NAVEEN BALARAJU
MACULAR DEGENERATION – A PILOT STUDY

Case Vs Control

```
Console  ~/Desktop/project/
Call:
 randomForest(formula = ME_case_control ~ ., data = age_data,        importance = TRUE, subset = train)
               Type of random forest: classification
                     Number of trees: 500
No. of variables tried at each split: 1

        OOB estimate of  error rate: 3.33%
Confusion matrix:
        case control class.error
case      16       0  0.00000000
control    1      13  0.07142857
>
> pred.case.control = predict(rf.case.control, newdata=age_data[-train ,])
> resp.test=age_data[-train ,"ME_case_control"]
> caret::confusionMatrix(resp.test,pred.case.control)
Confusion Matrix and Statistics

          Reference
Prediction case control
   case      11       0
   control    0      10

               Accuracy : 1
                 95% CI : (0.8389, 1)
    No Information Rate : 0.5238
    P-Value [Acc > NIR] : 1.267e-06

                  Kappa : 1
 Mcnemar's Test P-Value : NA

            Sensitivity : 1.0000
            Specificity : 1.0000
         Pos Pred Value : 1.0000
         Neg Pred Value : 1.0000
             Prevalence : 0.5238
         Detection Rate : 0.5238
   Detection Prevalence : 0.5238
      Balanced Accuracy : 1.0000
```

intermediate, control and advanced

```
Call:
 randomForest(formula = AMD_ini_level_WS_c ~ ., data = age_data,        importance = TRUE, subset = train)
                Type of random forest: classification
                      Number of trees: 500
No. of variables tried at each split: 1

        OOB estimate of  error rate: 13.33%
Confusion matrix:
                Advanced AMD Intermediate AMD None or early AMD class.error
Advanced AMD              5               3                 0       0.375
Intermediate AMD         1               7                 0       0.125
None or early AMD        0               0                14       0.000
>
> pred.ICA = predict(rf.ICA, newdata=age_data[-train ,])
> resp.ICA=age_data[-train ,"AMD_ini_level_WS_c"]
>
> caret::confusionMatrix(resp.ICA,pred.ICA)
Confusion Matrix and Statistics

                  Reference
Prediction          Advanced AMD Intermediate AMD None or early AMD
  Advanced AMD               2                0                 0
  Intermediate AMD           4                5                 0
  None or early AMD          0                0                10

Overall Statistics

               Accuracy : 0.8095
                 95% CI : (0.5809, 0.9455)
    No Information Rate : 0.4762
    P-Value [Acc > NIR] : 0.00185

                  Kappa : 0.7042
 Mcnemar's Test P-Value : NA
```

```
Statistics by Class:
```

	Class: Advanced AMD	Class: Intermediate AMD	Class: None or early AMD
Sensitivity	0.33333	1.0000	1.0000
Specificity	1.00000	0.7500	1.0000
Pos Pred Value	1.00000	0.5556	1.0000
Neg Pred Value	0.78947	1.0000	1.0000
Prevalence	0.28571	0.2381	0.4762
Detection Rate	0.09524	0.2381	0.4762
Detection Prevalence	0.09524	0.4286	0.4762
Balanced Accuracy	0.66667	0.8750	1.0000

Grouping control and intermediate into one class and advanced to other class

```
Call:
 randomForest(formula = AMD_ini_level_WS_c ~ ., data = age_data,      importance = TRUE, subset = train)
               Type of random forest: classification
                     Number of trees: 500
No. of variables tried at each split: 1

       OOB estimate of  error rate: 23.33%
Confusion matrix:
  1  2 class.error
1 2  4   0.6666667
2 3 21   0.1250000
>
> pred.CIvsA = predict(rf.CIvsA, newdata=age_data[-train ,])
> resp.CIvsA=age_data[-train ,"AMD_ini_level_WS_c"]
> caret::confusionMatrix(resp.CIvsA,pred.CIvsA)
Confusion Matrix and Statistics

          Reference
Prediction  1  2
         1  2  1
         2  1 16

              Accuracy : 0.9
                95% CI : (0.683, 0.9877)
   No Information Rate : 0.85
   P-Value [Acc > NIR] : 0.4049

                 Kappa : 0.6078
 Mcnemar's Test P-Value : 1.0000

           Sensitivity : 0.6667
           Specificity : 0.9412
        Pos Pred Value : 0.6667
        Neg Pred Value : 0.9412
            Prevalence : 0.1500
        Detection Rate : 0.1000
  Detection Prevalence : 0.1500
     Balanced Accuracy : 0.8039
```

Control vs Advanced

```
Call:
 randomForest(formula = AMD_ini_level_WS_c ~ ., data = dummy_data,      importance = TRUE, subset = train)
               Type of random forest: classification
                     Number of trees: 500
No. of variables tried at each split: 1

        OOB estimate of  error rate: 10%
Confusion matrix:
                  Advanced AMD None or early AMD class.error
Advanced AMD                 6                1  0.14285714
None or early AMD            1               12  0.07692308
>
> pred.CvsA = predict(rf.CvsA, newdata=dummy_data[-train ,])
> resp.CvsA=dummy_data[-train ,"AMD_ini_level_WS_c"]
> caret::confusionMatrix(resp.CvsA,pred.CvsA)
Confusion Matrix and Statistics

                   Reference
Prediction           Advanced AMD None or early AMD
  Advanced AMD                  3                0
  None or early AMD             0               12

               Accuracy : 1
                 95% CI : (0.782, 1)
    No Information Rate : 0.8
    P-Value [Acc > NIR] : 0.03518

                  Kappa : 1
 Mcnemar's Test P-Value : NA

            Sensitivity : 1.0
            Specificity : 1.0
         Pos Pred Value : 1.0
         Neg Pred Value : 1.0
             Prevalence : 0.2
         Detection Rate : 0.2
   Detection Prevalence : 0.2
      Balanced Accuracy : 1.0
```