

蒋豆芽

机器学习——KNN 已购

来自【机器学习面试题汇总与解析（蒋豆芽面试题总结）】 | 34 浏览 | 0 回复 | 2021-05-23



蒋豆芽

[+关注](#)

机器学习面试题汇总与解析——KNN

1. KNN介绍一下☆☆☆☆
2. KNN优缺点☆☆☆☆
3. KNN的K值怎么选☆☆☆☆
4. KNN数据需要归一化吗？☆☆☆☆
5. KNN三要素说一下☆☆☆☆
6. 欧式距离与曼哈顿距离区别☆☆☆☆
7. knn的k设置的过大会什么问题☆☆☆☆

- =====
- 本专栏适合于Python已经入门的学生或人士，有一定的编程基础。
 - 本专栏适合于算法工程师、机器学习、图像处理求职的学生或人士。
 - 本专栏针对面试题答案进行了优化，尽量做到好记、言简意赅。这才是一份面试题总结的正确打开方式。这样才方便背诵
 - 如专栏内容有错漏，欢迎在评论区指出或私聊我更改，一起学习，共同进步。
 - 相信大家都有着高尚的灵魂，请尊重我的知识产权，未经允许严禁各类机构和个人转载、传阅本专栏的内容。
- =====

关于机器学习算法书籍，我强烈推荐一本《百面机器学习算法工程师带你面试》，这个就很类似面经，还有讲解，写得比较好。私聊我进群。

关于深度学习算法书籍，我强烈推荐一本《解析神经网络——深度学习实践手册》，简称CNN book，通俗易懂。私聊我进群。

参考资料

B站机器学习视频：<https://space.bilibili.com/10781175/channel/detail?cid=133301>

蒋豆芽

KNN实战: <https://www.cnblogs.com/listenfwind/p/10685192.html>

读者可以把参考文章看看

个人理解

KNN是较为简单的机器学习方法了。最关键的就是理解“距离”的概念。

KNN的优缺点

优点:

1. 简单易用，相比其他算法，KNN算是比较简洁明了的算法。
2. 模型训练快
3. 预测效果好。
4. 对异常值不敏感

缺点:

1. 对内存要求较高，因为该算法存储了所有训练数据
2. 预测阶段可能很慢
3. 对不相关的功能和数据规模敏感
4. 对数据纲量敏感，所以数据要先归一化

1. KNN介绍一下☆☆☆☆

参考回答

KNN的全称是K Nearest Neighbors，意思是K个最近的邻居，KNN的原理就是当预测一个新的值x的时候，根据它距离最近的K个点是什么类别来判断x属于哪个类别。通过**方差**表示“距离”

答案解析

无。

类似的问题还有：

2. KNN优缺点☆☆☆☆

参考回答

KNN的优缺点

优点:

☰ 蒋豆芽

3. 预测效果好。
4. 对异常值不敏感

缺点：

1. 对内存要求较高，因为该算法存储了所有训练数据
2. 预测阶段可能很慢
3. 对不相关的功能和数据规模敏感
4. 对数据纲量敏感，所以数据要先归一化

答案解析

无。

3. KNN的K值怎么选☆☆☆☆

参考回答

通过不断验证K不同的取值来确定。

答案解析

无。

4. KNN数据需要归一化吗？☆☆☆☆

参考回答

KNN对数据**纲量**敏感，所以数据要先**归一化**。因为KNN使用的方差来反映“距离”，纲量对方差计算影响较大。

答案解析

无。

5. KNN三要素说一下☆☆☆☆

参考回答

1. k值的选取
2. 距离度量的方式。一般为欧式距离

答案解析

无。

6. 欧式距离与曼哈顿距离区别 ☆ ☆ ☆ ☆

参考回答

欧式距离：平方差求和再开方

曼哈顿距离：坐标差的绝对值求和

一般用欧式距离而非曼哈顿距离的**原因**：**欧式距离**可适用于不同空间，表示不同空间点之间的距离；**曼哈顿距离**则只计算水平或垂直距离，有维度的限制

答案解析

无。

7. knn的k设置的过大会有什么问题 ☆ ☆ ☆ ☆

参考回答

如果选择较大的K值，就相当于用较大领域中的训练实例进行预测，其优点是可以减少学习的估计误差，但缺点是学习的近似误差会增大。

我们考虑一种极端的情况，当k和整个样本数量一样的，KNN的分类结果总是取决于样本类别数量最多的那一类。这时模型的误差最大化。

答案解析

无。

[资源分享](#)[python](#)[机器学习](#)[算法工程师](#)[春秋招](#)[面试题](#)[软件开发](#)[面经](#)[举报](#)

收藏



赞

相关专栏



机器学习面试题汇总与解析（蒋豆芽面试题总结）

27篇文章 | 90订阅

已订阅

☰ 蒋豆芽

0条评论

0人喜欢



没有回复

请留下你的观点吧~

发布

 牛客博客，记录你的成长

[关于博客](#) | [意见反馈](#) | [免责声明](#) | [牛客网首页](#)