

百度机器学习笔试题

减肥的小猴子

Jul,12,2017

1 2004年

一简答题（30分每题10分）

1. 动态链接库和静态链接库优缺点

静态链接库的优点

- (1) 代码装载速度快，执行速度略比动态链接库快；
- (2) 只需保证在开发者的计算机中有正确的.LIB文件，在以二进制形式发布程序时不需考虑在用户的计算机上.LIB文件是否存在及版本问题，可避免DLL地域等问题。

动态链接库的优点

- (1) 更加节省内存并减少页面交换；
- (2) DLL文件与EXE文件独立，只要输出接口不变（即名称、参数、返回值类型和调用约定不变），更换DLL文件不会对EXE文件造成任何影响，因而极大地提高了可维护性和可扩展性；
- (3) 不同编程语言编写的程序只要按照函数调用约定就可以调用同一个DLL函数；
- (4) 适用于大规模的软件开发，使开发过程独立、耦合度小，便于不同开发者和开发组织之间进行开发和测试。

不足之处

- (1) 使用静态链接生成的可执行文件体积较大，包含相同的公共代码，造成浪费；
- (2) 使用动态链接库的应用程序不是自完备的，它依赖的DLL模块也要存在，如果使用载入时动态链接，程序启动时发现DLL不存在，系统将终止程序并给出错误信息。而使用运行时动态链接，系统不会终

止，但由于DLL中的导出函数不可用，程序会加载失败；速度比静态链接慢。当某个模块更新后，如果新模块与旧的模块不兼容，那么那些需要该模块才能运行的软件，统统撕掉。这在早期Windows中很常见。

2. 轮询调度和抢占式调度的区别

轮询调度算法的原理是每一次把来自用户的请求轮流分配给内部中的服务器，从1开始，直到N(内部服务器个数)，然后重新开始循环。

抢占式任务调度允许调度程序根据某种原则去暂停某个正在执行的进程，将已分配给该进程的处理机重新分配给另一进程。抢占方式的优点是，可以防止一个长进程长时间占用处理机，能为大多数进程提供更公平的服务，特别是能满足对响应时间有着较严格要求的实时任务的需求。

因为抢占式调度可能会暂停一些进程，需要记录进程的运行状态，较为复杂。轮询式只需要轮流分配资源，调度简单。

3. N折交叉验证的思想

将原始数据分成N组(一般是均分),将每个子集数据分别做一次验证集,其余的N-1组子集数据作为训练集,这样会得到N个模型,用这N个模型最终的验证集的分类准确率的平均数作为此N-CV下分类器的性能指标.N一般大于等于2,实际操作时一般从3开始取,只有在原始数据集合数据量小的时候才会尝试取2-CV可以有效的避免过学习以及欠学习状态的发生,最后得到的结果也比较具有说服力.

二算法题（45分，每题15）

1. 给个正整数，求比它大的最小不重复数，不重复意味相邻数位不相同，即“1101”是重复数，“1234”是不重复数。

```
1 class Solution {
2 public:
3     bool chongfu(int number)
4     {
5         map<int, int> record;
6         while(number)
7         {
8             int t=number%10;
9             if(record.find(t)==record.end())
10                 record[t] = 1;
11             else
12                 return true;
13             number = number/10;
14         }
15     }
```

```

    return false;
17 }
int minNum(int number)
19 {
    int t = number+1;
21 while(1)
    {
23     if(!chongfu(t))
        break;
25     else
        t++;
27 }
    return t;
29 }
};

```

2. 长度为N的字符串，求最长回文串

```

class solution{
2 public:
    string change(string s)
4 {
    int n = s.size();
6    string res(2*n+2, '#');
    res[0] = '$';
8    int c=2;
    for(int i=0;i<n;i++)
10 {
        res[c]=s[i];
12        c+=2;
14    }
    res[c]='\0';
16    return res;
    }
18 int Manacher(string s)
    {
20     int n=s.length();
    int mx=0,id=0,maxlen=0,p[2*n+2];
22     p[0]=0;
    for(int i=1;i<n;i++)
24 {
        if(i<mx)
26         p[i] = min((mx-i),p[2*id-1]);
        else
28         p[i]=1;
        while(s[i-p[i]]==s[i+p[i]])

```

```

30         p[i]++;
31         if(mx<i+p[i])
32         {
33             id=i;
34             mx=i+p[i];
35         }
36         max_len = max(max_len,p[i]-1);
37     }
38     return max_len;
39 }
40 };

```

3. 数轴上有点 $a[0], a[1], \dots, a[n-1]$, 给个长度为 L 的绳子, 问绳子能在数轴上覆盖多长。

```

class Solution {
2 public:
4     int RopeCover(int points[], int m, int L) {
5         int size = m;
6         if(size <= 0){
7             return 0;
8         }
9         int max = 0;
10        int start = 1, end = 0;
11        while(end < start){
12            if(points[start] - points[end] <= L){
13                if(start - end + 1 > max)
14                    max = start - end + 1;
15                ++start;
16            }
17            else
18                ++end;
19        }
20        return max;
21    }
22 };

```

三.函数题 (25分)

给了函数模型, 1问求解这个函数模型的损失量, 2问要你改进这个函数。(函数记不起来了)

略。。。