Prediction Assignment - Machine Learning

Femi Munatuz Zufa

## Introduction

This project aims to predict the manner in which weight lifting exercises are performed using sensor data. The classe variable is the target, and use machine learning techniques to classify it.

## Load Libraries

library(caret)  
library(randomForest)

library(readxl)train\_data <- read\_excel("pml-training.xlsx")

test\_data <- read\_excel("pml-testing.xlsx")

## Load Data

# Download and load data  
train\_url <- "https://d396qusza40orc.cloudfront.net/predmachlearn/pml-training.csv"  
test\_url <- "https://d396qusza40orc.cloudfront.net/predmachlearn/pml-testing.csv"  
  
train\_data <- read.csv(train\_url, na.strings = c("NA", "", "#DIV/0!"))  
test\_data <- read.csv(test\_url, na.strings = c("NA", "", "#DIV/0!"))

## Data Preprocessing

# Remove columns with too many NA values  
train\_data <- train\_data[, colSums(is.na(train\_data)) == 0]  
test\_data <- test\_data[, colSums(is.na(test\_data)) == 0]  
  
# Remove non-predictive columns  
train\_data <- train\_data[, -(1:7)]  
test\_data <- test\_data[, -(1:7)]

## Model Training & Evaluation

set.seed(123)  
  
# Split data into training and validation set  
trainIndex <- createDataPartition(train\_data$classe, p = 0.7, list = FALSE)  
train\_set <- train\_data[trainIndex, ]  
valid\_set <- train\_data[-trainIndex, ]  
  
# Train model using Random Forest  
model\_rf <- train(classe ~ ., data = train\_set, method = "rf", trControl = trainControl(method = "cv", number = 5))  
  
# Evaluate model  
predictions <- predict(model\_rf, valid\_set)  
confusionMatrix(predictions, valid\_set$classe)

## Final Predictions

```r final\_predictions <- predict(model\_rf, test\_data) final\_predictions