刷宝科技,互联网金融,,聊得挺愉快.貌似行业很有潜力,风控建模的

2个技术官,分别一个小时

第一个是风控模型技术负责人

,第二个是大数据技术员

简历上写什么,基本全部问到.

手写快速排序?(因为已背下,被两个技术官夸赞算法基础扎实,草泥马飞过,,)

手写SQL筛选(不难的 行筛选 手机号like 135% 这种答案)

对SVM的理解?空间映射的原理?(答复方差最大的特征,递归查出正交的特征,依次排序,自动封装,给定阈值)

PCA降维的原理?(答了同样的,笑了就)

看过周志华的机器学习没有?看过李航的统计学没有?XGboost了解没

SQL优化的方式?(答了查询语句+索引设置两方面,算满意)

为什么有时候SQL索引会失效?(答经常增删改,数据内存的指向地址变化了, 貌似不太对)

缺失值怎么处理(答,如果重要就保留待观察或者拉格朗日插值法,如果不重要就删了)

怎么检测异常值,(回答上四分图,下四分的箱型图,然后被指导还有另外一个常用的实际,)

给你一个手机号表,能获取什么信息(偏向数据清洗,判断哪些缺失例如是不是11位,哪些可能是随便填的)

表里是身份证号码,怎么提取出来年龄,判断注册是乱填的

拓展非常多,给你上课一样,,装老实,

1. tree,hadoop.二叉树,链式索引,内存地址啥的.这种谈谈理解.测试知识面,还问业余看什么书,有什么兴趣爱好.加班强度接受能力.

人脸识别用过什么工具做的答(咱们讲过的啥serachCV,自动逼近系数的那个,随便谈了下),社交网络分析igraph ,jieba,snownlp 情感分析,忒多,都会问到,写什么问什么.,做过什么kaggle比赛啥的.RMF模型(数据挖掘实战的第七章,航空客户价值模型有)

vivo数据分析,目前在上梅林,9月搬到宝安中心壹方城,很有钱的样子.已过技术面,约了27号非技术面,

面试官一个人,是团队负责人.一脸严肃两个小时,全是谈技术.

理论问题:

逻辑斯第回归的算法原理?

线性回归的原理和上面的区别?

数据规约,数据清洗,缺失值,特征处理,降维-模型建立,细致讲某个项目的工作过程.越细越好

爬虫的东西.

怎么确定降维的特征?舍去的原因.

针对项目:

量纲归一化均值化,的作用区别

特征中,时间选年还是月,为什么.

经纬度数据表,能提炼出什么价值信息?

,,,后面还有很多,忘了,,,