

三轮面试中问到的一些问题：

一、机器学习部分

1. 机器学习的整个流程是怎样的？
2. 如何处理样本中的缺失数据？
3. 做过的项目中用到了哪写机器学习算法？有没有遇到过什么问题？是怎么解决的？
4. 决策树有什么问题？怎么解决
5. 说一说随机森林和逻辑回归在分类问题上使用场景有什么不同？
6. SVM 算法推导
核函数是什么？有哪些？
函数间隔和几何间隔有什么区别？
加大数据量一定会提升 SVM 的准确率吗？为什么？
7. SVM 和 LR 的区别？
8. 逻辑回归为什么选用 sigmoid 函数？多分类怎么做？
9. 逻辑回归可以解决非线性分类问题吗？
10. 详细解释 L1 和 L2 正则化
11. 了解 xgboost 吗？和 gbdn 的区别？
12. 平时看过哪些机器学习方面的书籍？是怎么学习机器学习的？

二、JAVA 部分

1. Java 基础

HashMap 的实现原理，为什么不是线程安全的，并发情况下会有什么问题？

ConcurrentHashMap 怎么实现线程安全的

类的加载过程？

类加载器有哪些？双亲委派模型？有什么作用？

2. JVM

Java 内存管理模型

垃圾回收算法：CMS，G1

如何判断一个对象是否要被回收

3. Spring

Spring Bean 加载，解决循环引用

Spring AOP 原理

FactoryBean 和 BeanFactory 区别

4. 数据库

事务的隔离级别？平时用的是哪种？

MySQL 常用存储引擎，InnoDB 数据是怎么存储的

Left Join 是怎么执行的

聚簇索引，二级索引，联合索引

怎么判断一个查询走没走索引，like 走索引吗？

Hash 可以做索引吗？为什么 InnoDB 不使用 Hash 索引？

如何利用索引提升查询速率(任何优化一个慢查询)

5. MyBatis

MyBatis 执行一个 Select 查询的流程？

有哪些 Executor

插件原理

二级缓存机制？

6.多线程并发

线程生成方式有哪些? Callable 返回值?

ThreadLocal 原理

volatile 关键字原理

线程池参数

拒绝策略有哪些

同步队列哪几种? 区别?

线程通信方式有哪些

Synchronized 和 ReentrantLock 区别

编程实现两个线程循环打印 ABABAB

三、数据结构算法

双向链表如何判断有交叉? 如何找到交叉点?