1 The result of each algorithm

1.1 logistic regression

1.1.1 Running Results of Predict the 2017 Competition

```
[1] "Confusion Matrix for Logistic Regression (Year: 2017 )"
Confusion Matrix and Statistics

          Reference
Prediction Win Loss
      Win  998  270
      Loss 270  998

               Accuracy : 0.7871
                 95% CI : (0.7706, 0.8029)
    No Information Rate : 0.5
    P-Value [Acc > NIR] : <2e-16

                  Kappa : 0.5741

 Mcnemar's Test P-Value : 1

            Sensitivity : 0.7871
            Specificity : 0.7871
         Pos Pred Value : 0.7871
         Neg Pred Value : 0.7871
             Prevalence : 0.5000
         Detection Rate : 0.3935
   Detection Prevalence : 0.5000
      Balanced Accuracy : 0.7871

       'Positive' Class : Win
```

1.1.2 Running Results of Predict the 2018 Competition

```
[1] "Confusion Matrix for Logistic Regression (Year: 2018 )"
Confusion Matrix and Statistics

          Reference
Prediction Win Loss
      Win  531  155
      Loss 155  531

               Accuracy : 0.7741
                 95% CI : (0.751, 0.7959)
    No Information Rate : 0.5
    P-Value [Acc > NIR] : <2e-16

                  Kappa : 0.5481

 Mcnemar's Test P-Value : 1

            Sensitivity : 0.7741
            Specificity : 0.7741
         Pos Pred Value : 0.7741
         Neg Pred Value : 0.7741
             Prevalence : 0.5000
         Detection Rate : 0.3870
   Detection Prevalence : 0.5000
      Balanced Accuracy : 0.7741

       'Positive' Class : Win
```
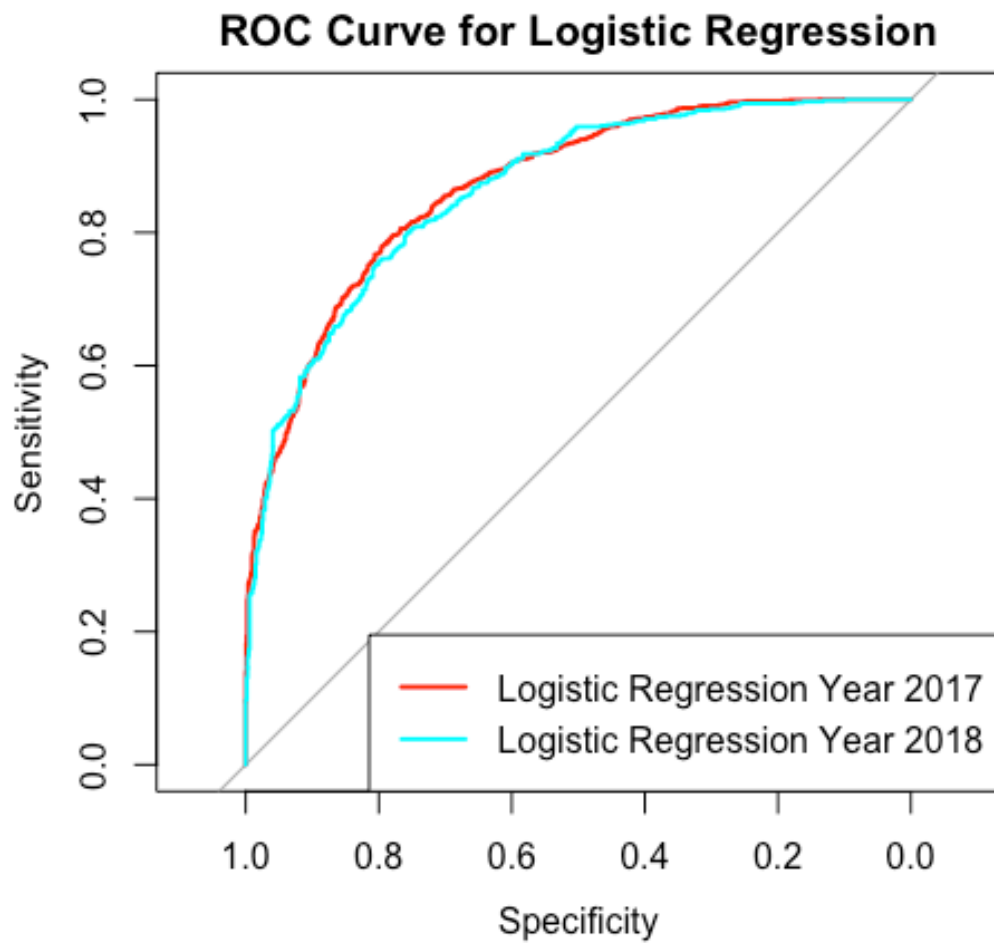
1.1.3 ROC of Predict the 2017 and 2018 Competition

ROC Curve for Logistic Regression


```
Logistic Regression Results:
          Year  Accuracy Precision    Recall        F1       AUC
Accuracy  2017 0.7870662 0.7870662 0.7870662 0.7870662 0.8704426
Accuracy1 2018 0.7740525 0.7740525 0.7740525 0.7740525 0.8647184
```

1.2 Random Forest

1.2.1 Running Results of Predict the 2017 Competition

```
[1] "Confusion Matrix for Random Forest (Year: 2017 )"
Confusion Matrix and Statistics

          Reference
Prediction Win Loss
      Win  970  283
      Loss 298  985

               Accuracy : 0.7709
                 95% CI : (0.754, 0.7871)
    No Information Rate : 0.5
    P-Value [Acc > NIR] : <2e-16

                  Kappa : 0.5418

 Mcnemar's Test P-Value : 0.5614

            Sensitivity : 0.7650
            Specificity : 0.7768
         Pos Pred Value : 0.7741
         Neg Pred Value : 0.7677
             Prevalence : 0.5000
         Detection Rate : 0.3825
   Detection Prevalence : 0.4941
      Balanced Accuracy : 0.7709

       'Positive' Class : Win
```

1.2.2 Running Results of Predict the 2017 Competition

```
[1] "Confusion Matrix for Random Forest (Year: 2018 )"
Confusion Matrix and Statistics

          Reference
Prediction Win Loss
      Win  513  179
      Loss 173  507

               Accuracy : 0.7434
                 95% CI : (0.7195, 0.7664)
    No Information Rate : 0.5
    P-Value [Acc > NIR] : <2e-16

                  Kappa : 0.4869

 Mcnemar's Test P-Value : 0.7899

            Sensitivity : 0.7478
            Specificity : 0.7391
         Pos Pred Value : 0.7413
         Neg Pred Value : 0.7456
             Prevalence : 0.5000
         Detection Rate : 0.3739
   Detection Prevalence : 0.5044
      Balanced Accuracy : 0.7434

       'Positive' Class : Win
```
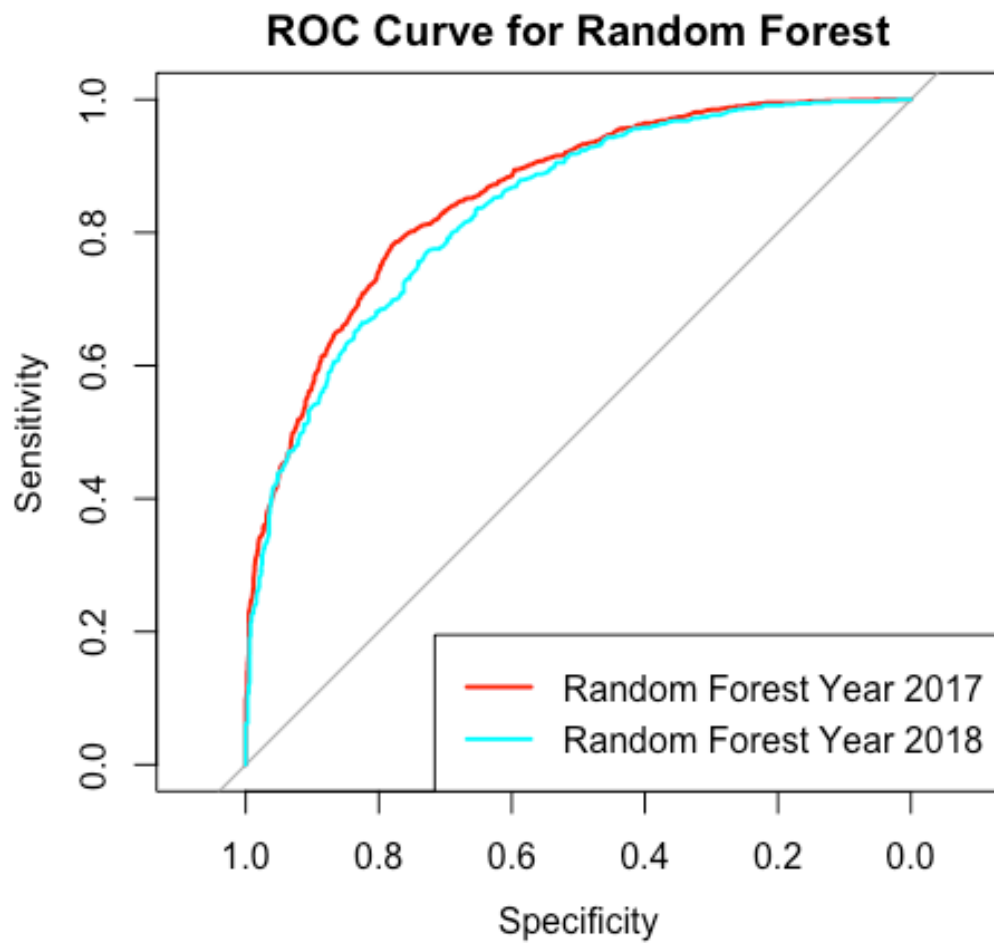
1.2.3 ROC of Predict the 2017 and 2018 Competition

ROC Curve for Random Forest


```
Random Forest Results:
          Year Accuracy Precision    Recall        F1        AUC
Accuracy  2017 0.773265 0.7747819 0.7705047 0.7726374 0.8542561
Accuracy1 2018 0.744898 0.7456140 0.7434402 0.7445255 0.8365424
```

1.3 SVM

1.3.1 Running Results of Predict the 2017 Competition

```
[1] "Confusion Matrix for Support Vector Machine (Year: 2017 )"
Confusion Matrix and Statistics

          Reference
Prediction Win Loss
      Win  998  270
      Loss 270  998

               Accuracy : 0.7871
                 95% CI : (0.7706, 0.8029)
    No Information Rate : 0.5
    P-Value [Acc > NIR] : <2e-16

                  Kappa : 0.5741

 Mcnemar's Test P-Value : 1

            Sensitivity : 0.7871
            Specificity : 0.7871
         Pos Pred Value : 0.7871
         Neg Pred Value : 0.7871
             Prevalence : 0.5000
         Detection Rate : 0.3935
   Detection Prevalence : 0.5000
      Balanced Accuracy : 0.7871

       'Positive' Class : Win
```
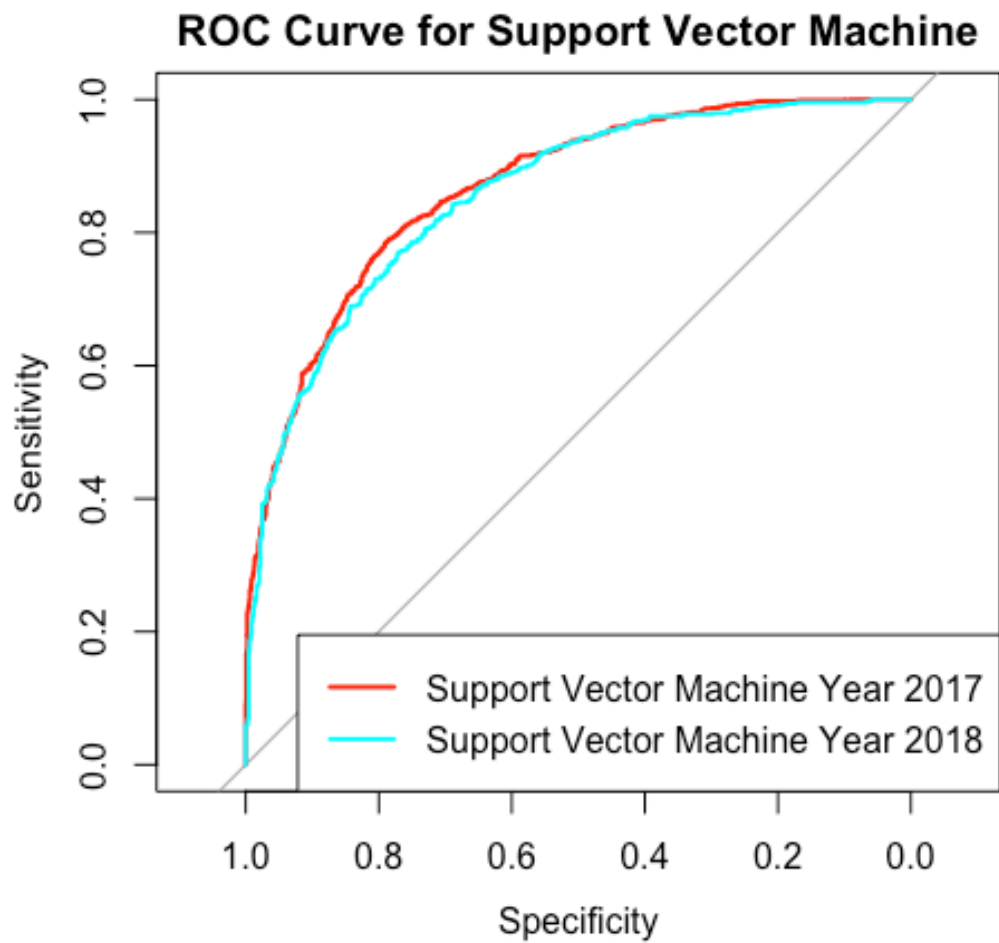
1.3.2 Running Results of Predict the 2018 Competition

```
[1] "Confusion Matrix for Support Vector Machine (Year: 2018 )"
Confusion Matrix and Statistics

          Reference
Prediction Win Loss
      Win  528  158
      Loss 158  528

               Accuracy : 0.7697
                 95% CI : (0.7465, 0.7917)
    No Information Rate : 0.5
    P-Value [Acc > NIR] : <2e-16

                  Kappa : 0.5394

 Mcnemar's Test P-Value : 1

            Sensitivity : 0.7697
            Specificity : 0.7697
         Pos Pred Value : 0.7697
         Neg Pred Value : 0.7697
             Prevalence : 0.5000
         Detection Rate : 0.3848
   Detection Prevalence : 0.5000
      Balanced Accuracy : 0.7697

       'Positive' Class : Win
```

1.3.3 ROC of Predict the 2017 and 2018 Competition

ROC Curve for Support Vector Machine


```
Support Vector Machine Results:
          Year  Accuracy Precision    Recall        F1        AUC
Accuracy  2017 0.7866719 0.7868982 0.7862776 0.7865878 0.8667895
Accuracy1 2018 0.7696793 0.7696793 0.7696793 0.7696793 0.8560931
```

1.4 Naive Bayes

1.4.1 Running Results of Predict the 2017 Competition

```
[1] "Confusion Matrix for Naive Bayes (Year: 2017 )"
Confusion Matrix and Statistics

          Reference
Prediction Win Loss
      Win  975  293
      Loss 293  975

               Accuracy : 0.7689
                 95% CI : (0.752, 0.7852)
    No Information Rate : 0.5
    P-Value [Acc > NIR] : <2e-16

                  Kappa : 0.5379

 Mcnemar's Test P-Value : 1

            Sensitivity : 0.7689
            Specificity : 0.7689
         Pos Pred Value : 0.7689
         Neg Pred Value : 0.7689
             Prevalence : 0.5000
         Detection Rate : 0.3845
   Detection Prevalence : 0.5000
      Balanced Accuracy : 0.7689

       'Positive' Class : Win
```

1.4.2 Running Results of Predict the 2018 Competition

```
[1] "Confusion Matrix for Naive Bayes (Year: 2018 )"
Confusion Matrix and Statistics

          Reference
Prediction Win Loss
      Win  516  170
      Loss 170  516

               Accuracy : 0.7522
                 95% CI : (0.7285, 0.7748)
    No Information Rate : 0.5
    P-Value [Acc > NIR] : <2e-16

                  Kappa : 0.5044

 Mcnemar's Test P-Value : 1

            Sensitivity : 0.7522
            Specificity : 0.7522
         Pos Pred Value : 0.7522
         Neg Pred Value : 0.7522
             Prevalence : 0.5000
         Detection Rate : 0.3761
   Detection Prevalence : 0.5000
      Balanced Accuracy : 0.7522

       'Positive' Class : Win
```
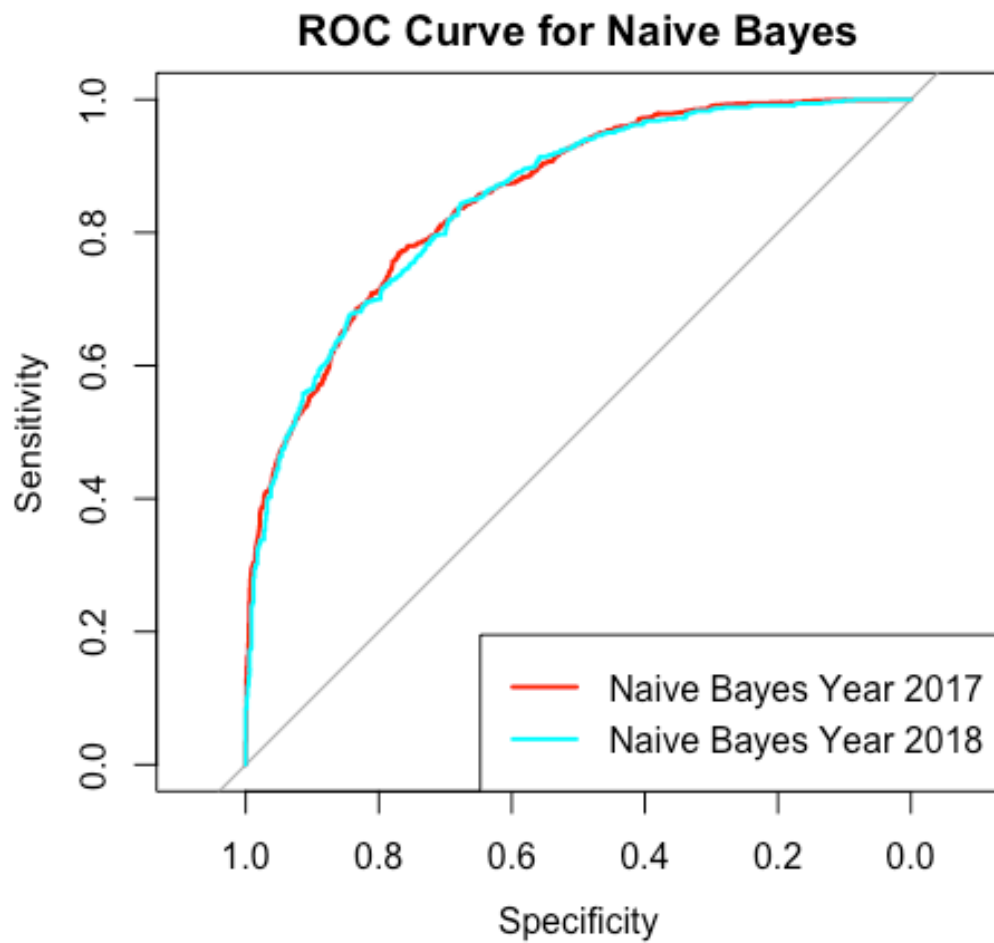
1.4.3 ROC of Predict the 2017 and 2018 Competition

ROC Curve for Naive Bayes


```
Naive Bayes Results:
         Year  Accuracy Precision    Recall        F1        AUC
Accuracy  2017 0.7689274 0.7689274 0.7689274 0.7689274 0.8533216
Accuracy1 2018 0.7521866 0.7521866 0.7521866 0.7521866 0.8497777
```

1.5 Neural Network

1.5.1 Running Results of Predict the 2017 Competition

```
[1] "Confusion Matrix for Neural Network (Year: 2017 )"
Confusion Matrix and Statistics

          Reference
Prediction  Win Loss
      Win   992  266
      Loss  276 1002

               Accuracy : 0.7863
                 95% CI : (0.7698, 0.8021)
    No Information Rate : 0.5
    P-Value [Acc > NIR] : <2e-16

                  Kappa : 0.5726

 Mcnemar's Test P-Value : 0.6991

            Sensitivity : 0.7823
            Specificity : 0.7902
         Pos Pred Value : 0.7886
         Neg Pred Value : 0.7840
             Prevalence : 0.5000
         Detection Rate : 0.3912
   Detection Prevalence : 0.4961
      Balanced Accuracy : 0.7863

       'Positive' Class : Win
```

1.5.2 Running Results of Predict the 2018 Competition

```
[1] "Confusion Matrix for Neural Network (Year: 2018 )"
Confusion Matrix and Statistics

          Reference
Prediction Win Loss
      Win 528  154
      Loss 158  532

               Accuracy : 0.7726
                 95% CI : (0.7495, 0.7945)
    No Information Rate : 0.5
    P-Value [Acc > NIR] : <2e-16

                  Kappa : 0.5452

 Mcnemar's Test P-Value : 0.8651

            Sensitivity : 0.7697
            Specificity : 0.7755
         Pos Pred Value : 0.7742
         Neg Pred Value : 0.7710
             Prevalence : 0.5000
         Detection Rate : 0.3848
   Detection Prevalence : 0.4971
      Balanced Accuracy : 0.7726

       'Positive' Class : Win
```

## 1.5.3 ROC of Predict the 2017 and 2018 Competition

ROC Curve for Neural Network
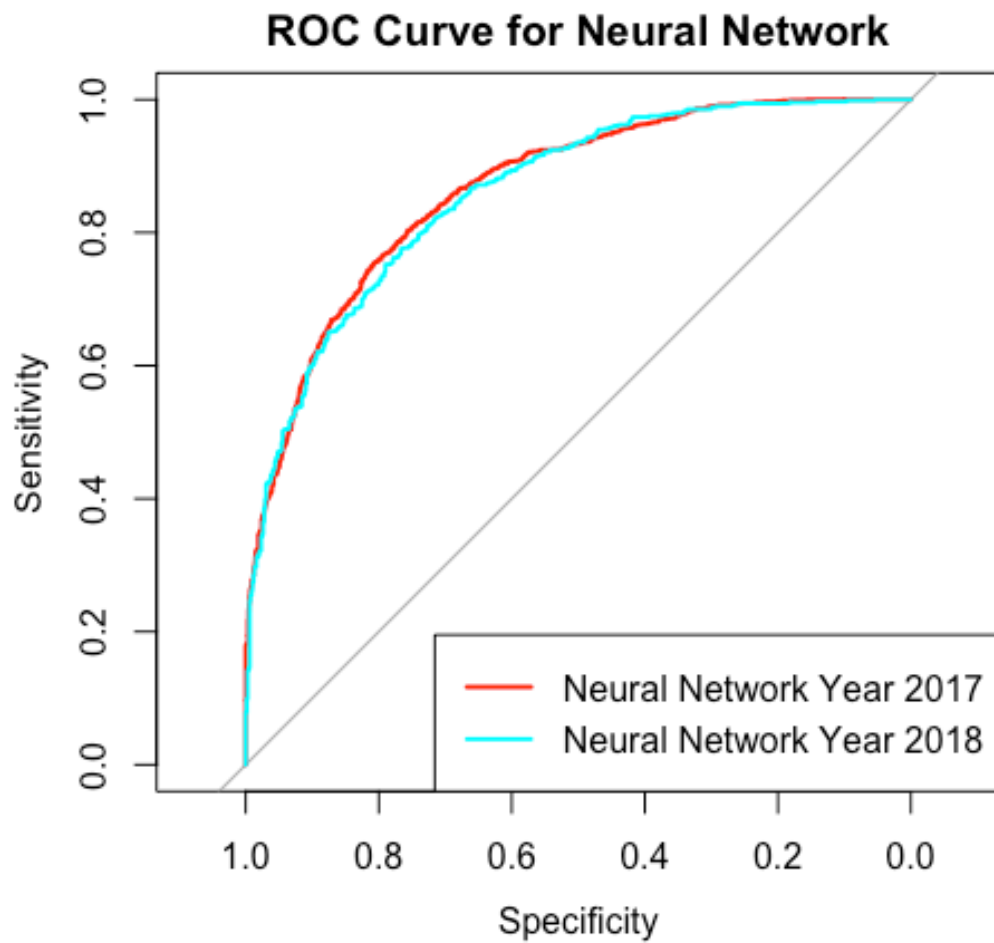

```
Neural Network Results:
         Year  Accuracy Precision   Recall        F1        AUC
Accuracy  2017 0.7850946 0.7839749 0.7870662 0.7855175 0.8671826
Accuracy1 2018 0.7660350 0.7641100 0.7696793 0.7668845 0.8624786
```

2 Comparison Algorithm

Comparison of five algorithms on Accuracy, Precision, Recall, F1-Score, AUC and ROC curves:

Year	Algorithm	Accuracy	Precision	Recall	F1-Score	AUC
2017	*Logistic Regression*	0.7871	0.7871	0.7871	0.7871	0.8704
	Random Forest	0.7709	0.7748	0.7650	0.7695	0.8535
	SVM	0.7871	0.7869	0.7871	0.7871	0.8664
	Naive Bayes	0.7689	0.7689	0.7689	0.7689	0.8533
	Neural Network	0.7863	0.7840	0.7823	0.7854	0.8671

Year	Algorithm	Accuracy	Precision	Recall	F1-Score	AUC
2018	*Logistic Regression*	0.7741	0.7741	0.7741	0.7741	0.8647
	Random Forest	0.7434	0.7456	0.7478	0.7446	0.8376
	SVM	0.7697	0.7697	0.7697	0.7697	0.8562
	Naive Bayes	0.7522	0.7522	0.7522	0.7522	0.8498
	Neural Network	0.7726	0.7641	0.7697	0.7719	0.8604

Accuaracy

Precision

Recall

F1-Score

ROC Curves for All Algorithms

AUC

	2017	2018
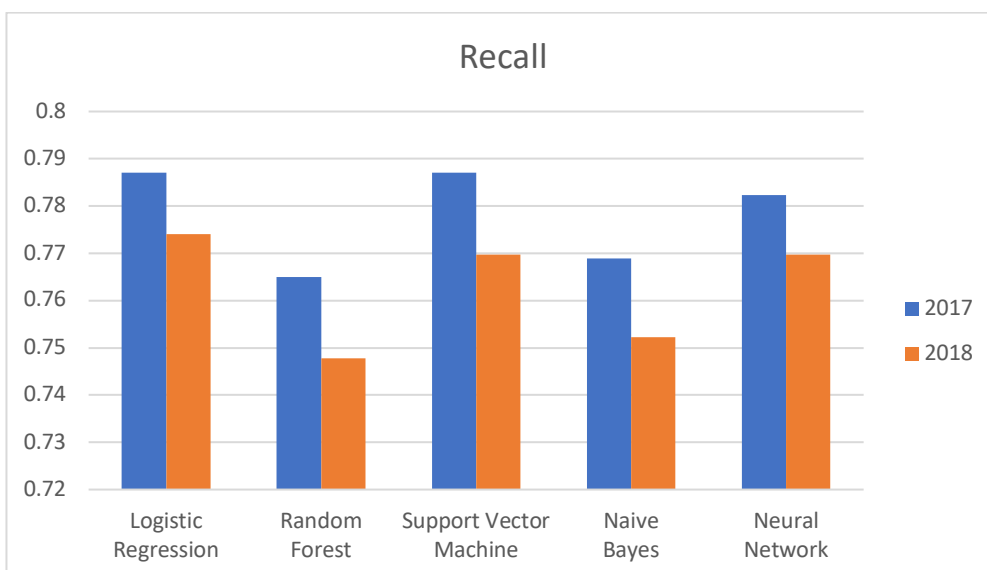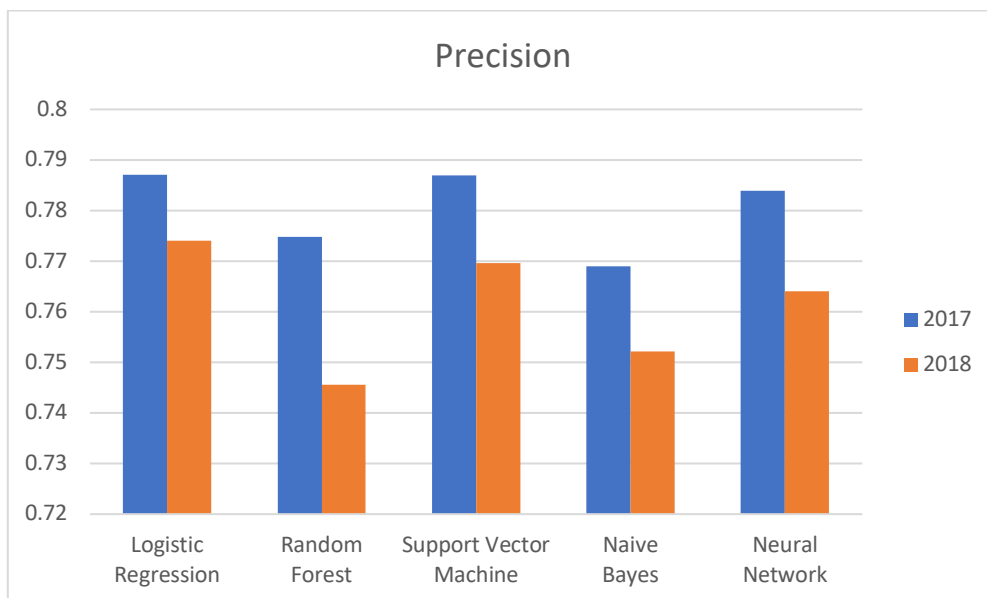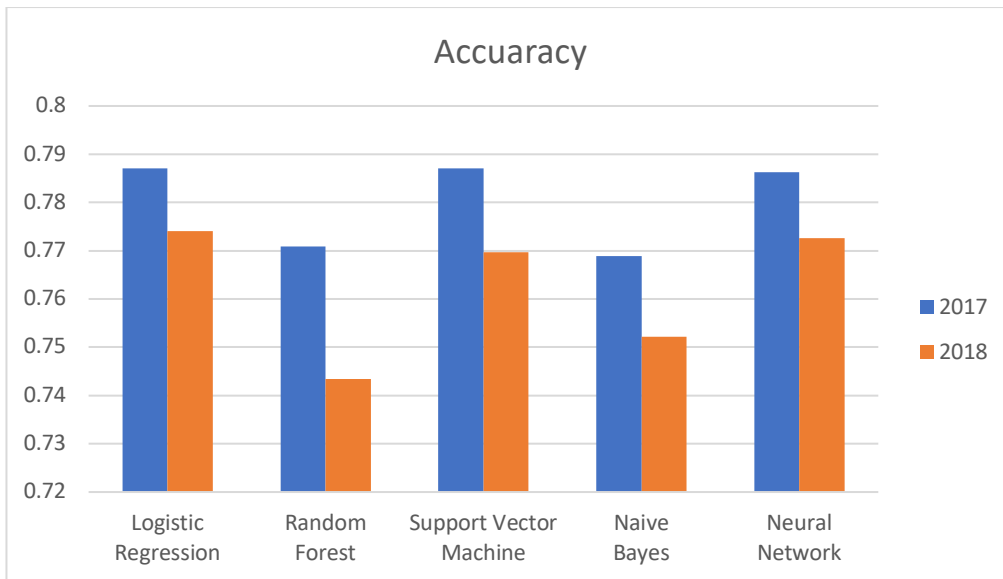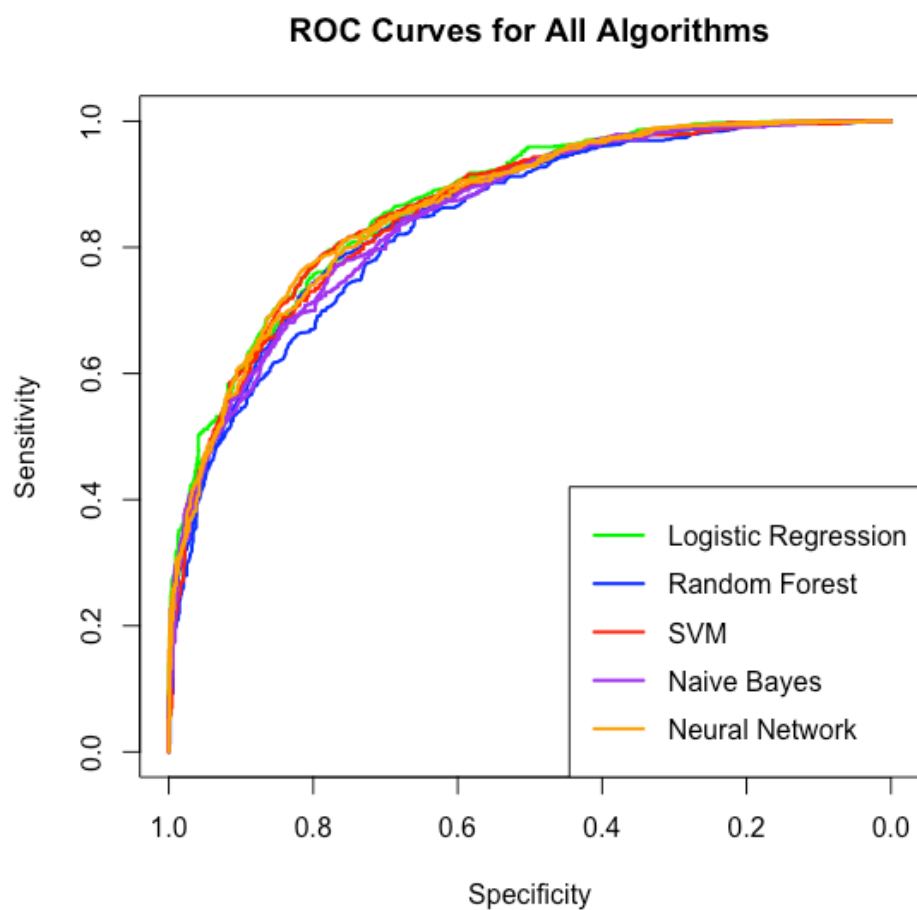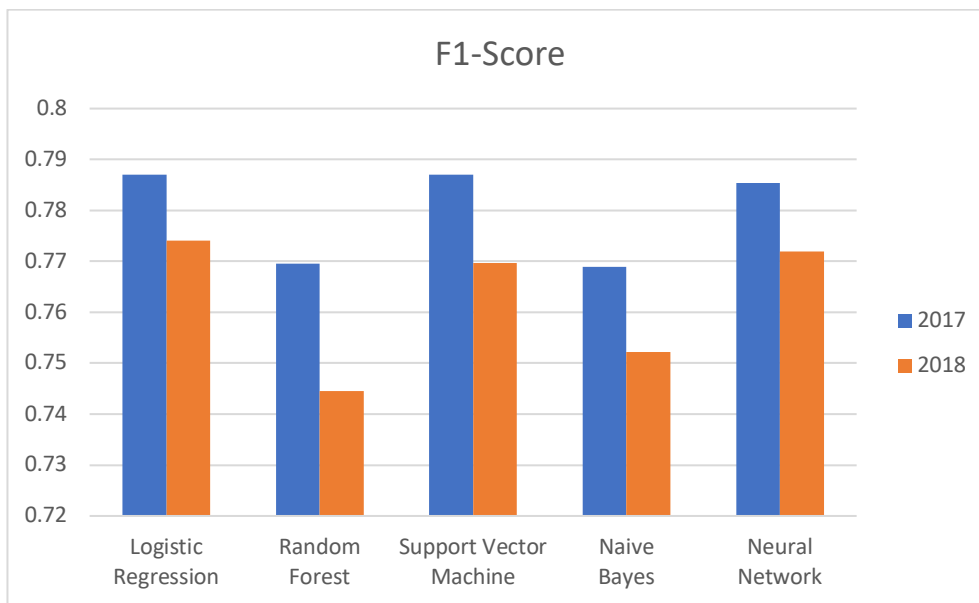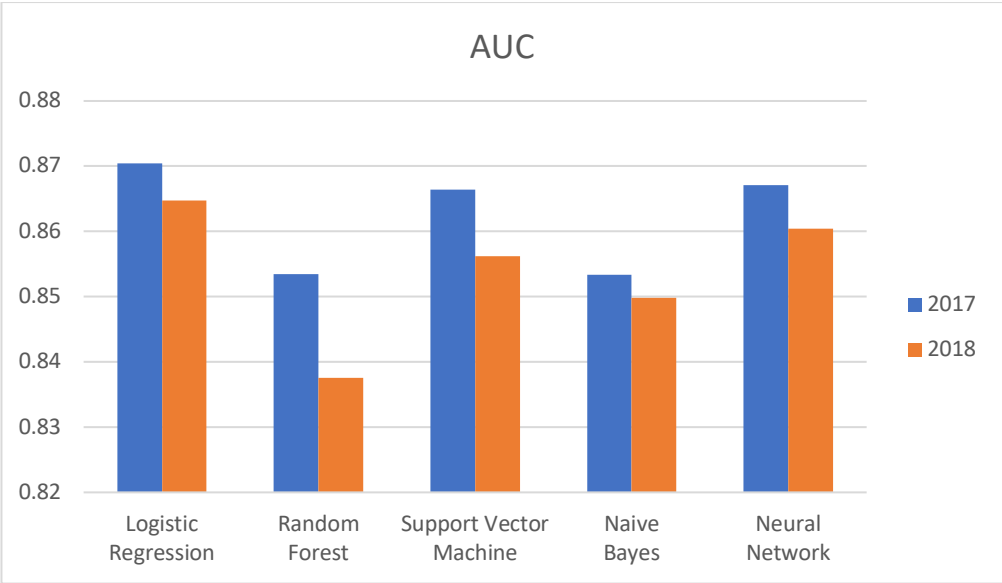
3 Stacking

3.1 Running Results of Predict the 2017 Competition

```
[1] "Confusion Matrix for Stacking (Year: 2017 )"
Confusion Matrix and Statistics

          Reference
Prediction Win Loss
      Win  987  285
      Loss 281  983

               Accuracy : 0.7768
                 95% CI : (0.7601, 0.7929)
    No Information Rate : 0.5
    P-Value [Acc > NIR] : <2e-16

                  Kappa : 0.5536

 Mcnemar's Test P-Value : 0.8997

            Sensitivity : 0.7784
            Specificity : 0.7752
         Pos Pred Value : 0.7759
         Neg Pred Value : 0.7777
             Prevalence : 0.5000
         Detection Rate : 0.3892
   Detection Prevalence : 0.5016
      Balanced Accuracy : 0.7768

       'Positive' Class : Win
```

3.2 Running Results of Predict the 2018 Competition

```
[1] "Confusion Matrix for Stacking (Year: 2018 )"
Confusion Matrix and Statistics

          Reference
Prediction Win Loss
      Win  518  164
      Loss 168  522

               Accuracy : 0.758
                 95% CI : (0.7345, 0.7805)
    No Information Rate : 0.5
    P-Value [Acc > NIR] : <2e-16

                  Kappa : 0.516

 Mcnemar's Test P-Value : 0.8692

            Sensitivity : 0.7551
            Specificity : 0.7609
         Pos Pred Value : 0.7595
         Neg Pred Value : 0.7565
             Prevalence : 0.5000
         Detection Rate : 0.3776
   Detection Prevalence : 0.4971
      Balanced Accuracy : 0.7580

       'Positive' Class : Win
```
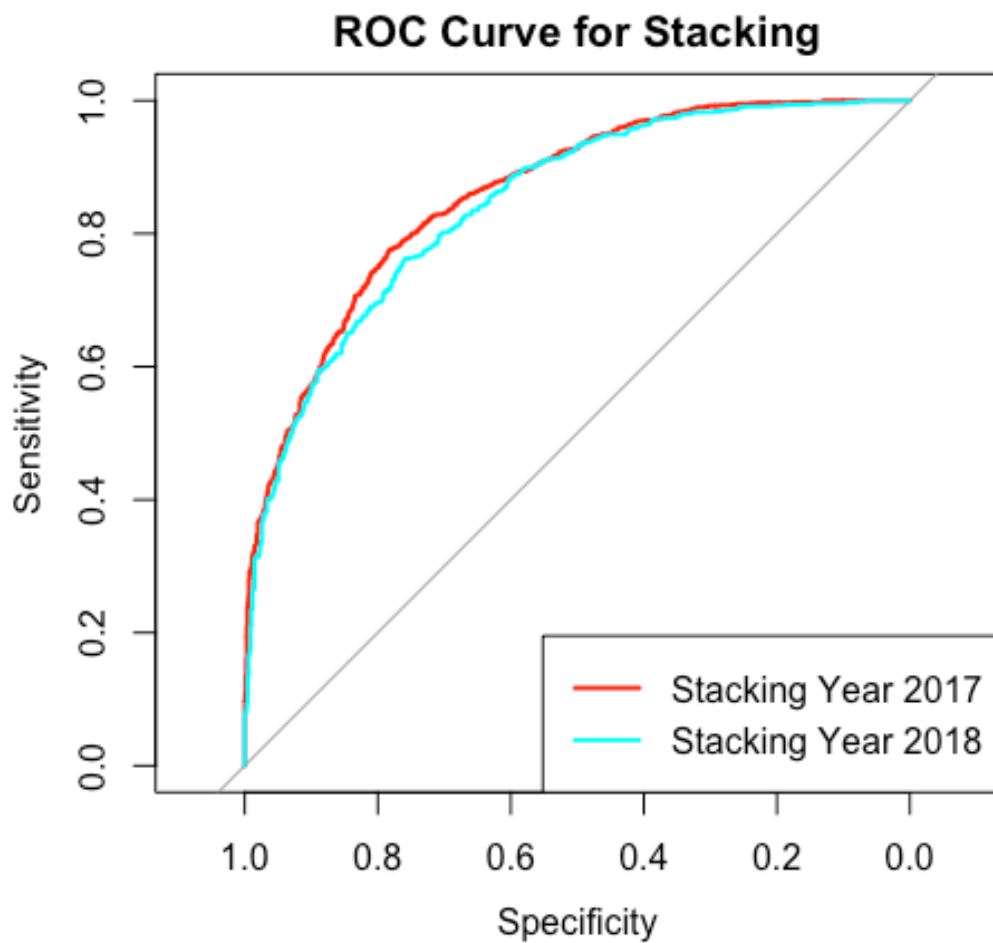
3.3 Running Results of Predict the 2017 Competition

ROC Curve for Stacking


```
Stacking Results:
          Year  Accuracy Precision    Recall        F1       AUC
Accuracy  2017 0.7803628 0.7810277 0.7791798 0.7801026 0.8602503
Accuracy1 2018 0.7572886 0.7561684 0.7594752 0.7578182 0.8458402
```

4 Compare the integrated algorithm with Random Forest and Naive Bayes

Year	Algorithm	Accuracy	Precision	Recall	F1-Score	AUC
2017	*Random Forest*	0.7709	0.7748	0.7650	0.7695	0.8535
	Naive Bayes	0.7689	0.7689	0.7689	0.7689	0.8533
	Stacking	0.7776	0.7810	0.7784	0.7778	0.8619

Year	Algorithm	Accuracy	Precision	Recall	F1-Score	AUC
2018	*Random Forest*	0.7434	0.7456	0.7478	0.7446	0.8376
	Naive Bayes	0.7522	0.7522	0.7522	0.7522	0.8498
	Stacking	0.7566	0.7562	0.7595	0.7573	0.8471

Precision

Recall

F1-Score

AUC

	Random Forest	Naive Bayes	Stacking
2017			
2018			