 个回答

LR,gbdt,libfm这三种模型分别适合处理什么类型的特征,为了取得较好效果他们对特征有何要求？ 10 个回答

xgboost本质上是树模型，能进行线性回归拟合么，用xgboost拟合出来的只能是曲线不能出直线？ 9 个回答

神经网络能否代替决策树算法？ 32 个回答

机器学习算法中 GBDT 和 XGBOOST 的区别有哪些？ 36 个回答

相关推荐

**机器学习实践指南**

阿图尔·特里帕蒂

14 人读过

阅读

**机器学习实践指南：案例应用解析**

16 人读过

阅读

**机器学习案例实战**

赵卫东

20 人读过

阅读

teambition

轻量级项目管理工具

10人以下团队免费使用



是 $w_q(x_i)$ 过大的话，这个近似相等就不存在了。这也是我们一般的shrinkage(也叫eta, learning rate)取得很小的原因吧，毕竟Boosting的整体过程是在模仿gradient/steepest descent这一数值优化方法。

在XGBoost调整时，max_depth, min_child_weight and gamma就能够比较好的控制树复杂度了。那个针对叶子节点得分的L2/L1 regularization基本不用调节。只要shrinkage够小，叶子节点得分多少不是问题。

可参考saedsayad.com/docs/xgbo...这个课件。

编辑于 2017-11-30

▲ 赞同 2 ▼ 收起评论 分享 ★ 收藏 ♥ 喜欢 ...

4 条评论

⇌ 切换为时间排序



dalong chen (提问者)

2 年前

嗯嗯，从泰勒近似的角度解释结点输出值需要尽可能小是很合理的。但是你提到的shrinkage 在xgboost中对应的是哪个参数啊？以我的理解正好对应结点输出，不知道对不对

👍 赞



仰望sky (作者) 回复 dalong chen (提问者)

2 年前

shrinkage是GBM中术语。在XGBoost中就是eta或者learning_rate。在不同接口(R, Python, sklearn等)中，可能称呼不同。

👍 赞



dalong chen (提问者) 回复 仰望sky (作者)

2 年前

嗯，非常感谢，不过还是有疑问，我对learning_rate的理解还是停留在比如在梯度下降中更新参数的步长，但是在xgboost中没有明显的通过梯度下降更新参数这样的操作，所以learning_rate在xgboost还是没搞清楚指的是什么。

👍 赞



Honfung.Wong 回复 dalong chen (提问者)

1 个月前

xgb 的 eta 体现在新一代的树，在放入加法模型时，额外给了一个小于 1 的权重，目的是降低当前树的影响，给后续更多学习空间

👍 赞

写下你的评论...



✍ 写回答

1 个回答被折叠 (为什么?)

刘看山 · 知乎指南 · 知乎协议 · 知乎隐私保护



应用 · 工作 · 申请开通知乎机构号

侵权举报 · 网上有害信息举报专区

京 ICP 证 110745 号

京 ICP 备 13052560 号 - 1

京公网安备 11010802010035 号

互联网药品信息服务资格证书

(京) - 非经营性 - 2017 - 0067

违法和不良信息举报: 010-82716601

儿童色情信息举报专区

证照中心

联系我们 © 2020 知乎

