Homework 3 SVM

Yawen Zhang

1. I have used the Sklearn implementation of SVM and used them in this analysis.

2. Comment on the performance of giving different C and kernels (linear and RBF)

1) **For linear kernel**, as C , Margin (M) , Number of support vectors for each class , both Train and Hold-out data prediction accuracy , also, computing time .

Additionally, as C increases from 0.01 to 100, it is obvious that the prediction accuracy of training data is higher than hold-out data, which can be recoginized as a sign of overfitting.

2) **For RBF kernel**,

3) **By comparing linear and RBF kernel**, we can see that: given the same C, linear kernel computes much faster than RBF kernel.

**kernel\_1: linear**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| C | 0.0001 | 0.001 | 0.01 | 0.1 | 1 | 10 | 100 |
| Margin (M) | 1.871 | 0.931 | 0.5 | 0.239 | 0.094 | 0.041 | 0.021 |
| Number of  surpport vectors | #3: 2778  #8: 2778 | #3: 1252  #8: 1256 | #3: 667  #8: 676 | #3: 484  #8: 500 | #3: 423  #8: 425 | #3: 400  #8: 398 | #3: 378  #8: 401 |
| Train\_accuracy | 0.938 | 0.964 | 0.971 | 0.977 | 0.982 | 0.984 | 0.986 |
| Hold\_out\_accuracy | 0.950 | 0.968 | 0.970 | 0.968 | 0.967 | 0.964 | 0.960 |

**kernel\_2: RBF** (gamma = 1/n\_features)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| C | 0.0001 | 0.001 | 0.01 | 0.1 | 1 | 10 | 100 |
| Number of  surpport vectors | #3: 4842  #8: 4842 | #3: 4842  #8: 4842 | #3: 4380  #8: 4380 | #3: 2079  #8: 2078 | #3: 957  #8: 957 | #3: 541  #8: 573 | #3: 381  #8: 443 |
| Train\_accuracy | 0.513 | 0.513 | 0.894 | 0.954 | 0.972 | 0.985 | 0.998 |
| Hold\_out\_accuracy | 0.509 | 0.509 | 0.905 | 0.963 | 0.973 | 0.979 | 0.989 |

3. C = 0.01 with linear kernel

   

   