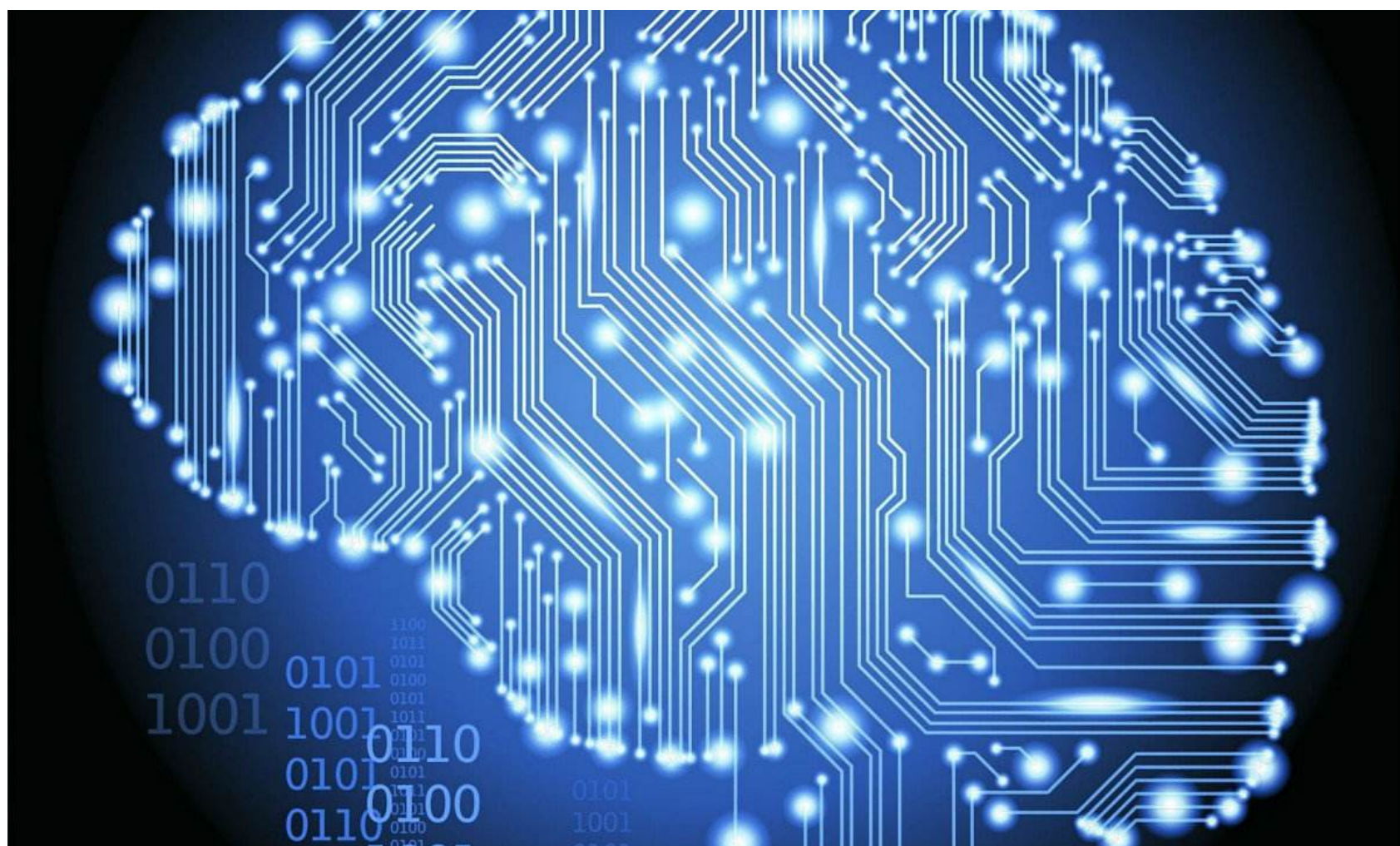


SFS与SBS特征选择算法

[2018年4月3日](#) by [xyjisaw](#)



(1)序列前向选择(SFS , Sequential Forward Selection)

算法描述：特征子集 X 从空集开始，每次选择一个特征 x 加入特征子集 X ，使得特征函数 $J(X)$ 最优。简单说就是，每次都选择一个使得评价函数的取值达到最优的特征加入，其实就是一种简单的贪心算法。

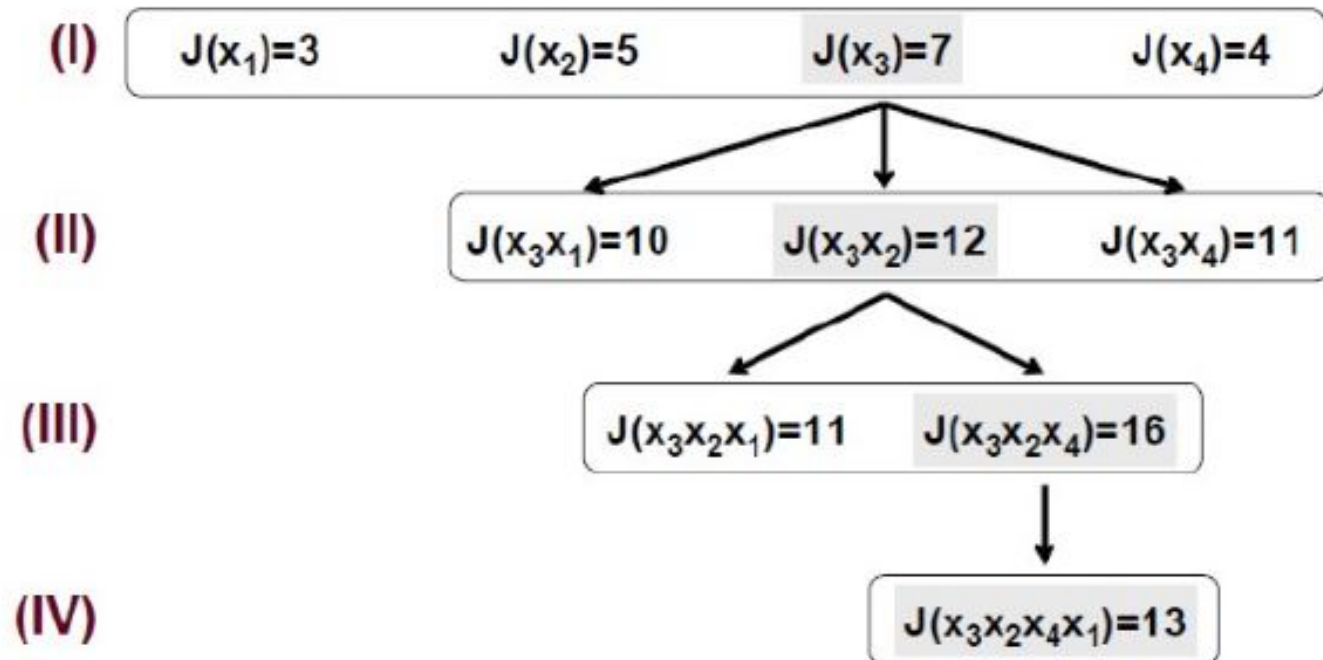
1. Start with the empty set $Y_0 = \{\emptyset\}$
2. Select the next best feature $x^* = \operatorname{argmax}_{x \notin Y_k} [J(Y_k + x)]$
3. Update $Y_{k+1} = Y_k + x^*$, $k = k + 1$
4. Go to 2

- Assuming the objective function $J(\underline{X})$ below, perform a Sequential Forward Selection to completion

$$J(\underline{X}) = -2x_1x_2 + 3x_1 + 5x_2 - 2x_1x_2x_3 + 7x_3 + 4x_4 - 2x_1x_2x_3x_4$$

- where x_k are *indicator variables* that determine if the k -th feature has been selected ($x_k=1$) or not ($x_k=0$)

Solution



算法评价：缺点是只能加入特征而不能去除特征。例如：特征A完全依赖于特征B与C，可以认为如果加入了特征B与C则A就是多余的。假设序列前向选择算法首先将A加入特征集，然后将B与C加入，那么特征子集中就包含了多余的特征A。

代码：

```
%----4.17编 顺序前进法特征选择 成功!
```

```

1
2
3
4
5
6 %----4.17编 顺序前进法特征选择 成功!
7 clear;
```

```
8   clc;
9   %-----特征导入 请自行修改
10  M=512;N=512;
11  load coourfeature16_0521_Aerial1 %%%共生矩阵 96.14%
12  wfeature{1}=coourfeature(:,1);
13  wfeature{2}=coourfeature(:,2);
14  wfeature{3}=coourfeature(:,3);
15  load fufeature_0521_SARAerial1_512%%复小波 98.26%
16  for i=1:13
17      wfeature{3+i}=wavefeature(:,i);
18  end
19  load wavefeature_0521_SARAerial1_512%%非下采样小波 97.58%
20  for i=1:7
21      wfeature{16+i}=wavefeature(:,i);
22  end
23  load wavefeature_0521_Aerial1%%小波 97.65%
24  for i=1:7
25      wfeature{23+i}=wavefeature(:,i);
26  end
27  % load rwt_cofeature96_0423_lsy1
28  % for i=1:96
29  %     wfeature{30+i}=feature(:,i);
30  % end
31  %%%%%%%%%-----归一化
32  [m n]=size(wfeature{1});
```

```

33 for j=1:30%一共30组特征 这里 请自行修改
34     mx=max(wfeature{j});
35     mi=min(wfeature{j});
36     mxx=(mx-mi);
37     mii=ones([m n])*mi;
38     wfeature{j}=(wfeature{j}-mii)./mxx;
39 end
40 %%-----SFS 先选4个特征尝试
41 chosen=[];%%表示已选的特征
42 chosen=[chosen 1];
43 Jc=0;%%选出的 J 值
44 for j=1:5 %选5个特征
45     J=zeros([1 30]);
46     for i=2:30 %一共30组特征 这里 请自行修改
47         [mm nn]=size(chosen);
48         for p=1:nn
49             if i==chosen(p)
50                 J(i)=0;
51                 break;
52             else
53                 J(i)=J(i)-sum(sum((wfeature{i}-wfeature{chosen(p)}).^2));
54             end
55         end
56     end
57 end

```

```
58     mi=min(J);
59     for i=1:30
60         if J(i)==0
61             J(i)=mi;
62         end
63     end
64     ma=max(J);
65     for i=1:30
66         if J(i)==ma
67             chosen=[chosen i];
68             break;
69         end
70     end
71 end
72 save Aerial1_6t_chosen chosen
73 [mm nn]=size(chosen);
74 tezh=[];
75 for i=1:nn
76     tezh=[tezh wfeature{chosen(i)}];
77 end
78 %%%%%%%%%%聚类
79 [IDC,U]=kmeans(tezh,2);
80     cc(IDC==1,1)=0;
81     cc(IDC==2,1)=0.75;
82 g=reshape(cc,M,N);
```

```
83 figure,imshow(g);
84
85
86
87
88
```

(2)序列后向选择(SBS , Sequential Backward Selection)

算法描述： 从特征全集O开始， 每次从特征集O中剔除一个特征x， 使得剔除特征x后评价函数值达到最优。

算法评价： 序列后向选择与序列前向选择正好相反， 它的缺点是特征只能去除不能加入。

1. Start with the full set $Y_0=X$

2. Remove the worst feature $x^- = \operatorname{argmax}_{x \in Y_k} [J(Y_k - x)]$

3. Update $Y_{k+1}=Y_k-x^-$; $k=k+1$

4. Go to 2

代码：

%----4.17编 顺序后退法特征选择

```
1
2
3
4
5 %----4.17编 顺序后退法特征选择
6 clear;
```



```
7   clc;
8   %-----特征导入 请自行修改
9   A=imread('lsy1.gif');
10  [M N]=size(A);
11  load coourfeature_0414_lsy1 %%%共生矩阵 96.14%
12  feature{1}=coourfeature(:,1);
13  feature{2}=coourfeature(:,2);
14  feature{3}=coourfeature(:,3);
15  load fuwavefeature_0413_lsy1 %%复小波 98.26%
16  for i=1:13
17      feature{3+i}=wavefeature(:,i);
18  end
19  load wavefeature_0413_feixia_lsy1 %%%非下采样小波 97.58%
20  for i=1:7
21      feature{16+i}=wavefeature(:,i);
22  end
23  load wavefeature_0417_lsy1 %%小波 97.65%
24  for i=1:7
25      feature{23+i}=wavefeature(:,i);
26  end
27  %%%%%%%%%-----归一化-归一化
28  [m n]=size(feature{1});
29  for j=1:30%一共30组特征 这里 请自行修改
30      mx=max(feature{j});
31      mi=min(feature{j});
```

```

32     mxx=(mx-mi);
33     mii=ones([m n])*mi;
34     feature{j}=(feature{j}-mii)./mxx;
35 end
36 %%-----SBS
37 chosen=[];dele=[];
38 for i=1:30
39     chosen=[chosen i];
40 end
41 for j=1:24  %%删10个,留20个
42     J=zeros([1 30]);ii=0; %J(1)是删1的结果,J(2)是删除2 的结果.....
43     for i=1:30 %???dele 是必要的么????????????????????????????%一共30组特征
44         这里 请自行修改
45
46         [mm nn]=size(chosen);
47         for p=1:nn
48             if sum(i==dele)~=0
49                 J(i)=0;
50                 break;
51             else
52                 for q=1:nn
53                     if (chosen(q)~=i) & (chosen(p)~=i)
54                         J(i)=J(i)-sum(sum((feature{chosen(q)}-feature{chosen(p)}).^2))
55                     end
56                 end
57             end
58         end
59     end
60 end

```



```
57     end
58 end
59     mi=min(J);
60     for cc=1:30
61         if J(cc)==0
62             J(cc)=mi;
63         end
64     end
65     [ma we]=max(J);
66     dele=[dele we];
67     for dd=1:nn
68         if chosen(dd)==we
69             chosen(dd)=[];
70
71     end
72
73 end
74 % chosen=[2 4 5 6 7 8 9 11 12 13 14 19 20 22 23 26 27 28 29 30];
75 [mm nn]=size(chosen);
76 tezh=[];
77 for i=1:nn
78     tezh=[tezh feature{chosen(i)}];
79 end
80 %%%%%%%%%%聚类
81 [IDC,U]=kmeans(tezh,2);
```

```
82     cc(IDC==1,1)=0;
83     cc(IDC==2,1)=0.75;
84 g=reshape(cc,M,N);
85 figure,imshow(g);
86 %%%%%%%%%%%%%%计算正确率
87 ju=ones(M)*0.75;
88 for i=1:M
89     for j=1:M/2
90         ju(i,j)=0;
91     end
92 end
93 ju2=g-ju;
94 prob=prod(size(find(ju2~=0)))/(m*n)
95 1-prob
96
97
98
99
```

另外，SFS与SBS都属于贪心算法，容易陷入局部最优值。

读者评分

[评分人数: 0 平均分: 0]

赏