Practical Machine Learning Assignment

First, we install the required packages and load library.

```
# Install packages and load library
install.packages('randomForest')
library('randomForest')
install.packages('caret')
library('caret')
install.packages('e1071')
library('e1071')
```

Next, we read the data and remove the NA value.

```
# Read data frame
a=read.csv('pml-training.csv',na.strings=c('','NA'))
b=a[,!apply(a,2,function(x) any(is.na(x)) )]
c=b[,-c(1:7)]
```

We split the data into two group with the ratio 60% train and 40% test

```
#For cross validation, split our testing data into sub groups 60:40

subGrps=createDataPartition(y=c$classe, p=0.6, list=FALSE)
subTraining=c[subGrps,]
subTesting=c[-subGrps, ]
dim(subTraining);dim(subTesting)

[1] 11776    53
[1] 7846   53
```

We next build the model using the training data. Here we use random forest algorithm. When the model is built, we continue to predict the outcome using test data and check the confusion matrix.

```
# Build prediction model based on random forest paradigm


model=randomForest(classe~., data=subTraining, method='class')

pred=predict(model,subTesting, type='class')

z=confusionMatrix(pred,subTesting$classe)

save(z,file='test.RData')


# Load data

load('test.RData')

z$table
```

The result is good with the accuracy over 99%. So we will use this model to test on the test data.

Prediction	Reference A	B	C	D	E
A	2229	20	0	0	0
B	0	1495	4	0	0
C	3	3	1363	19	0
D	0	0	1	1266	5
E	0	0	0	1	1437

Accuracy

0.9928626

```
# Check the accuracy of model


z$overall[1]
```

We load the test data and remove the NA values.

```
# Analyze the test data set
# Read data frame
d=read.csv('pml-testing.csv',na.strings=c('','NA'))
e=d[,!apply(d,2,function(x) any(is.na(x)) )]
f=e[,-c(1:7)]
```

Run the prediction using the random forest model.

```
# Predict using our model

predicted=predict(model,f,type='class')
save(predicted,file='predicted.RData')
```

The result is printed out.

```
# Check the result

load('predicted.RData')

# Display results
print(predicted)
```

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
B A B A A E D B A A B C B A E E A B B B
Levels: A B C D E