prediction_assignment.Rmd

Noel Namai

Sunday, July 20, 2014

Function to read data:

```r
readData <- function(file){
    data <- read.csv(file,
                     header=TRUE,
                     na.strings = c("NA", "", "#DIV/0!"),
                     stringsAsFactors=FALSE,
                     sep=",")
    data
    }
```

Function to clean data:

```r
cleanData <- function(df) {
    #df <- df[complete.cases(df),]
    df <- df[,-1]
    df <- df[sapply(df, function(df) !any(is.na(df)))]
    df
    }
```

Function to create features:

```r
creatFeatures <- function(df) {
    df$cvtd_timestamp <- strptime(df$cvtd_timestamp, "%m/%d/%Y %H:%M")
    df$weekday <- weekdays(df$cvtd_timestamp)
    df$month <- months(df$cvtd_timestamp)
    df$time <- format.Date(df$cvtd_timestamp, format="%H")
    df$ampm <- format.Date(df$cvtd_timestamp, format="%p")
    df
    }
```

Read training data:

```r
training <- readData("./data/training.csv")
dim(training)
```

```
## [1] 19622    160
```

Clean data:

```r
training <- cleanData(training)
training <- creatFeatures(training)
dim(training)
```

```
## [1] 19622     63
```

```r
library(caret)
```

```
## Warning: package 'caret' was built under R version 3.1.1
```

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

```r
#fit <- train(classe ~ ., method="rpart", data=training)
#fit
```