**币尔大数据笔试题**

1. 设函数在x0处可导， ，则当 时，必有（ ）
2. dy是h的同价无穷小
3.  是h的同阶无穷小
4. dy是h的高阶无穷小
5. 是h的高阶无穷小
6. 将3粒黄豆随机放入4个杯子，则杯中盛黄豆最多为一粒的概率为（ ）
7.  B.  C.  D.
8. 设 ，则等于 （ ）
9.  B.  C. D. 
10. 某学校五个绿化小组一天植树的棵树如下: 10, 10, 12, x, 8, 如果这组数据的平均数与众数相等，那么这组数据的中位数是（ ）
11. 8 B. 9 C. 10 D. 12
12. 从邻接矩阵 可见，该图中含有顶点的个数为（ ）
13. 5 B. 9 C. 4 D. 3
14. 对于JVM内存配置参数:

-Xmx10240m -Xms10240m -Xmm5120m -XXSurvivorRatio=3

其最小内存值和Survivor区总大小分别是（ ）

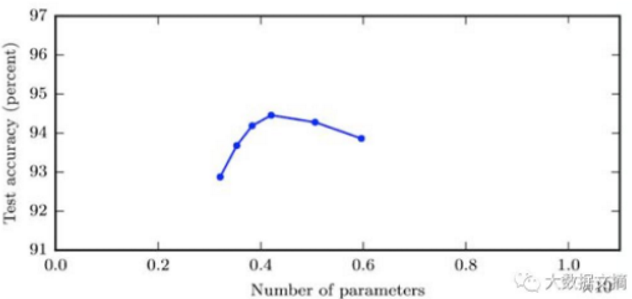
1. 5120m，1024m
2. 5120m，2048m
3. 10240m，1024m
4. 10240m，2048m
5. 下面属于面向对象的特征的是（ ）
6. 重载 B. 重写 C. 封装 D. 解耦
7. Python不支持的数据类型有（ ）
8. char B. int C. float D. list
9. 下列哪个语句在Python中是非法的（ ）
10.  B.  C.  D. 
11. 关于死锁的说法不正确的是（ ）
12. 竞争可剥夺资源会产生死锁
13. 竞争临时资源会产生死锁
14. 在发生死锁时，必然存在一个进程—资源的环形链
15. 如果进程在一次性申请其所需的全部资源成功后才运行，就不会发生死锁。
16. 在数据结构的讨论中，把数据结构从逻辑上分为（ ）
17. 内部结构与外部结构
18. 静态结构与动态结构
19. 线性结构与非线性结构
20. 紧凑结构与非紧凑结构
21. 下面程序段的时间复杂度为（ ）

for (int i=0; i<m; i++)

for(int j=0; j<n; j++)

a[i]a[j] = i\*j;

1.  B.  C.  D. 
2. 已知栈的最大容量为4，若进栈序列为1,2,3,4,5,6，且进栈和出栈可以穿插进行，则可能出现的出栈序列为（ ）
3. 5,4,3,2,1 B. 2,3,5,6,1,4 C. 3,2,5,4,1,6 D. 1,4,6,5,2,3
4. 评估模型之后，得出模型存在偏差，下列哪种方法可能解决这一问题：
5. 减少模型中的特征
6. 增加模型中的特征
7. 增加更多的数据
8. 减少一些数据
9. 训练好的三层卷积神经网络的精确度（Accuracy）vs参数数量（比如特征核的数量）的图如下：



从图中趋势可见，如果增加神经网络的宽度，精度会先增大再减小。这一现象出现最可能的原因是：

1. 当核数量增加，其相关性增加，导致过拟合
2. 当核数量增加，神经网络的预测功效降低
3. 即便核数量（number of kernels）增加，只有一部分核被用于预测
4. 以上都不可能
5. 假设有数据集(x,t)={(x1,t1),(x2,t2),…(xi,ti)}和如下的神经网络模型：



其中： 

 为该神经网络的第i层权重矩阵

 为第i层偏移量

 为第i层输出

整个神经网络的输出为，损失函数为 ，

求 ， ，

1. 给定一个字符串，其中的字符包含英文字母和三种括号：花括号{ }、中括号[ ]、圆括号( )，即它仅由 “( ) [ ] { }” 这六个字符组成。用python语言编程判断该字符串是否有效，即字符串中括号是否匹配。括号匹配要求括号必须以正确的顺序配对，如 “{ [ab ]c ( ) }” 或 “n[ ( { f}[ ] ) ]” 等为正确的格式，而*“*[ ( aab] )”或*“*{ cdd[ ( d) }” 或*“*( {ef } ] )” 均为不正确的格式。
2. Given a string, find the **longest substring** without repeating characters.

**Examples:**

Given "abcabcbb", the answer is "abc".

Given "bbbbb", the answer is "b".

Given "pwwkew", the answer is "wke", with the length of 3. Note that the answer must be a **substring**, "pwke" is a *subsequence* and not a substring.

Use python to solve it.

1. 杭州有多少辆汽车 ? 如不能给出准确数字，结合自己所知的相关数据推测一个范围。

1、聊天机器人的发展历程，有哪些技术手段以及各自的优缺点

2、制作一款陪聊机器人，如何解决如下问题：

a、能回通顺的句子

b、有恒定的背景信息

c、能增量学习

d、同一个问题能多样回复

e、有上下文逻辑衔接的回复

f、记住不同聊天者的信息，面对特定的聊天者，能结合他的信息回复