Practical Machine Learning Final Report: Exercise Prediction

I have downloaded the file, reading it then creating partition.

Reading and creating partition in the File

```
library(caret)
```

```
## Loading required package: lattice
```

```
## Loading required package: ggplot2
```

```
library(randomForest)
```

```
## randomForest 4.6-12
```

```
## Type rfNews() to see new features/changes/bug fixes.
```

```
##
## Attaching package: 'randomForest'
```

```
## The following object is masked from 'package:ggplot2':
##
##     margin
```

```
library(rpart)
train_in <- read.csv("pml-training.csv")
validation <- read.csv('pml-testing.csv', header=T)
set.seed(127)
training_sample <- createDataPartition(y=train_in$classe, p=0.7, list=FALSE)
training <- train_in[training_sample, ]
testing <- train_in[-training_sample, ]
dim(training)
```

```
## [1] 13737   160
```

```
dim(testing)
```

```
## [1] 5885  160
```

Removing near 0 variables and then removing variable

```
NZV <- nearZeroVar(training)
training <- training[, -NZV]
testing  <- testing[, -NZV]
dim(training)
```

```
## [1] 13737   109
```

```
dim(testing)
```

```
## [1] 5885  109
```

```
#Now removing variables that are mostly NA
AllNA    <- sapply(training, function(x) mean(is.na(x))) > 0.95
training <- training[, AllNA==FALSE]
testing  <- testing[, AllNA==FALSE]
dim(training)
```

```
## [1] 13737    59
```

```
dim(testing)
```

```
## [1] 5885   59
```

Model building

The three model types I'm going to test are:

1.Random forest decision trees (rf) 2.Decision trees with CART (rpart) 3.Method: Generalized Boosted Model

1. Random forest decision trees (rf)

```
 controlRF <- trainControl(method="cv", number=3, verboseIter=FALSE)
    modFitRandForest <- train(classe ~ ., data=training, method="rf",trControl=controlRF)
    predictRandForest <- predict(modFitRandForest, newdata=testing)
    confMatRandForest <- confusionMatrix(predictRandForest, testing$classe)
    confMatRandForest
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    A    B    C    D    E
##          A 1674    0    0    0    0
##          B    0 1139    0    0    0
##          C    0    0 1026    0    0
##          D    0    0    0  964    0
##          E    0    0    0    0 1082
##
## Overall Statistics
##
##                Accuracy : 1
##                  95% CI : (0.9994, 1)
##     No Information Rate : 0.2845
##     P-Value [Acc > NIR] : < 2.2e-16
##
##                   Kappa : 1
##  Mcnemar's Test P-Value : NA
##
## Statistics by Class:
##
##                      Class: A Class: B Class: C Class: D Class: E
## Sensitivity            1.0000   1.0000   1.0000   1.0000   1.0000
## Specificity            1.0000   1.0000   1.0000   1.0000   1.0000
## Pos Pred Value         1.0000   1.0000   1.0000   1.0000   1.0000
## Neg Pred Value         1.0000   1.0000   1.0000   1.0000   1.0000
## Prevalence             0.2845   0.1935   0.1743   0.1638   0.1839
## Detection Rate         0.2845   0.1935   0.1743   0.1638   0.1839
## Detection Prevalence   0.2845   0.1935   0.1743   0.1638   0.1839
## Balanced Accuracy      1.0000   1.0000   1.0000   1.0000   1.0000
```

2. Decision trees

```
modFitDecTree <- rpart(classe ~ ., data=training, method="class")
predictDecTree <- predict(modFitDecTree, newdata=testing, type="class")
confMatDecTree <- confusionMatrix(predictDecTree, testing$classe)
confMatDecTree
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    A    B    C    D    E
##          A 1673    0    0    0    0
##          B    1 1138    0    0    0
##          C    0    1 1025    0    0
##          D    0    0    1  964    0
##          E    0    0    0    0 1082
##
## Overall Statistics
##
##                Accuracy : 0.9995
##                  95% CI : (0.9985, 0.9999)
##     No Information Rate : 0.2845
##     P-Value [Acc > NIR] : < 2.2e-16
##
##                   Kappa : 0.9994
##  Mcnemar's Test P-Value : NA
##
## Statistics by Class:
##
##                      Class: A Class: B Class: C Class: D Class: E
## Sensitivity            0.9994   0.9991   0.9990   1.0000   1.0000
## Specificity            1.0000   0.9998   0.9998   0.9998   1.0000
## Pos Pred Value         1.0000   0.9991   0.9990   0.9990   1.0000
## Neg Pred Value         0.9998   0.9998   0.9998   1.0000   1.0000
## Prevalence             0.2845   0.1935   0.1743   0.1638   0.1839
## Detection Rate         0.2843   0.1934   0.1742   0.1638   0.1839
## Detection Prevalence   0.2843   0.1935   0.1743   0.1640   0.1839
## Balanced Accuracy      0.9997   0.9995   0.9994   0.9999   1.0000
```

3. Method: Generalized Boosted Model

```
controlGBM <- trainControl(method = "repeatedcv", number = 5, repeats = 1)
modFitGBM  <- train(classe ~ ., data=training, method = "gbm",trControl = controlGBM, verbose = FALSE)
```

```
## Loading required package: gbm
```

```
## Loading required package: survival
```

```
##
## Attaching package: 'survival'
```

```
## The following object is masked from 'package:caret':
##
##     cluster
```

```
## Loading required package: splines
```

```
## Loading required package: parallel
```

```
## Loaded gbm 2.1.1
```

```
## Loading required package: plyr
```

```
modFitGBM$finalModel
```

```
## A gradient boosted model with multinomial loss function.
## 50 iterations were performed.
## There were 80 predictors of which 1 had non-zero influence.
```

```
predictGBM <- predict(modFitGBM, newdata=testing)
confMatGBM <- confusionMatrix(predictGBM, testing$classe)
confMatGBM
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    A    B    C    D    E
##          A 1673    0    0    0    0
##          B    1 1138    0    0    0
##          C    0    1 1025    0    0
##          D    0    0    1  964    0
##          E    0    0    0    0 1082
##
## Overall Statistics
##
##                Accuracy : 0.9995
##                  95% CI : (0.9985, 0.9999)
##     No Information Rate : 0.2845
##     P-Value [Acc > NIR] : < 2.2e-16
##
##                   Kappa : 0.9994
##  Mcnemar's Test P-Value : NA
##
## Statistics by Class:
##
##                      Class: A Class: B Class: C Class: D Class: E
## Sensitivity            0.9994   0.9991   0.9990   1.0000   1.0000
## Specificity            1.0000   0.9998   0.9998   0.9998   1.0000
## Pos Pred Value         1.0000   0.9991   0.9990   0.9990   1.0000
## Neg Pred Value         0.9998   0.9998   0.9998   1.0000   1.0000
## Prevalence             0.2845   0.1935   0.1743   0.1638   0.1839
## Detection Rate         0.2843   0.1934   0.1742   0.1638   0.1839
## Detection Prevalence   0.2843   0.1935   0.1743   0.1640   0.1839
## Balanced Accuracy      0.9997   0.9995   0.9994   0.9999   1.0000
```

Model Assessment (Out of sample error)

The accuracy of the 3 regression modeling methods above are:

Random Forest : 1 Decision Tree : 0.9995 GBM : 0.9995 In that case, the Random Forest model will be applied to predict the 20 quiz results (testing dataset) as shown below.

```
predictTEST <- predict(modFitRandForest, newdata=testing)
predictTEST
```

```
##    [1] A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
##   [35] A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
##   [69] A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
##  [103] A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
##  [137] A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
##  [171] A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
##  [205] A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
##  [239] A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
##  [273] A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
##  [307] A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
##  [341] A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
##  [375] A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
##  [409] A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
##  [443] A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
##  [477] A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
##  [511] A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
##  [545] A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
##  [579] A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
##  [613] A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
##  [647] A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
##  [681] A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
##  [715] A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
##  [749] A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
##  [783] A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
##  [817] A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
##  [851] A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
##  [885] A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
##  [919] A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
##  [953] A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
##  [987] A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
## [1021] A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
```

```
## [1055] A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
## [1089] A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
## [1123] A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
## [1157] A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
## [1191] A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
## [1225] A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
## [1259] A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
## [1293] A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
## [1327] A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
## [1361] A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
## [1395] A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
## [1429] A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
## [1463] A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
## [1497] A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
## [1531] A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
## [1565] A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
## [1599] A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
## [1633] A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
## [1667] A A A A A A A A B B B B B B B B B B B B B B B B B B B B B B B B B
## [1701] B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B
## [1735] B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B
## [1769] B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B
## [1803] B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B
## [1837] B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B
## [1871] B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B
## [1905] B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B
## [1939] B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B
## [1973] B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B
## [2007] B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B
## [2041] B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B
## [2075] B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B
## [2109] B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B
## [2143] B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B
## [2177] B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B
## [2211] B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B
## [2245] B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B
## [2279] B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B
## [2313] B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B
## [2347] B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B
## [2381] B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B
## [2415] B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B
## [2449] B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B
## [2483] B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B
## [2517] B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B
## [2551] B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B
## [2585] B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B
## [2619] B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B
## [2653] B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B
## [2687] B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B
## [2721] B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B
## [2755] B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B
## [2789] B B B B B B B B B B B B B B B B B B B B B B B B B C C C C C C C C C
## [2823] C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C
## [2857] C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C
## [2891] C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C
## [2925] C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C
## [2959] C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C
## [2993] C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C
## [3027] C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C
## [3061] C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C
## [3095] C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C
## [3129] C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C
## [3163] C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C
## [3197] C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C
## [3231] C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C
## [3265] C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C
## [3299] C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C
## [3333] C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C
## [3367] C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C
## [3401] C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C
## [3435] C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C
## [3469] C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C
## [3503] C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C
## [3537] C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C
## [3571] C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C
## [3605] C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C
```

```
## [3639] C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C
## [3673] C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C
## [3707] C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C
## [3741] C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C
## [3775] C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C
## [3809] C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C D D D
## [3843] D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D
## [3877] D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D
## [3911] D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D
## [3945] D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D
## [3979] D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D
## [4013] D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D
## [4047] D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D
## [4081] D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D
## [4115] D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D
## [4149] D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D
## [4183] D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D
## [4217] D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D
## [4251] D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D
## [4285] D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D
## [4319] D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D
## [4353] D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D
## [4387] D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D
## [4421] D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D
## [4455] D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D
## [4489] D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D
## [4523] D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D
## [4557] D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D
## [4591] D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D
## [4625] D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D
## [4659] D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D
## [4693] D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D
## [4727] D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D
## [4761] D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D
## [4795] D D D D D D D D D E E E E E E E E E E E E E E E E E E E E E E E E E
## [4829] E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E
## [4863] E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E
## [4897] E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E
## [4931] E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E
## [4965] E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E
## [4999] E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E
## [5033] E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E
## [5067] E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E
## [5101] E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E
## [5135] E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E
## [5169] E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E
## [5203] E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E
## [5237] E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E
## [5271] E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E
## [5305] E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E
## [5339] E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E
## [5373] E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E
## [5407] E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E
## [5441] E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E
## [5475] E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E
## [5509] E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E
## [5543] E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E
## [5577] E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E
## [5611] E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E
## [5645] E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E
## [5679] E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E
## [5713] E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E
## [5747] E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E
## [5781] E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E
## [5815] E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E
## [5849] E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E E
## [5883] E E E
## Levels: A B C D E
```