拼多多+猫眼+京东云+猎豹移动机器学习面试分享

**拼多多一二面**：讲项目，针对项目提问（没什么特别的）

基础算法： 给定一棵多叉树，每个节点有一个编号，现在要对节点排序，要求对于每个节点，它的父节点排在它后面,如果一棵树有N个节点，那么肯定有N-1条边。输入的数据形式为：Map.Entry<int, int> key: 父节点编号，value: 子节点编号（也就是边的集合）

**猫眼二面**：比较惨的一次面试，自己挖坑了，讲了实验室工作，但是研究方向很久没搞了而且不打算做那块所以也没深入研究，结果就GG了，让别人觉得研究方向都讲不清楚不好印象。

算法题：数组找前K大的元素，在线写有点紧张，就直接写了快排。应该用快排的思想但不实际排序

**京东云二面：**针对项目提问

1.讲了下stacking模型以及做模型融合的一些工作

2.序列模式挖掘抛出两个问题：1.在分布式情况下，如果我想尽可能多的挖掘出多的频繁序列，怎么办？2.在分布式情况下，如果我要求尽可能准确的挖掘序列模式，怎么做？

答：没在分布式环境中做过，举例子如果FPTree 算法可以并行的建树，如果挖掘更多的话，我们可以随着频繁序列的增长而降低阈值，面试官基本认可，说有一个对应公式，方向应该没问题。

第二问当时答的不好，面试官说的意思是对各个机器上计算出来的频繁序列零容忍，即只要出现一个项集不是频繁的，则其他机器包含他的序列模式都应该抛弃。（**有没有大神指点下**）

3.怎样去优化SVM算法模型的？

答：没弄懂面试官的意思，把SVM讲了一遍。

其实应该答SMO算法的，SMO算法是支持向量机的快速算法，不断的将原二次规划问题分解为只有两个变量的二次规划子问题求解，直到所有变量满足KTT条件，这样通过启发式的方法得到原二次规划问题的最优解。（**欢迎大佬补充KKT条件和凸二次规划的理论知识**）

**猎豹移动一面：**讲项目，针对项目提问（没什么特别的）

1.讲下LSTM算法模型

2.讲下CNN算法模型

3.现有底层是tensorflow的keras框架，如果现在有一个tensorflow训练好的模型，keras怎么读取？

答：一脸蒙逼（**还请哪位大神指点一二**）

**总结：拼多多算法岗感觉也会有大量的开发工作，感觉考察机器学习的不太多，编程做的不对，主要难点我觉得输入形式不是树的结构，感觉可能GG了；自我感觉有大厂的实习经验还是有用的，前提是参与的项目得十分清楚，不会被考到。基本算法题能力薄弱**