Machine Learning Jiecao Chen ID:4746311
HOMEWORK #3

Question 1

a. Bagging: training B binary decision trees using all features, then use the majority vote. We are using the Information Gain to decide the **best** feature, As the features are continious, we need to choose the threshold to split the feature.

My way to choose the threshold: suppose the feature value lies in $[a, b]$

$$\delta = (b - a)/N$$

then we choose the threshold from $\{a + i\ \delta \mid 1 \leq i \leq N\}$ that can maximize the **Information Gain**.

In practicle, I set $N = 20$

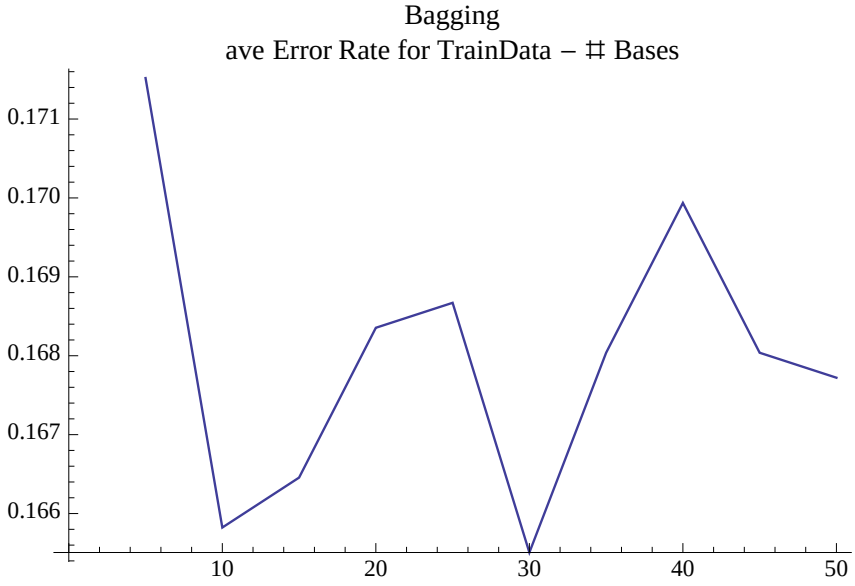
Following is the output

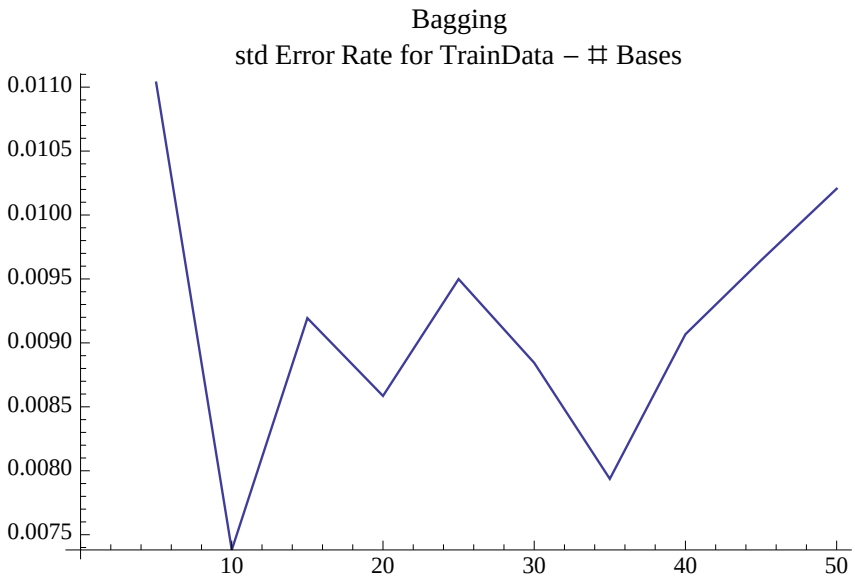
```
$ python myBagging2.py Ionosphere.csv 5 10 15 20 25 30 35 40 45 50
# Fold 0
   5 Bases: [0.18037974683544303, 0.11428571428571428]
   10 Bases: [0.17088607594936708, 0.11428571428571428]
   15 Bases: [0.18037974683544303, 0.11428571428571428]
   20 Bases: [0.18037974683544303, 0.11428571428571428]
   25 Bases: [0.17405063291139242, 0.11428571428571428]
   30 Bases: [0.18037974683544303, 0.11428571428571428]
   35 Bases: [0.17721518987341772, 0.11428571428571428]
   40 Bases: [0.17721518987341772, 0.11428571428571428]
   45 Bases: [0.18037974683544303, 0.11428571428571428]
   50 Bases: [0.18037974683544303, 0.11428571428571428]
# Fold 1
   5 Bases: [0.15822784810126583, 0.2571428571428571]
   10 Bases: [0.15822784810126583, 0.2571428571428571]
   15 Bases: [0.15822784810126583, 0.2571428571428571]
   20 Bases: [0.15822784810126583, 0.2571428571428571]
   25 Bases: [0.15822784810126583, 0.2571428571428571]
   30 Bases: [0.15822784810126583, 0.2571428571428571]
   35 Bases: [0.15822784810126583, 0.2571428571428571]
   40 Bases: [0.15822784810126583, 0.2571428571428571]
   45 Bases: [0.15822784810126583, 0.2571428571428571]
   50 Bases: [0.15822784810126583, 0.2571428571428571]
# Fold 2
   5 Bases: [0.17405063291139242, 0.17142857142857143]
   10 Bases: [0.16139240506329114, 0.17142857142857143]
   15 Bases: [0.16772151898734178, 0.17142857142857143]
   20 Bases: [0.16139240506329114, 0.17142857142857143]
   25 Bases: [0.16139240506329114, 0.17142857142857143]
```

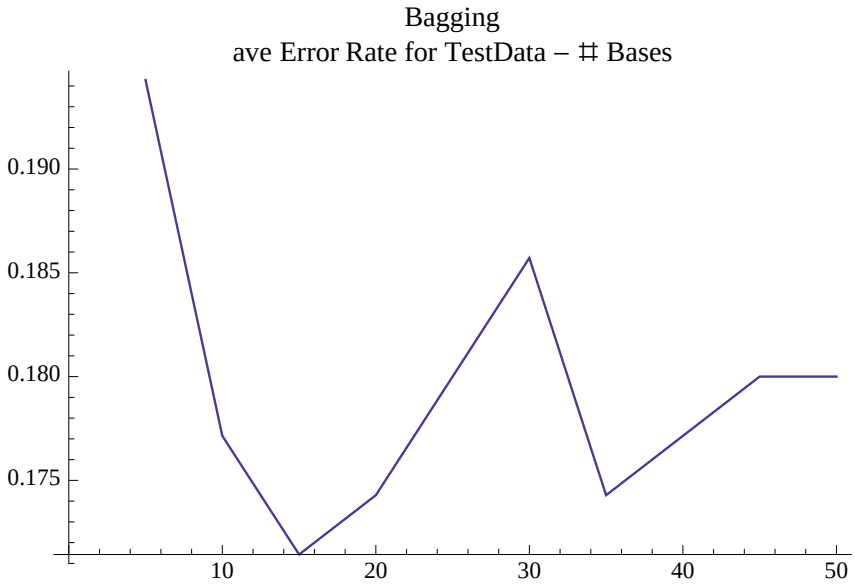
```
   30 Bases: [0.17088607594936708, 0.17142857142857143]
   35 Bases: [0.16772151898734178, 0.17142857142857143]
   40 Bases: [0.17405063291139242, 0.17142857142857143]
   45 Bases: [0.16139240506329114, 0.17142857142857143]
   50 Bases: [0.16139240506329114, 0.17142857142857143]
# Fold 3
   5 Bases: [0.15822784810126583, 0.2571428571428571]
   10 Bases: [0.1550632911392405, 0.2571428571428571]
   15 Bases: [0.15822784810126583, 0.2571428571428571]
   20 Bases: [0.16139240506329114, 0.2571428571428571]
   25 Bases: [0.15822784810126583, 0.2571428571428571]
   30 Bases: [0.1550632911392405, 0.2571428571428571]
   35 Bases: [0.15822784810126583, 0.2571428571428571]
   40 Bases: [0.15822784810126583, 0.2571428571428571]
   45 Bases: [0.15822784810126583, 0.2571428571428571]
   50 Bases: [0.15822784810126583, 0.2571428571428571]
# Fold 4
   5 Bases: [0.1518987341772152, 0.3142857142857143]
   10 Bases: [0.1550632911392405, 0.34285714285714286]
   15 Bases: [0.14873417721518986, 0.3142857142857143]
   20 Bases: [0.1550632911392405, 0.34285714285714286]
   25 Bases: [0.1550632911392405, 0.34285714285714286]
   30 Bases: [0.1550632911392405, 0.34285714285714286]
   35 Bases: [0.1550632911392405, 0.34285714285714286]
   40 Bases: [0.1550632911392405, 0.34285714285714286]
   45 Bases: [0.1550632911392405, 0.34285714285714286]
   50 Bases: [0.1550632911392405, 0.34285714285714286]
# Fold 5
   5 Bases: [0.17405063291139242, 0.17142857142857143]
   10 Bases: [0.17405063291139242, 0.17142857142857143]
   15 Bases: [0.17405063291139242, 0.14285714285714285]
   20 Bases: [0.17405063291139242, 0.14285714285714285]
   25 Bases: [0.17405063291139242, 0.17142857142857143]
   30 Bases: [0.17405063291139242, 0.17142857142857143]
   35 Bases: [0.17088607594936708, 0.14285714285714285]
   40 Bases: [0.17405063291139242, 0.17142857142857143]
   45 Bases: [0.17405063291139242, 0.17142857142857143]
   50 Bases: [0.17405063291139242, 0.17142857142857143]
# Fold 6
   5 Bases: [0.18670886075949367, 0.2857142857142857]
   10 Bases: [0.17405063291139242, 0.11428571428571428]
   15 Bases: [0.16455696202531644, 0.11428571428571428]
   20 Bases: [0.16772151898734178, 0.11428571428571428]
   25 Bases: [0.16772151898734178, 0.11428571428571428]
   30 Bases: [0.1550632911392405, 0.2]
```

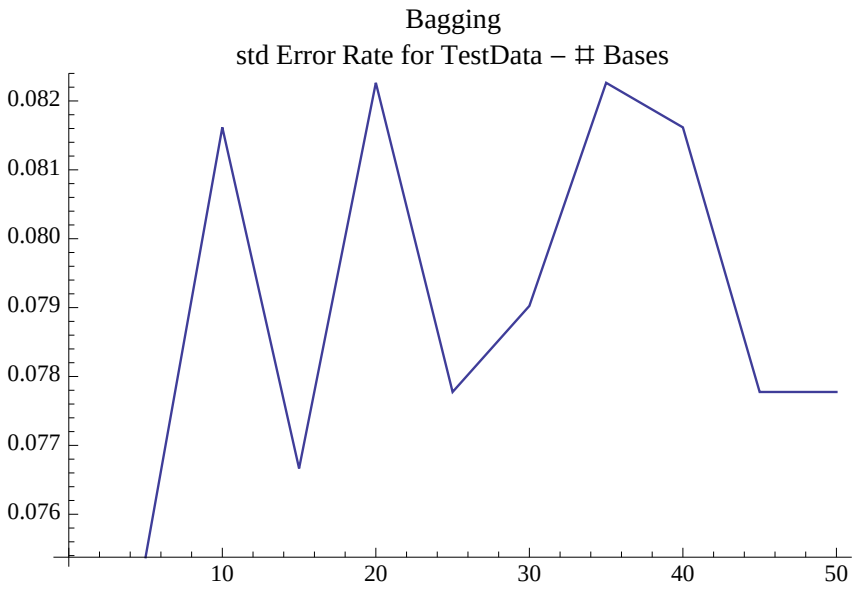
```
   35 Bases: [0.16772151898734178, 0.11428571428571428]
   40 Bases: [0.17405063291139242, 0.11428571428571428]
   45 Bases: [0.16772151898734178, 0.11428571428571428]
   50 Bases: [0.16455696202531644, 0.11428571428571428]
# Fold 7
   5 Bases: [0.18354430379746836, 0.08571428571428572]
   10 Bases: [0.17405063291139242, 0.05714285714285714]
   15 Bases: [0.17721518987341772, 0.05714285714285714]
   20 Bases: [0.17405063291139242, 0.05714285714285714]
   25 Bases: [0.18354430379746836, 0.08571428571428572]
   30 Bases: [0.17405063291139242, 0.05714285714285714]
   35 Bases: [0.17721518987341772, 0.05714285714285714]
   40 Bases: [0.18037974683544303, 0.05714285714285714]
   45 Bases: [0.18354430379746836, 0.08571428571428572]
   50 Bases: [0.18354430379746836, 0.08571428571428572]
# Fold 8
   5 Bases: [0.17405063291139242, 0.11428571428571428]
   10 Bases: [0.16772151898734178, 0.11428571428571428]
   15 Bases: [0.17088607594936708, 0.11428571428571428]
   20 Bases: [0.18037974683544303, 0.11428571428571428]
   25 Bases: [0.18037974683544303, 0.11428571428571428]
   30 Bases: [0.16772151898734178, 0.11428571428571428]
   35 Bases: [0.17721518987341772, 0.11428571428571428]
   40 Bases: [0.18037974683544303, 0.11428571428571428]
   45 Bases: [0.17721518987341772, 0.11428571428571428]
   50 Bases: [0.18037974683544303, 0.11428571428571428]
# Fold 9
   5 Bases: [0.17405063291139242, 0.17142857142857143]
   10 Bases: [0.16772151898734178, 0.17142857142857143]
   15 Bases: [0.16455696202531644, 0.17142857142857143]
   20 Bases: [0.17088607594936708, 0.17142857142857143]
   25 Bases: [0.17405063291139242, 0.17142857142857143]
   30 Bases: [0.16455696202531644, 0.17142857142857143]
   35 Bases: [0.17088607594936708, 0.17142857142857143]
   40 Bases: [0.16772151898734178, 0.17142857142857143]
   45 Bases: [0.16455696202531644, 0.17142857142857143]
   50 Bases: [0.16139240506329114, 0.17142857142857143]
# average errors in form of [trainData, testData]:
   5 bases: [0.17151898734177218, 0.1942857142857143]
   10 bases: [0.16582278481012658, 0.17714285714285716]
   15 bases: [0.16645569620253164, 0.17142857142857143]
   20 bases: [0.16835443037974684, 0.1742857142857143]
   25 bases: [0.16867088607594938, 0.18]
   30 bases: [0.16550632911392404, 0.18571428571428572]
   35 bases: [0.1680379746835443, 0.1742857142857143]
```

```
   40 bases: [0.1699367088607595, 0.17714285714285716]
   45 bases: [0.1680379746835443, 0.18]
   50 bases: [0.16772151898734178, 0.18]
# std derivation in form of [trainData, testData]:
   5 bases: [0.011035187198837145, 0.07537660547584525]
   10 bases: [0.007380951765626967, 0.08161632489763257]
   15 bases: [0.009193569016667057, 0.07666518779999278]
   20 bases: [0.008585227826740843, 0.08226388599364554]
   25 bases: [0.009498943683420021, 0.07777518622180685]
   30 bases: [0.008843790261064636, 0.07902466677679618]
   35 bases: [0.007936668483534462, 0.08226388599364554]
   40 bases: [0.009067435937907847, 0.08161632489763257]
   45 bases: [0.009645411806410233, 0.07777518622180685]
   50 bases: [0.01020538955480829, 0.07777518622180685]
```




Bagging
std Error Rate for TrainData – ⌗ Bases
0.0110
0.0105
0.0100
0.0095
0.0090
0.0085
0.0080
0.0075
10
20
30
40
50


Bagging
ave Error Rate for TestData – ⌗ Bases
0.190
0.185
0.180
0.175
10
20
30
40
50



b. Random Forest:
The only difference between **Bagging** and **Random Forest** is that **Random Forest** are using random m features rather than all features.

There is the output

```
python myRForest2.py Ionosphere.csv 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

# Fold 0
   1 random features: [0.25, 0.2571428571428571]
   2 random features: [0.25, 0.2571428571428571]
   3 random features: [0.29430379746835444, 0.34285714285714286]
   4 random features: [0.27848101265822783, 0.2571428571428571]
   5 random features: [0.2721518987341772, 0.2857142857142857]
   6 random features: [0.2626582278481013, 0.37142857142857144]
   7 random features: [0.25949367088607594, 0.2857142857142857]
   8 random features: [0.2721518987341772, 0.2571428571428571]
   9 random features: [0.29430379746835444, 0.4]
   10 random features: [0.2974683544303797, 0.37142857142857144]
   11 random features: [0.3227848101265823, 0.4]
   12 random features: [0.29430379746835444, 0.34285714285714286]
   13 random features: [0.29430379746835444, 0.34285714285714286]
   14 random features: [0.34810126582278483, 0.4]
   15 random features: [0.3259493670886076, 0.4]
   16 random features: [0.29430379746835444, 0.34285714285714286]
   17 random features: [0.3227848101265823, 0.4]
   18 random features: [0.31329113924050633, 0.4]
   19 random features: [0.2689873417721519, 0.34285714285714286]
   20 random features: [0.26582278481012656, 0.4]
   21 random features: [0.310126582278481, 0.4]
   22 random features: [0.3069620253164557, 0.4]
```

```
   23 random features: [0.22468354430379747, 0.34285714285714286]
   24 random features: [0.2689873417721519, 0.4]
   25 random features: [0.29430379746835444, 0.4]
   26 random features: [0.2088607594936709, 0.34285714285714286]
   27 random features: [0.2120253164556962, 0.34285714285714286]
   28 random features: [0.2120253164556962, 0.34285714285714286]
   29 random features: [0.1962025316455696, 0.34285714285714286]
   30 random features: [0.16139240506329114, 0.2857142857142857]
   31 random features: [0.16139240506329114, 0.2857142857142857]
   32 random features: [0.16139240506329114, 0.2857142857142857]
   33 random features: [0.16139240506329114, 0.2857142857142857]
   34 random features: [0.16139240506329114, 0.2857142857142857]
# Fold 1
   1 random features: [0.2563291139240506, 0.2]
   2 random features: [0.2563291139240506, 0.2]
   3 random features: [0.3069620253164557, 0.22857142857142856]
   4 random features: [0.2848101265822785, 0.2]
   5 random features: [0.2911392405063291, 0.22857142857142856]
   6 random features: [0.2563291139240506, 0.2]
   7 random features: [0.310126582278481, 0.22857142857142856]
   8 random features: [0.30063291139240506, 0.22857142857142856]
   9 random features: [0.2911392405063291, 0.22857142857142856]
   10 random features: [0.3670886075949367, 0.2857142857142857]
   11 random features: [0.3069620253164557, 0.22857142857142856]
   12 random features: [0.31329113924050633, 0.22857142857142856]
   13 random features: [0.34810126582278483, 0.2857142857142857]
   14 random features: [0.3069620253164557, 0.22857142857142856]
   15 random features: [0.3069620253164557, 0.22857142857142856]
   16 random features: [0.310126582278481, 0.22857142857142856]
   17 random features: [0.27848101265822783, 0.22857142857142856]
   18 random features: [0.3037974683544304, 0.2571428571428571]
   19 random features: [0.29430379746835444, 0.2571428571428571]
   20 random features: [0.30063291139240506, 0.2571428571428571]
   21 random features: [0.30063291139240506, 0.2571428571428571]
   22 random features: [0.2911392405063291, 0.2571428571428571]
   23 random features: [0.29430379746835444, 0.2571428571428571]
   24 random features: [0.27848101265822783, 0.2571428571428571]
   25 random features: [0.2310126582278481, 0.2]
   26 random features: [0.2310126582278481, 0.2]
   27 random features: [0.2310126582278481, 0.2]
   28 random features: [0.22151898734177214, 0.2]
   29 random features: [0.20569620253164558, 0.2]
   30 random features: [0.17405063291139242, 0.17142857142857143]
   31 random features: [0.17405063291139242, 0.17142857142857143]
   32 random features: [0.17405063291139242, 0.17142857142857143]
```

```
   33 random features: [0.17405063291139242, 0.17142857142857143]
   34 random features: [0.17405063291139242, 0.17142857142857143]
# Fold 2
   1 random features: [0.24367088607594936, 0.3142857142857143]
   2 random features: [0.24367088607594936, 0.3142857142857143]
   3 random features: [0.24367088607594936, 0.3142857142857143]
   4 random features: [0.2689873417721519, 0.34285714285714286]
   5 random features: [0.25949367088607594, 0.34285714285714286]
   6 random features: [0.24367088607594936, 0.3142857142857143]
   7 random features: [0.27531645569620256, 0.34285714285714286]
   8 random features: [0.25, 0.3142857142857143]
   9 random features: [0.2721518987341772, 0.37142857142857144]
   10 random features: [0.2911392405063291, 0.37142857142857144]
   11 random features: [0.3227848101265823, 0.42857142857142855]
   12 random features: [0.31329113924050633, 0.42857142857142855]
   13 random features: [0.2911392405063291, 0.37142857142857144]
   14 random features: [0.29430379746835444, 0.37142857142857144]
   15 random features: [0.2974683544303797, 0.37142857142857144]
   16 random features: [0.27848101265822783, 0.37142857142857144]
   17 random features: [0.2911392405063291, 0.4]
   18 random features: [0.2721518987341772, 0.37142857142857144]
   19 random features: [0.2848101265822785, 0.34285714285714286]
   20 random features: [0.25, 0.2857142857142857]
   21 random features: [0.23734177215189872, 0.2571428571428571]
   22 random features: [0.23734177215189872, 0.2571428571428571]
   23 random features: [0.23734177215189872, 0.2571428571428571]
   24 random features: [0.2974683544303797, 0.34285714285714286]
   25 random features: [0.2689873417721519, 0.2857142857142857]
   26 random features: [0.2689873417721519, 0.3142857142857143]
   27 random features: [0.23734177215189872, 0.22857142857142856]
   28 random features: [0.22151898734177214, 0.2]
   29 random features: [0.22468354430379747, 0.2571428571428571]
   30 random features: [0.18670886075949367, 0.2]
   31 random features: [0.17721518987341772, 0.2]
   32 random features: [0.17088607594936708, 0.2]
   33 random features: [0.16455696202531644, 0.2]
   34 random features: [0.16772151898734178, 0.2]
# Fold 3
   1 random features: [0.25, 0.2571428571428571]
   2 random features: [0.3575949367088608, 0.37142857142857144]
   3 random features: [0.25, 0.2571428571428571]
   4 random features: [0.2721518987341772, 0.3142857142857143]
   5 random features: [0.30063291139240506, 0.2857142857142857]
   6 random features: [0.28164556962025317, 0.2571428571428571]
   7 random features: [0.26582278481012656, 0.2571428571428571]
```

```
   8 random features: [0.28164556962025317, 0.2571428571428571]
   9 random features: [0.2848101265822785, 0.2571428571428571]
   10 random features: [0.29430379746835444, 0.3142857142857143]
   11 random features: [0.2974683544303797, 0.3142857142857143]
   12 random features: [0.2974683544303797, 0.3142857142857143]
   13 random features: [0.2974683544303797, 0.3142857142857143]
   14 random features: [0.3069620253164557, 0.3142857142857143]
   15 random features: [0.34177215189873417, 0.37142857142857144]
   16 random features: [0.31962025316455694, 0.3142857142857143]
   17 random features: [0.30063291139240506, 0.2857142857142857]
   18 random features: [0.2911392405063291, 0.2571428571428571]
   19 random features: [0.3322784810126582, 0.3142857142857143]
   20 random features: [0.3259493670886076, 0.3142857142857143]
   21 random features: [0.3259493670886076, 0.3142857142857143]
   22 random features: [0.2911392405063291, 0.2571428571428571]
   23 random features: [0.2879746835443038, 0.2571428571428571]
   24 random features: [0.24367088607594936, 0.2]
   25 random features: [0.2310126582278481, 0.2]
   26 random features: [0.24050632911392406, 0.22857142857142856]
   27 random features: [0.24367088607594936, 0.22857142857142856]
   28 random features: [0.22784810126582278, 0.22857142857142856]
   29 random features: [0.2310126582278481, 0.2]
   30 random features: [0.17405063291139242, 0.17142857142857143]
   31 random features: [0.18037974683544303, 0.17142857142857143]
   32 random features: [0.17405063291139242, 0.17142857142857143]
   33 random features: [0.17405063291139242, 0.17142857142857143]
   34 random features: [0.16772151898734178, 0.17142857142857143]
# Fold 4
   1 random features: [0.24367088607594936, 0.3142857142857143]
   2 random features: [0.24367088607594936, 0.3142857142857143]
   3 random features: [0.2848101265822785, 0.42857142857142855]
   4 random features: [0.2626582278481013, 0.4]
   5 random features: [0.27848101265822783, 0.42857142857142855]
   6 random features: [0.2626582278481013, 0.4]
   7 random features: [0.2689873417721519, 0.42857142857142855]
   8 random features: [0.2721518987341772, 0.42857142857142855]
   9 random features: [0.28164556962025317, 0.42857142857142855]
   10 random features: [0.29430379746835444, 0.42857142857142855]
   11 random features: [0.31645569620253167, 0.42857142857142855]
   12 random features: [0.2911392405063291, 0.42857142857142855]
   13 random features: [0.310126582278481, 0.42857142857142855]
   14 random features: [0.31329113924050633, 0.45714285714285713]
   15 random features: [0.3322784810126582, 0.45714285714285713]
   16 random features: [0.3322784810126582, 0.45714285714285713]
   17 random features: [0.3291139240506329, 0.42857142857142855]
```

```
   18 random features: [0.27848101265822783, 0.4]
   19 random features: [0.2626582278481013, 0.4]
   20 random features: [0.24050632911392406, 0.34285714285714286]
   21 random features: [0.2120253164556962, 0.3142857142857143]
   22 random features: [0.28164556962025317, 0.4]
   23 random features: [0.22151898734177214, 0.34285714285714286]
   24 random features: [0.3037974683544304, 0.42857142857142855]
   25 random features: [0.21518987341772153, 0.34285714285714286]
   26 random features: [0.23417721518987342, 0.34285714285714286]
   27 random features: [0.2848101265822785, 0.4]
   28 random features: [0.21518987341772153, 0.34285714285714286]
   29 random features: [0.20569620253164558, 0.3142857142857143]
   30 random features: [0.20253164556962025, 0.3142857142857143]
   31 random features: [0.16139240506329114, 0.3142857142857143]
   32 random features: [0.15822784810126583, 0.2571428571428571]
   33 random features: [0.16139240506329114, 0.2571428571428571]
   34 random features: [0.15822784810126583, 0.2571428571428571]
# Fold 5
   1 random features: [0.24367088607594936, 0.3142857142857143]
   2 random features: [0.24367088607594936, 0.3142857142857143]
   3 random features: [0.29430379746835444, 0.34285714285714286]
   4 random features: [0.2721518987341772, 0.3142857142857143]
   5 random features: [0.2626582278481013, 0.3142857142857143]
   6 random features: [0.27531645569620256, 0.34285714285714286]
   7 random features: [0.2721518987341772, 0.34285714285714286]
   8 random features: [0.28164556962025317, 0.34285714285714286]
   9 random features: [0.3259493670886076, 0.37142857142857144]
   10 random features: [0.29430379746835444, 0.34285714285714286]
   11 random features: [0.29430379746835444, 0.34285714285714286]
   12 random features: [0.30063291139240506, 0.34285714285714286]
   13 random features: [0.34177215189873417, 0.37142857142857144]
   14 random features: [0.34177215189873417, 0.37142857142857144]
   15 random features: [0.3037974683544304, 0.34285714285714286]
   16 random features: [0.33860759493670883, 0.37142857142857144]
   17 random features: [0.310126582278481, 0.34285714285714286]
   18 random features: [0.30063291139240506, 0.34285714285714286]
   19 random features: [0.3227848101265823, 0.3142857142857143]
   20 random features: [0.3227848101265823, 0.34285714285714286]
   21 random features: [0.2911392405063291, 0.3142857142857143]
   22 random features: [0.3259493670886076, 0.34285714285714286]
   23 random features: [0.3069620253164557, 0.34285714285714286]
   24 random features: [0.23734177215189872, 0.2857142857142857]
   25 random features: [0.23734177215189872, 0.2571428571428571]
   26 random features: [0.22784810126582278, 0.2571428571428571]
   27 random features: [0.27531645569620256, 0.2857142857142857]
```

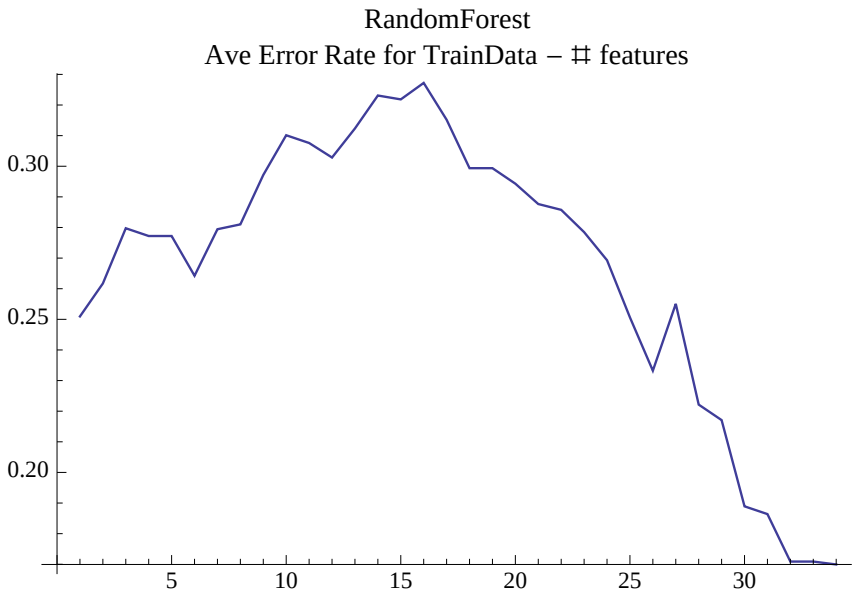
```
   28 random features: [0.21518987341772153, 0.2571428571428571]
   29 random features: [0.19936708860759494, 0.2571428571428571]
   30 random features: [0.20569620253164558, 0.2571428571428571]
   31 random features: [0.2088607594936709, 0.2571428571428571]
   32 random features: [0.16455696202531644, 0.14285714285714285]
   33 random features: [0.16455696202531644, 0.14285714285714285]
   34 random features: [0.16139240506329114, 0.14285714285714285]
# Fold 6
   1 random features: [0.26582278481012656, 0.11428571428571428]
   2 random features: [0.26582278481012656, 0.11428571428571428]
   3 random features: [0.31645569620253167, 0.14285714285714285]
   4 random features: [0.2911392405063291, 0.14285714285714285]
   5 random features: [0.29430379746835444, 0.17142857142857143]
   6 random features: [0.26582278481012656, 0.11428571428571428]
   7 random features: [0.3037974683544304, 0.11428571428571428]
   8 random features: [0.2848101265822785, 0.11428571428571428]
   9 random features: [0.3322784810126582, 0.17142857142857143]
   10 random features: [0.3227848101265823, 0.17142857142857143]
   11 random features: [0.31645569620253167, 0.14285714285714285]
   12 random features: [0.31962025316455694, 0.2]
   13 random features: [0.3291139240506329, 0.2]
   14 random features: [0.3575949367088608, 0.2]
   15 random features: [0.31645569620253167, 0.14285714285714285]
   16 random features: [0.3575949367088608, 0.2]
   17 random features: [0.3291139240506329, 0.2]
   18 random features: [0.3322784810126582, 0.17142857142857143]
   19 random features: [0.3449367088607595, 0.2]
   20 random features: [0.3037974683544304, 0.17142857142857143]
   21 random features: [0.3037974683544304, 0.17142857142857143]
   22 random features: [0.25949367088607594, 0.17142857142857143]
   23 random features: [0.310126582278481, 0.17142857142857143]
   24 random features: [0.2879746835443038, 0.17142857142857143]
   25 random features: [0.2310126582278481, 0.14285714285714285]
   26 random features: [0.23734177215189872, 0.14285714285714285]
   27 random features: [0.2310126582278481, 0.14285714285714285]
   28 random features: [0.2310126582278481, 0.14285714285714285]
   29 random features: [0.2310126582278481, 0.11428571428571428]
   30 random features: [0.2120253164556962, 0.14285714285714285]
   31 random features: [0.18037974683544303, 0.11428571428571428]
   32 random features: [0.16772151898734178, 0.14285714285714285]
   33 random features: [0.17088607594936708, 0.14285714285714285]
   34 random features: [0.17088607594936708, 0.14285714285714285]
# Fold 7
   1 random features: [0.25, 0.2571428571428571]
   2 random features: [0.25, 0.2571428571428571]
```

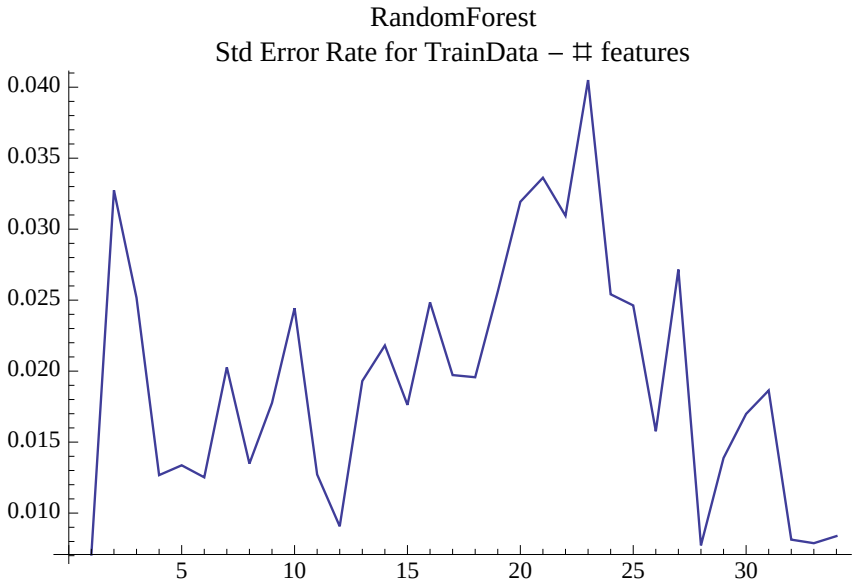
```
   3 random features: [0.25, 0.2571428571428571]
   4 random features: [0.27848101265822783, 0.2571428571428571]
   5 random features: [0.2689873417721519, 0.2571428571428571]
   6 random features: [0.25, 0.2571428571428571]
   7 random features: [0.25, 0.2571428571428571]
   8 random features: [0.2848101265822785, 0.2571428571428571]
   9 random features: [0.30063291139240506, 0.2571428571428571]
   10 random features: [0.34177215189873417, 0.2857142857142857]
   11 random features: [0.31329113924050633, 0.2571428571428571]
   12 random features: [0.3037974683544304, 0.2571428571428571]
   13 random features: [0.3037974683544304, 0.2571428571428571]
   14 random features: [0.34810126582278483, 0.2857142857142857]
   15 random features: [0.34177215189873417, 0.2571428571428571]
   16 random features: [0.3512658227848101, 0.2857142857142857]
   17 random features: [0.34810126582278483, 0.2857142857142857]
   18 random features: [0.27531645569620256, 0.22857142857142856]
   19 random features: [0.3037974683544304, 0.2571428571428571]
   20 random features: [0.2974683544303797, 0.2571428571428571]
   21 random features: [0.2848101265822785, 0.2571428571428571]
   22 random features: [0.23417721518987342, 0.17142857142857143]
   23 random features: [0.2468354430379747, 0.17142857142857143]
   24 random features: [0.29430379746835444, 0.2571428571428571]
   25 random features: [0.2563291139240506, 0.17142857142857143]
   26 random features: [0.22468354430379747, 0.17142857142857143]
   27 random features: [0.29430379746835444, 0.2571428571428571]
   28 random features: [0.22468354430379747, 0.17142857142857143]
   29 random features: [0.22468354430379747, 0.14285714285714285]
   30 random features: [0.18037974683544303, 0.11428571428571428]
   31 random features: [0.1962025316455696, 0.11428571428571428]
   32 random features: [0.18037974683544303, 0.11428571428571428]
   33 random features: [0.18037974683544303, 0.11428571428571428]
   34 random features: [0.18037974683544303, 0.11428571428571428]
# Fold 8
   1 random features: [0.2468354430379747, 0.2857142857142857]
   2 random features: [0.2468354430379747, 0.2857142857142857]
   3 random features: [0.2974683544303797, 0.3142857142857143]
   4 random features: [0.3037974683544304, 0.3714285714285714]
   5 random features: [0.26582278481012656, 0.2857142857142857]
   6 random features: [0.2848101265822785, 0.3142857142857143]
   7 random features: [0.27848101265822783, 0.3142857142857143]
   8 random features: [0.2848101265822785, 0.3142857142857143]
   9 random features: [0.29430379746835444, 0.34285714285714286]
   10 random features: [0.29430379746835444, 0.3142857142857143]
   11 random features: [0.28164556962025317, 0.2857142857142857]
   12 random features: [0.30063291139240506, 0.3142857142857143]
```

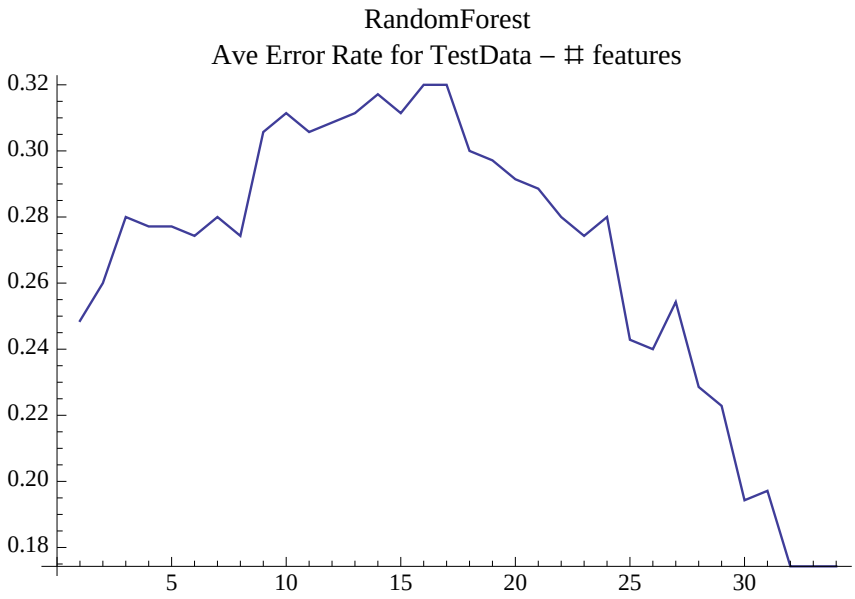
```
   13 random features: [0.2974683544303797, 0.3142857142857143]
   14 random features: [0.3069620253164557, 0.3142857142857143]
   15 random features: [0.3037974683544304, 0.3142857142857143]
   16 random features: [0.33860759493670883, 0.4]
   17 random features: [0.3291139240506329, 0.4]
   18 random features: [0.30063291139240506, 0.34285714285714286]
   19 random features: [0.28164556962025317, 0.3142857142857143]
   20 random features: [0.2911392405063291, 0.3142857142857143]
   21 random features: [0.3069620253164557, 0.37142857142857144]
   22 random features: [0.31962025316455694, 0.37142857142857144]
   23 random features: [0.310126582278481, 0.37142857142857144]
   24 random features: [0.23417721518987342, 0.2857142857142857]
   25 random features: [0.2848101265822785, 0.2857142857142857]
   26 random features: [0.21518987341772153, 0.2571428571428571]
   27 random features: [0.2879746835443038, 0.3142857142857143]
   28 random features: [0.21518987341772153, 0.2571428571428571]
   29 random features: [0.21518987341772153, 0.2571428571428571]
   30 random features: [0.18037974683544303, 0.2]
   31 random features: [0.2120253164556962, 0.2571428571428571]
   32 random features: [0.17088607594936708, 0.2]
   33 random features: [0.17088607594936708, 0.2]
   34 random features: [0.17088607594936708, 0.2]
# Fold 9
   1 random features: [0.25949367088607594, 0.17142857142857143]
   2 random features: [0.25949367088607594, 0.17142857142857143]
   3 random features: [0.25949367088607594, 0.17142857142857143]
   4 random features: [0.25949367088607594, 0.17142857142857143]
   5 random features: [0.27848101265822783, 0.17142857142857143]
   6 random features: [0.25949367088607594, 0.17142857142857143]
   7 random features: [0.310126582278481, 0.22857142857142856]
   8 random features: [0.2974683544303797, 0.22857142857142856]
   9 random features: [0.29430379746835444, 0.22857142857142856]
   10 random features: [0.3037974683544304, 0.22857142857142856]
   11 random features: [0.3037974683544304, 0.22857142857142856]
   12 random features: [0.29430379746835444, 0.22857142857142856]
   13 random features: [0.310126582278481, 0.22857142857142856]
   14 random features: [0.3069620253164557, 0.22857142857142856]
   15 random features: [0.34810126582278483, 0.22857142857142856]
   16 random features: [0.3512658227848101, 0.22857142857142856]
   17 random features: [0.31329113924050633, 0.22857142857142856]
   18 random features: [0.3259493670886076, 0.22857142857142856]
   19 random features: [0.2974683544303797, 0.22857142857142856]
   20 random features: [0.3449367088607595, 0.22857142857142856]
   21 random features: [0.3037974683544304, 0.22857142857142856]
   22 random features: [0.310126582278481, 0.17142857142857143]
```

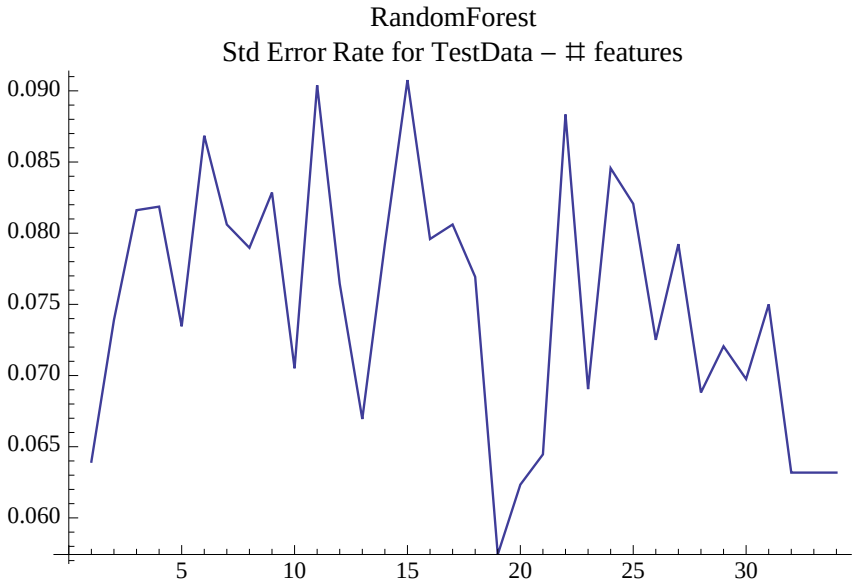
```
   23 random features: [0.3449367088607595, 0.22857142857142856]
   24 random features: [0.2468354430379747, 0.17142857142857143]
   25 random features: [0.2563291139240506, 0.14285714285714285]
   26 random features: [0.24367088607594936, 0.14285714285714285]
   27 random features: [0.25316455696202533, 0.14285714285714285]
   28 random features: [0.23734177215189872, 0.14285714285714285]
   29 random features: [0.23734177215189872, 0.14285714285714285]
   30 random features: [0.2120253164556962, 0.08571428571428572]
   31 random features: [0.2120253164556962, 0.08571428571428572]
   32 random features: [0.18670886075949367, 0.05714285714285714]
   33 random features: [0.18670886075949367, 0.05714285714285714]
   34 random features: [0.18670886075949367, 0.05714285714285714]
# average errors in form of [trainData, testData]:
   1 random features: [0.2509493670886076, 0.24857142857142853]
   2 random features: [0.2617088607594937, 0.25999999999999995]
   3 random features: [0.27974683544303797, 0.2799999999999999]
   4 random features: [0.27721518987341776, 0.27714285714285714]
   5 random features: [0.27721518987341776, 0.2771428571428571]
   6 random features: [0.26424050632911394, 0.27428571428571424]
   7 random features: [0.27943037974683543, 0.27999999999999997]
   8 random features: [0.2810126582278481, 0.2742857142857143]
   9 random features: [0.29715189873417724, 0.3057142857142857]
   10 random features: [0.310126582278481, 0.3114285714285714]
   11 random features: [0.30759493670886073, 0.3057142857142857]
   12 random features: [0.30284810126582273, 0.3085714285714286]
   13 random features: [0.31234177215189873, 0.3114285714285715]
   14 random features: [0.3231012658227848, 0.31714285714285717]
   15 random features: [0.32183544303797473, 0.3114285714285714]
   16 random features: [0.32721518987341774, 0.31999999999999995]
   17 random features: [0.31518987341772153, 0.31999999999999995]
   18 random features: [0.2993670886075949, 0.3]
   19 random features: [0.2993670886075949, 0.29714285714285715]
   20 random features: [0.29430379746835433, 0.2914285714285715]
   21 random features: [0.28765822784810124, 0.2885714285714286]
   22 random features: [0.28575949367088604, 0.2799999999999999]
   23 random features: [0.2784810126582279, 0.27428571428571435]
   24 random features: [0.2693037974683544, 0.2799999999999999]
   25 random features: [0.2506329113924051, 0.24285714285714283]
   26 random features: [0.23322784810126582, 0.24]
   27 random features: [0.2550632911392405, 0.2542857142857142]
   28 random features: [0.2221518987341772, 0.22857142857142856]
   29 random features: [0.2170886075949367, 0.2228571428571428]
   30 random features: [0.18892405063291143, 0.19428571428571426]
   31 random features: [0.18639240506329113, 0.19714285714285712]
   32 random features: [0.1708860759493671, 0.17428571428571427]
```

```
   33 random features: [0.1708860759493671, 0.17428571428571427]
   34 random features: [0.1699367088607595, 0.17428571428571427]
# std derivation in form of [trainData, testData]:
   1 random features: [0.007083237115696003, 0.06395151224456969]
   2 random features: [0.03273596135993671, 0.07390009803644322]
   3 random features: [0.02516569031907728, 0.08161632489763257]
   4 random features: [0.012674040756013158, 0.0818659930393966]
   5 random features: [0.0133662734695841, 0.07345691504104251]
   6 random features: [0.01251901293978006, 0.0868496237346895]
   7 random features: [0.020265522337449826, 0.08060991988380506]
   8 random features: [0.013485617564976274, 0.07897299977763002]
   9 random features: [0.01775821246404835, 0.08285714285714287]
   10 random features: [0.024430707677998462, 0.07050835816716038]
   11 random features: [0.012721361545722648, 0.09039595439746499]
   12 random features: [0.009067435937907847, 0.07645193234434086]
   13 random features: [0.019303797468354433, 0.06694499722205703]
   14 random features: [0.021807907789898797, 0.07923099785064028]
   15 random features: [0.017622349104209938, 0.09075645813867766]
   16 random features: [0.02484529927391205, 0.07959079015533782]
   17 random features: [0.01972207145681378, 0.08060991988380506]
   18 random features: [0.019569145359164944, 0.07693092581620721]
   19 random features: [0.02559968153943884, 0.05742786069211938]
   20 random features: [0.03192910950792263, 0.062335497797918374]
   21 random features: [0.0336233246461854, 0.06446008098673417]
   22 random features: [0.03094315383777123, 0.0883407133356589]
   23 random features: [0.04050138419183837, 0.06904597699196899]
   24 random features: [0.025413185459384622, 0.08456370621113567]
   25 random features: [0.02462667906119458, 0.08206518066482898]
   26 random features: [0.015762543551154768, 0.07250615737399727]
   27 random features: [0.027156250573322705, 0.07923099785064028]
   28 random features: [0.00772566811122386, 0.06880911187881311]
   29 random features: [0.013880830505988168, 0.0720544012166771]
   30 random features: [0.016985742911797014, 0.06975174637562116]
   31 random features: [0.018638676357755974, 0.07499659856232392]
   32 random features: [0.008129894037129827, 0.06318098396427424]
   33 random features: [0.007879683289866287, 0.06318098396427424]
   34 random features: [0.008378609047388432, 0.06318098396427424]
```


RandomForest
Ave Error Rate for TrainData – ⌗ features
0.30
0.25
0.20
5
10
15
20
25
30


RandomForest
Std Error Rate for TrainData – ⌗ features
0.040
0.035
0.030
0.025
0.020
0.015
0.010
5
10
15
20
25
30


RandomForest
Ave Error Rate for TestData – ⌗ features
0.32
0.30
0.28
0.26
0.24
0.22
0.20
0.18
5
10
15
20
25
30


RandomForest
Std Error Rate for TestData – ⌗ features
0.090
0.085
0.080
0.075
0.070
0.065
0.060
5
10
15
20
25
30

1. attached in last pages

2. (a) We will prove $VC(\mathcal{F}) = d+1$ by proving $VC(\mathcal{F}) \le d+1$ and $VC(\mathcal{F}) \ge d+1$

- Lower Bound

  Consider the set $\{P_0, P_1, \cdots, P_d\}$, where $P_i \in \mathbb{R}^d$

  Let $P_0 = (0, 0, \cdots, 0)^T$, $P_i = (0, 0, \cdots, 0, 1, 0, \cdots 0)^T$ ($i > 0$)

  (1 is the ith)

  Now, for any assignment $f(P_i) = C_i$ (where $C_i \in \{-1, 1\}$)

  Let $W = (C_1, C_2, \cdots, C_d)$, $W_0 = \frac{C_0}{2}$

  We can see that $f(P_i) = \text{Sign}(W^T P_i + W_0) = C_i$ for $0 \le i \le d$

  So we can conclude that $VC(\mathcal{F}) \ge d+1$

- Upper Bound

  here we will use "Radon's Theorem":

  Theorem Let S be a set of $(n+2)$ points in n dimensions, Then S can be partitioned into two (disjoint) subset S_1 and S_2 whose convex hulls intersect.

  ps: convex hull of S is defined as $\{\sum_{x_i \in S} \lambda_i x_i : \lambda_i \le 0 \text{ and } \sum \lambda_i = 1\}$

  ~~it can be seen that if all points of S is in one side of a~~

  Suppose S has $n+2$ points in n dimensions, by Radon's Theorem, S could be partitioned into two sets S_1 and S_2 whose convex hulls intersect. Let p be a point in that intersection. Assume that there exist a hyperplane

  $$W \cdot X_i \le -W_0, \quad \forall X_i \in \text{convex hull of } S_1$$
  $$W \cdot X_i \ge -W_0, \quad \forall X_i \in \text{convex hull of } S_2$$

  so that $W \cdot p \le -W_0$

  Which is contradicted by

  $$W \cdot p = \sum_{i: X_i \in S_2} \lambda_i X_i > \left(\sum_{i: X_i \in S_2} \lambda_i\right) \min(W \cdot X_i) = \min(W \cdot X_i) > -W_0$$
  $$i: X_i \in S_2 \qquad i: X_i \in S_2$$

  it can be seen that if all points of S is in one side of a hyperplanes, then so does its convex hull. next page

So no set of $n+2$ points can be shattered and $VC(\mathcal{F}) = d+1$ ✓

PS: proof of Randon's Theorem could be found in ~~wiki page~~. Wikipedia page.

(b) $R_n(\mathcal{F}) = \mathbb{E}\left[\sup_{f\in\mathcal{F}} \frac{1}{n}\sum_{i=1}^{n} \rho_i f(Z_i)\right]$ ... ①

$\mathbb{E}\left[\sup_{f\in\mathcal{F}} \frac{1}{n}\sum_{i=1}^{n} \rho_i (f(X_i') - f(X_i))\right]$

$\le \mathbb{E}\left[\sup_{f\in\mathcal{F}} \frac{1}{n}\sum \rho_i f(X_i') - \inf_{f\in\mathcal{F}} \frac{1}{n}\sum \rho_i f(X_i)\right]$ $\because -\inf(f) = \sup(-f)$

$= \mathbb{E}\left[\sup \frac{1}{n}\sum \rho_i f(X_i') + \sup \frac{1}{n}\sum -\rho_i f(X_i)\right]$ ... ②

$= \underbrace{\mathbb{E}\,\mathbb{E}\sup \frac{1}{n}\sum \rho_i f(X_i')]}_{③} + \underbrace{\mathbb{E}\left[\sup \frac{1}{n}\sum -\rho_i f(X_i')\right]}_{④}$ ... ⑤

③ $= R_n(\mathcal{F})$, for ④, note that ρ_i take -1, 1 ~~both~~ with probability $\frac{1}{2}$, then ④ = ③ $= R_n(\mathcal{F})$

$\therefore$ ⑤ $= 2R_n(\mathcal{F})$ thus We proved what we need.

(c) $R_n(\mathcal{C}_2(\mathcal{F})) = \mathbb{E}\left[\sup_{f\in\mathcal{C}_2(\mathcal{F})} \frac{1}{n}\sum \rho_i f(X_i)\right]$

$= \mathbb{E}\left[\sup_{\substack{f_1,f_2\in\mathcal{F} \\ \alpha\in[0,1]}} \underbrace{\frac{1}{n}\sum \rho_i [\alpha f_1(X_i) + (1-\alpha) f_2(X_i)]}_{②}\right]$ ... ①

as ~~$\mathcal{F}$ is finite~~, ~~Let's fix α~~, ~~suppose $f_1, f_2 \in \mathcal{F}$~~

~~help sup ② get it~~

~~for $f_1, f_2 \in \mathcal{F}$~~

Next page

for given $f_1, f_2 \in \bar{F}$, [3]

let $f(\alpha) = \frac{1}{n}\sum \rho_i (\alpha f_1(x_i) + (1-\alpha) f_2(x_i))$, $\alpha \in [0,1]$

$= \alpha\left(\frac{1}{n}\sum \rho_i f_1(x_i)\right) + (1-\alpha)\left(\frac{1}{n}\sum \rho_i f_2(x_i)\right)$

Which imply

$$\min\{f(0), f(1)\} \le f(\alpha) \le \max\{f(0), f(1)\}$$

which imply $\sup_{\alpha\in[0,1]} f(\alpha) = \max\{f(0), f(1)\}$ ... ③

$\because \bar{F}$ finite

how

$$\sup_{f_1,f_2\in\bar{F},\ \alpha\in[0,1]} ② = \sup_{f_1,f_2\in\bar{F}} \sup_{\alpha\in[0,1]} ②$$

$$= \sup_{f_1,f_2\in\bar{F}} \max\left\{\frac{1}{n}\sum \rho_i f_1(x_i), \frac{1}{n}\sum \rho_i f_2(x_i)\right\}$$

$$= \sup_{f\in\bar{F}} \frac{1}{n}\sum \rho_i f(x_i) = R_n(\bar{F})$$

So $R_n(\mathrm{CO}_2(\bar{F})) = R_n(\bar{F})$

3. (a) $f(w) = h(w) + \lambda g(w)$

$$\min \sum_{i=1}^{m} h_i(w_i) + r(\bar{z})$$

subject to $w_i - \bar{z} = 0$, $i = 1, 2, \cdots, m$

① $w_i^{k+1} := \underset{w_i}{\operatorname{argmin}} \left(h_i(w_i) + (\rho/2)\| w_i - \bar{z}^k + u_i^k \|_2^2 \right)$

② $\bar{z}^{k+1} := \underset{\bar{z}}{\operatorname{argmin}} \left(r(\bar{z}) + (m\rho/2)\| \bar{z} - \bar{w}^{k+1} - \bar{u}^k \|_2^2 \right)$

③ $u_i^{k+1} := u_i^k + w_i^{k+1} - \bar{z}^{k+1}$

[4]

(b) ① and ③ can be executed in parallel

② need require access to more than one mini-batch as we need to compute $\bar{u}$ and $\bar{w}$

(c) Yes, solve the optimization problem

①, ② requires inner iterations, eg. Gradient Descent.

4. (a) Let $f(w) = \langle \nabla h(w_t), w \rangle + \frac{1}{2\eta_t} \|w - w_t\|^2$

Let $\nabla f(w) = \nabla h(w_t) + \frac{1}{\eta_t}(w - w_t) = 0$

$\Rightarrow w = w_t - \eta_t \nabla h(w_t)$

it exactly matches the Gradient descent update

$$w_{t+1} = w_t - \eta_t \nabla h(w_t)$$

(b) denote $w = (x_1, \cdots, x_d)$ $\nabla h(w_t) = (\alpha_1, \cdots, \alpha_d)$

$w_t = (\beta_1, \cdots, \beta_d)$

consider $f(x_i) = \alpha_i x_i + \lambda |x_i| + \frac{1}{2\eta_t}(x_i - \beta_i)^2$

$$\frac{\partial f(x_i)}{\partial x_i} = \alpha_i + \lambda \operatorname{sign}(x_i) + \frac{x_i - \beta_i}{\eta_t} \quad \cdots ①$$

Set ① = 0

we have

$x_i = \beta_i - \alpha_i \eta_t - \lambda \eta_t \, sign(x_i)$ [5]

- case 1: $(\beta_i - \alpha_i \eta_t) - \lambda \eta_t > 0$

  We have $x_i = \beta_i - \alpha_i \eta_t - \lambda \eta_t$

- case 2: $(\beta_i - \alpha_i \eta_t) + \lambda \eta_t < 0$

  we have $x_i = (\beta_i - \alpha_i \eta_t) + \lambda \eta_t$

- case 3: $-\lambda \eta_t \le \beta_i - \alpha_i \eta_t \le \lambda \eta_t$

$$2\eta_t f(x_i) = (x_i - \beta_i)^2 + 2\alpha_i \eta_t x_i + 2\lambda \eta_t |x_i|$$
$$= x_i^2 - 2(\beta_i - \alpha_i \eta_t) x_i + 2\lambda \eta_t |x_i| + \beta_i^2 \quad \cdots (2)$$

by $-\lambda \eta_t \le \beta_i - \alpha_i \eta_t \le \lambda \eta_t$

$$(2) \ge x_i^2 - 2|\beta_i - \alpha_i \eta_t| |x_i| + 2\lambda \eta_t |x_i| + \beta_i^2$$
$$= x_i^2 + 2(\underbrace{\lambda \eta_t - |\beta_i - \alpha_i \eta_t|}_{\ge 0}) \underbrace{|x_i|}_{\ge 0} + \beta_i^2$$
$$\ge x_i^2 + \beta_i^2$$

So $2\eta_t f(x_i) \ge 0 = 2\eta_t f(0)$

$\therefore x_i = 0$ will minimize $f(x_i)$

thus
$$x_i = \begin{cases} \beta_i - \alpha_i \eta_t - \lambda \eta_t, & \beta_i - \alpha_i \eta_t \ge \lambda \eta_t \\ \beta_i - \alpha_i \eta_t + \lambda \eta_t, & \beta_i - \alpha_i \eta_t < \lambda \eta_t \\ 0, & |\beta_i - \alpha_i \eta_t| \le \lambda \eta_t \end{cases} \quad \cdots (3)$$

this is the closed form for each component of W_{t+1}

(c) from ③ in (b), we know that if α_i is given, then x_i require $O(1)$

So if $\nabla h(w_t)$ is given, computing w_{t+1} requires $O(d)$

now consider $\nabla h(w_t)$

$$\nabla h(w_t) = -\sum_{i=1}^{N} \underbrace{\left\{ y_i x_i^T - \frac{x_i \exp(w^T x_i)}{1+\exp(w^T x_i)} \right\}}_{④}$$

each ④ requires $O(d)$

thus $\nabla h(w)$ requires $O(dN)$

and

$O(d) \& O(dN) \Rightarrow O(dN)$ that's what we need.

(4) NO. Let $w = (w_1, w_2, \cdots, w_d)$, $\nabla h(w_t) = (\alpha_1, \alpha_2, \cdots, \alpha_t)$

consider

$$\alpha_j = \frac{\partial h(w_t)}{\partial w_j} = -\sum_{i=1}^{N} \left\{ y_i x_{ij} - \frac{x_{ij} \exp(w^T x_i)}{1+\exp(w^T x_i)} \right\}$$

we can see that to calculate α_j, we need to compute the term $\exp(w^T x_i)$ whose complexity depends on d, and α_j is required for each iteration in stochastic gradient descent (if it exists)

which suggests that the complexity of each iteration depends on d.