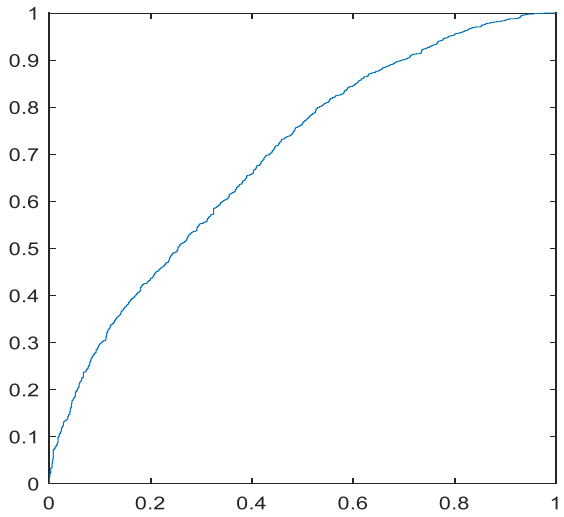
Q1:

```
>> trpre = pre(1:3000,:);
tepre = pre(3001:end,:);
trres=res(2:3001,:);
teres=res(3001:end,:);
factors=glmfit(trpre, trres, 'binomial');
```

For training data:

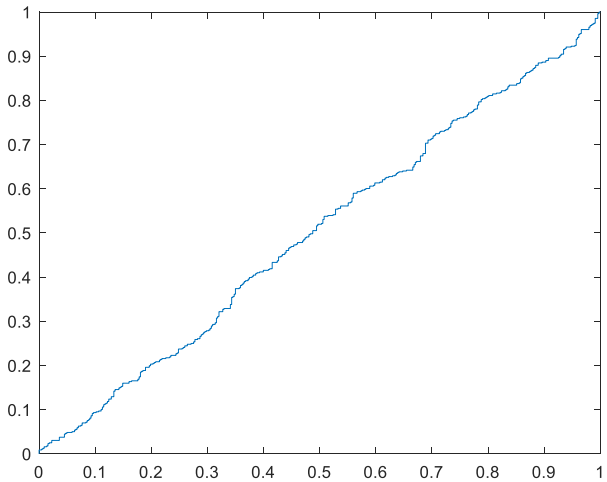
```
>> prob1=glmval(factors, trpre, 'logit');
[X, Y, thre1, AUC2]=perfcurve(trres, prob1,1);
plot(X,Y)
[AUC2]
```



AUC = 0.6952

For testing data:

```
>> prob2=glmval(factors, tepre, 'logit');
[X, Y, thre2, AUC3]=perfcurve(teres, prob2,1);
plot(X,Y)
[AUC3]
```



AUC = 0.5004

Q2:

After applied x(isinf(x)) = 0, the data and AUCdidn't change. Hence, there is no infinity number in the data.

Q3:

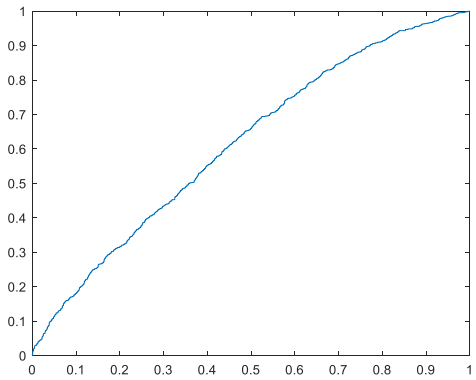
```
>> pre2=pre(2:end,:)-pre(1:(end-1),:);
res2=res(2:end,:);
fpre=[pre(2:end,:) pre2(:,:)];
trainfpre=fpre(1:3000,:);
testfpre=fpre(3001:end,:);
trainres2=res2(1:3000,:);
testres2=res2(3001:end,:);
[feaInd]=featureSelection(trainfpre,trainres2);
```

```
Start forward sequential feature selection:
Initial columns included:  none
Columns that can not be included:  none
Step 1, used initial columns, criterion value 4033.98
Step 2, added column 175, criterion value 4023.1
Step 3, added column 299, criterion value 4007.74
Step 4, added column 94, criterion value 3998.46
Step 5, added column 430, criterion value 3986.69
Step 6, added column 63, criterion value 3981.3
Step 7, added column 570, criterion value 3975.11
Step 8, added column 387, criterion value 3969.5
Step 9, added column 298, criterion value 3964.05
Step 10, added column 53, criterion value 3958.54
Step 11, added column 187, criterion value 3953.4
Step 12, added column 118, criterion value 3948.2
Step 13, added column 830, criterion value 3943.21
Step 14, added column 627, criterion value 3937.93
Step 15, added column 651, criterion value 3932.85
Step 16, added column 775, criterion value 3926.4
Step 17, added column 203, criterion value 3922.49
Step 18, added column 403, criterion value 3917.53
Step 19, added column 279, criterion value 3913.57
Step 20, added column 179, criterion value 3908.81
Step 21, added column 679, criterion value 3904.4
Step 22, added column 466, criterion value 3900.41
Final columns included:  53 63 94 118 175 179 187 203 279 298 299 387 403 430 466 570 627 651 679 775 830
```

Applying all the features chosen:

Training data:

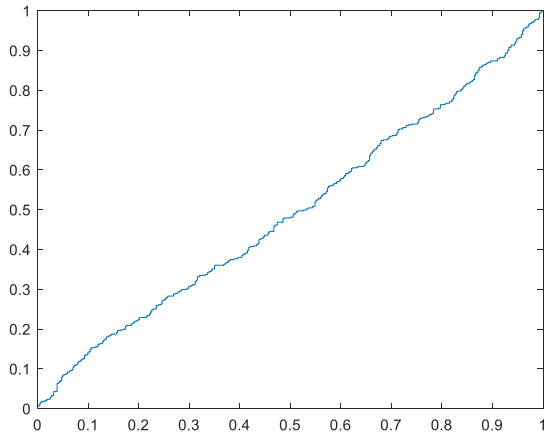
```
>> trainpre2=trainfpre(:,feaInd);
factors=glmfit(trainpre2, trainres2, 'binomial');
trainprob=glmval(factors, trainpre2, 'logit');
[X2, Y2, thre2, AUC2]=perfcurve(trainres2, trainprob,1);
trainpre2=trainfpre(:,feaInd);
factors=glmfit(trainpre2, trainres2, 'binomial');
trainprob=glmval(factors, trainpre2, 'logit');
[X2, Y2, thre2, AUC2]=perfcurve(trainres2, trainprob,1);
plot(X2,Y2)
>> [AUC2]
```



AUC = 0.6159

Testing data:

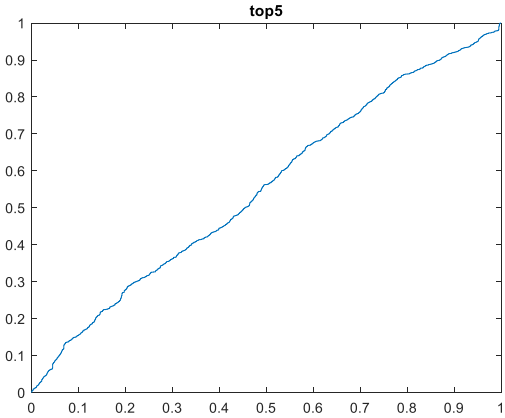
```
>> testpre2=testfpre(:,feaInd);
testprob=glmval(factors, testpre2, 'logit');
[X3, Y3, thre3, AUC3]=perfcurve(testres2, testprob,1);
plot(X3,Y3)
[AUC3]
```



AUC = 0.493

Using top 5:

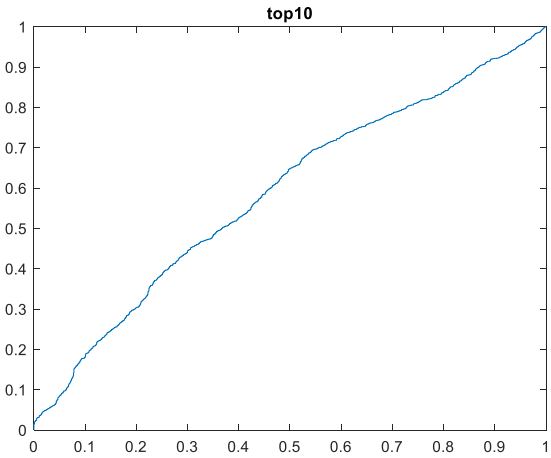
```
>> newpredictor1 = newpredictor(:,[53,63,94,118,175]);
factors11=glmfit(newpredictor1, response, 'binomial');
prob=glmval(factors11, newpredictor1, 'logit');
[X, Y, thre1, AUC1]=perfcurve(response, prob,1);
plot(X,Y),title('top5');
[AUC1]
```



AUC = 0.5519

Using top 10:

```
>> newpredictor2 = newpredictor(:,[53,63,94,118,175,179,187,203,279,298]);
factors=glmfit(newpredictor2, response, 'binomial');
prob=glmval(factors, newpredictor2, 'logit');
[X, Y, thre2, AUC2]=perfcurve(response, prob,1);
plot(X,Y),title('top10');
[AUC2]
```



AUC = 0.5883