

[笔试题目] 腾讯2015年9月基础研究笔试题

原创 Eastmount 2015-10-04 20:58:12 12285 收藏

版权

分类专栏: [面试工作](#) 文章标签: [腾讯](#) [笔试题](#) [自然语言处理](#)



Python+TensorFlow人工智能

该专栏为人工智能入门专栏，采用Python3和TensorFlow实现人工智能相关算法。前期介绍安装流程、基础语法、神经网络、可视化等，中间讲解CNN、RNN、LSTM等代码，后续复现图像处理...



Eastmount

¥9.90

[订阅博主](#)

面试职位

- 应聘职位：基础研究
- 岗位描述：探索和创新前沿互联网技术，将高新技术转化为产品，参与腾讯产品的开发与技术实现
- 岗位要求：2016年毕业，自然语言处理、多媒体、分布式网络、数据分析、智能计算、安全等互联网相关技术具有较深研究和专业积累，具有较好的创新能力和编程开发能力，具备较强的团队合作能力

选择题

- 第1题 求定积分结果 $\int_0^1 1 dx$
答案：1
- 第2题 已知总标准差=25，随着样本增加均值怎么变化。
- 第3题 已知K是第四名，A的年龄不是最大但比D老，D不是第二名，年龄第2小的是第二名，第三名的年龄大于第一名，B比第三名年轻，求A、B、D和K的排名顺序。
答案：D>B>A>K
- 第4题 $y=\sin(x)+\cos(x)$ 的极大值
答案： $\sqrt{2}$
- 第5题 $\max\{Cx \mid Ax \leq b, x \geq 0\}$ 为最优解， $z \geq 0$ ，当C、b被v和v/2乘时一下结果正确的是...
- 第6题 u检验的应用条件是：两样本来源得总体符合正太分布...
- 第7题 设某商品需求满足： $y_i = \beta_0 + \beta_1 * x_i + \varepsilon_i$ ，参数估计量...
- 第8题 齐次线性方程组 $Ax=0$ 存在无穷解，其行列式m行、n列实数矩阵，则满足：
A. $m > n$ B. $m < n$ C. $\text{rank}(A) = m$ D. $\text{rank}(A) = n$
- 第9题 已知 $x_0=(1,1)$ ，求 $f(x) = 4 * x_1 + 6 * x_2 - 2 * x_1^2 - 2 * x_1 * x_2 - 2 * x_2^2$ 极大值。 x_1 坐标为：
A. (1,1) B. (-1,1) C. (-1/2,1) D. (1/2,1)
- 第10题 相同样本量下，重复抽样与不重复抽样平均误差大小关系是：相同、重复抽样大、不重复抽样大选项
- 第11题 n阶矩阵A行列式 $|A| = 1$ ，则 $|A^{-1}| = ?$
- 第12题 一对夫妇两孩子，已知一个是男孩，求两个都是男孩的概率。猜测答案是0.5
- 第13题 $x_1, x_2, x_3 \dots x_n$ 来自 (μ, σ^2) 一个样本， \bar{x} 为均值， S^2 为方差。其中答案包括： $S^2/\sigma^2 \sim \chi(n-1)$
- 第14题 随机变量x概率分布： $x=2 f(x)=0.2$ 、 $x=4 f(x)=0.3$ 、 $x=7 f(x)=0.4$ 、 $x=8 f(x)=0.1$ ，x的标准差是。
- 第15题 一盒装有6个乒乓球，4个位新球，第一次随机取2个，使用后放回盒子；第二次随机取2个，求第二次取全新的概率。
- 第16题 商品销售量、方差分析，excel表分析颜色特征对结果的影响。

- 第17题 k_1 、 k_2 是矩阵A两个不同特征值， a 与 b 是A分别属于 k_1 、 k_2 的特征向量，则 a 、 b 满足，选项包括线性无关、线性相关等。
- 第18题 青蛙掉入21米井里，每天向上爬10米，晚上再掉下7米，问几天爬出。
答案：5天
- 第19题 人工批量盆景虎皮兰，平均高度70cm，标准差5cm，随机搬出100盆，则说法正确的是：选项包括至少75盆以上在70-80cm之间等。
- 第20题 行列式的值是多少：

$$\begin{vmatrix} a^2 & ab & b^2 \\ 2a & a+b & 2b \\ 1 & 1 & 1 \end{vmatrix}$$

- 第21题 数据库系统描述错误的是：
A.DBMS是数据库管理系统简称
B.Hive是开源的数据库系统
C.数据库事务有ACID属性
D.SQL中create属于DML语言

答案：D

分析：终于看到第一道与编程、OS、DB相关的题目了，泪流满面啊！故简单讲解下：DBMS就是数据库管理系统（Database Management System），事务包括ACID性质（原子性Atomicity、一致性Consistency、隔离性Isolation、持久性Durability），Hive是属于Hadoop的数据仓库工具。而create属于DDL（数据库模式定义语言），Insert、delete、update属于DML（数据库操纵语言）、Select属于DQL（数据库查询语言）、grant属于DCL（数据库控制语言）。

- 第22题 甲乙两人比射箭，水平一样。甲射101箭，乙射100箭，求甲射中次数比乙射中次数多的概率。
- 第23题 一个对数组A（n个元数组）排序，代码如下：

```
for i=i to length(A)-1
    j=i
    while j>0 and A[j-1]>A[j]
        swap A[j] and A[j-1]
    j=j-1
```

求代码的时间复杂度： $O(n^2)$

- 第24题 mysql中选课表learn(stu_id int, course_id int)，其中学生id和课程id，则获取每个学生所选课程个数的语句是：
A.select stu_id, count(course_id) from learn;
B.select stu_id, sum(course_id) from learn;
C.select stu_id, count(course_id) from learn group by student_id;
D.select stu_id, sum(course_id) from learn group by student_id;

分析：count计数和sum求和的区别，group by分组

- 第25题 两家工厂A和B，技术相同，合格率一样平均80%，A大规模工厂，B小规模工厂，没有出现故障某家企业合格率仅为70%，问哪家工厂的可能性更大？

问答题

第一题

进行特征选择时，即分析哪些协变量对目标变量有较大影响时，有哪些常用的方法？你本人更看好其中哪些？为什么？

第二题

网络上流传一个故事，关于数据的神奇性：某商场向一位未婚少女邮寄一些用于保胎的商品目标和折扣券，其父亲发火，后证实该少女确实怀

孕，解释一下如何做到的？

个人从一下几个方面答的，可能存在错误或不足之处：

- 1.搜索引擎与推荐系统及算法
- 2.数据挖掘相关的关联规则挖掘算法，经典尿布与啤酒
- 3.知识图谱 问答系统实现理解人类语言（毕设相关）

第三题

QQ空间相册替换墙换一换功能。把聚类好的人脸进行圈人标注，系统根据用户历史标注记录，分析每张照片最可能身份，用户进行确认和否认。每次只从聚类人脸中抽取5张进行展示。假设利用信息是每个人脸的相似度矩阵 G ，而且第 $k+1$ 次抽样并不知道第 k 次抽出的样本，可用人脸数量 N 始终大于50张。

- (1) 如何设计抽样算法，抽样碰撞率尽可能低；
- (2) 如果抽样完成后，抽出5张人脸从中删除，如何设计抽样算法，使得相邻两次间抽样碰撞率尽可能低？

总结

面试了阿里和腾讯的NLP和数据挖掘相关的算法方向，个人感觉腾讯的题目偏重数学统计概率，其中选择题就3道与程序和数据库相关，而且比较简单，更多的是数学知识；而三个大题感觉还是不错的，很贴近生活和腾讯的实际产品，让我也学到很多知识。但是阿里的题目都至少偏向算法点，自己需要学习的东西还是非常多，也认识到了自己很多不足之处！接下来继续努力吧。

第一次使用Markdown编辑器，发现它latex编辑数学公式和线下博客优势确实明显，但同时我也喜欢自己的CSDN编码风格：微软雅黑18px字体。推荐：[Markdown输入数学公式](#)

当时一边做题一边抄题很辛苦的，但是感触也颇多，不敢发表出来。最后希望该博客内容对你有所帮助，题目所有权归腾讯公司所有，我只是想分享给大家学习，还望贵公司海涵~

(By: Eastmount 2015-9-7 下午2点 <http://blog.csdn.net/eastmount/>)