

prediction-RandomForest.R

mamarisg

Thu Jun 23 16:26:21 2016

```
library(randomForest)
```

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

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

```
library(ggplot2)
```

```
##
```

```
## Attaching package: 'ggplot2'
```

```
## The following object is masked from 'package:randomForest':
```

```
##
```

```
##      margin
```

```
dirpath <- "~/Doctorate/svm-gpuperf/"  
setwd(paste(dirpath, sep=""))
```

```
gpus <- read.table("./R-code/deviceInfo.csv", sep="," , header=T)  
NoGPU <- dim(gpus)[1]
```

```
apps <- c("matMul_gpu_uncoalesced", "matMul_gpu", "matMul_gpu_sharedmem_uncoalesced", "matMul_gpu_sharedmem",  
         "matrix_sum_normal", "matrix_sum_coalesced",  
         "dotProd", "vectorAdd", "subSeqMax")
```

```
Parameters <- c("gpu_name", "gpu_id", "AppName", "AppId", "Input.Size", "Duration",  
               "max_clock_rate", "num_of_cores",  
               "Achieved.Occupancy",  
               "totalLoadGM", "totalStoreGM", "totalLoadSM", "totalStoreSM",  
               "Floating.Point.Operations.Single.Precision.",  
               "L2.Read.Transactions", "L2.Write.Transactions",  
               "blockSize", "GridSize", "totalThreads"  
)
```

```
DataAppGPU <- read.csv(file = paste("./R-code/Datasets/CleanData/App-GPU-CC-All.csv", sep = ""))  
DataAppGPU <- rbind(DataAppGPU[c(Parameters)])
```

```
result <- data.frame()
```

```
for (CC in c(1:10)){  
  for( j in 1:9) {  
  
    if (j <= 4){  
      Data <- subset(DataAppGPU, AppId == j & Input.Size >= 4096 )  
    } else if (j > 4 & j < 7) {
```

```

    Data <- subset(DataAppGPU, AppId == j & Input.Size >= 4096 )
  } else{
    Data <- subset(DataAppGPU, AppId == j & Input.Size >= 71303168)
  }

  if (j == 3 | j == 4 | j == 9){
    print(j)
  } else {
    Data$totalLoadSM <- NULL
    Data$totalStoreSM <- NULL
  }

Data <- Data[complete.cases(Data),]
# Data[["max_clock_rate"]] <- scale(Data[["max_clock_rate"]], center = FALSE, scale = max(Data[

trainingSet <- subset(Data, gpu_id != CC)
testSet <- subset(Data, gpu_id == CC )

dim(Data)
dim(trainingSet)
dim(testSet)

trainingSet$AppName <- NULL
trainingSet$gpu_name <- NULL
trainingSet$AppId <- NULL
trainingSet$gpu_id <- NULL

trainingSet$max_clock_rate <- NULL
trainingSet$num_of_cores <- NULL
trainingSet$Achieved.Occupancy <- NULL
trainingSet$blockSize <- NULL
# trainingSet$GridSize <- NULL
# trainingSet$totalThreads <- NULL
trainingSet$inst_issued2 <- NULL
trainingSet$L2.Read.Transactions <- NULL
trainingSet$L2.Write.Transactions <- NULL
trainingSet$totalStoreGM <- NULL

TestDuration <- testSet["Duration"]
Size <- testSet["Input.Size"]
App <- testSet["AppName"]
Gpu <- testSet["gpu_name"]
Block <- testSet["blockSize"]

testSet$AppName <- NULL
testSet$gpu_name <- NULL
testSet$Duration <- NULL
testSet$AppId <- NULL
testSet$gpu_id <- NULL

testSet$max_clock_rate <- NULL
testSet$num_of_cores <- NULL

```

```

testSet$Achieved.Occupancy <- NULL
testSet$blockSize <- NULL
# testSet$GridSize <- NULL
# testSet$totalThreads <- NULL
testSet$inst_issued2 <- NULL
testSet$L2.Read.Transactions <- NULL
testSet$L2.Write.Transactions <- NULL
testSet$totalStoreGM <- NULL

fit <- randomForest(trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
print(fit)
summary(fit)
predictions <- predict(fit, testSet)

mse <- mean((as.matrix(TestDuration) - predictions)^2)
mae <- mean(abs(as.matrix(TestDuration) - predictions))
mape <- mean(abs(as.matrix(TestDuration) - predictions)/predictions)
# mpe <- mean(as.matrix(TestDuration) - predictions/predictions)
# smape = mean((abs(as.matrix(predictions) - TestDuration) / (abs(TestDuration) + abs(predictions))))

Acc <- predictions/TestDuration
AccMin <- min(Acc)
AccMean <- mean(as.matrix(Acc))
AccMedian <- median(as.matrix(Acc))
AccMax <- max(Acc)
AccSD <- sd(as.matrix(Acc))

Tempresult <- data.frame(Gpu, App, Size, Block, TestDuration, predictions, Acc, AccMin, AccMax, AccSD)

result <- rbind(result, Tempresult)

}
}

```

```

##          |      Out-of-bag      |
## Tree |      MSE  %Var(y) |
## 100 |      5.277   27.11 |
## 200 |      5.307   27.26 |
## 300 |      5.19    26.66 |
## 400 |      5.181   26.62 |
## 500 |      5.179   26.60 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##              Type of random forest: regression
##              Number of trees: 500
## No. of variables tried at each split: 1
##
##              Mean of squared residuals: 5.178875
##              % Var explained: 73.4
##          |      Out-of-bag      |
## Tree |      MSE  %Var(y) |
## 100 |      0.7535   25.88 |

```

```

## 200 | 0.7671 26.35 |
## 300 | 0.7642 26.25 |
## 400 | 0.7435 25.53 |
## 500 | 0.7487 25.72 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 1
##
##           Mean of squared residuals: 0.7487151
##           % Var explained: 74.28
## [1] 3
##           | Out-of-bag |
## Tree | MSE %Var(y) |
## 100 | 5.162 21.52 |
## 200 | 5.028 20.97 |
## 300 | 4.887 20.38 |
## 400 | 4.984 20.78 |
## 500 | 4.833 20.15 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 2
##
##           Mean of squared residuals: 4.832846
##           % Var explained: 79.85
## [1] 4
##           | Out-of-bag |
## Tree | MSE %Var(y) |
## 100 | 0.3278 10.78 |
## 200 | 0.3662 12.05 |
## 300 | 0.3593 11.82 |
## 400 | 0.3791 12.47 |
## 500 | 0.374 12.30 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 2
##
##           Mean of squared residuals: 0.3740101
##           % Var explained: 87.7
##           | Out-of-bag |
## Tree | MSE %Var(y) |
## 100 | 0.000282 89.31 |
## 200 | 0.0002833 89.70 |
## 300 | 0.0002832 89.67 |
## 400 | 0.0002833 89.70 |
## 500 | 0.0002832 89.69 |

```

```

##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 1
##
##           Mean of squared residuals: 0.0002832491
##           % Var explained: 10.31
##           |           Out-of-bag           |
## Tree |           MSE %Var(y) |
## 100 | 3.767e-06   75.26 |
## 200 | 3.793e-06   75.78 |
## 300 | 3.802e-06   75.96 |
## 400 | 3.798e-06   75.89 |
## 500 | 3.801e-06   75.94 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 1
##
##           Mean of squared residuals: 3.800893e-06
##           % Var explained: 24.06
##           |           Out-of-bag           |
## Tree |           MSE %Var(y) |
## 100 | 2.32e-05    17.44 |
## 200 | 2.3e-05     17.30 |
## 300 | 2.29e-05    17.22 |
## 400 | 2.297e-05   17.27 |
## 500 | 2.293e-05   17.24 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 1
##
##           Mean of squared residuals: 2.293065e-05
##           % Var explained: 82.76
##           |           Out-of-bag           |
## Tree |           MSE %Var(y) |
## 100 | 8.862e-06    39.36 |
## 200 | 8.83e-06     39.21 |
## 300 | 8.818e-06    39.16 |
## 400 | 8.805e-06    39.11 |
## 500 | 8.81e-06     39.13 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 1

```

```

##
##           Mean of squared residuals: 8.810015e-06
##           % Var explained: 60.87
## [1] 9
##      |      Out-of-bag      |
## Tree |      MSE %Var(y) |
## 100 | 0.0001243      7.34 |
## 200 | 0.0001286      7.59 |
## 300 | 0.0001231      7.26 |
## 400 | 0.0001218      7.19 |
## 500 | 0.0001227      7.24 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 2
##
##           Mean of squared residuals: 0.0001227269
##           % Var explained: 92.76
##      |      Out-of-bag      |
## Tree |      MSE %Var(y) |
## 100 |      6.485      34.94 |
## 200 |      6.508      35.06 |
## 300 |      6.466      34.84 |
## 400 |      6.441      34.70 |
## 500 |      6.556      35.32 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 1
##
##           Mean of squared residuals: 6.556088
##           % Var explained: 64.68
##      |      Out-of-bag      |
## Tree |      MSE %Var(y) |
## 100 |      0.8342      29.08 |
## 200 |      0.8262      28.80 |
## 300 |      0.8077      28.16 |
## 400 |      0.8028      27.99 |
## 500 |      0.7936      27.67 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 1
##
##           Mean of squared residuals: 0.7936439
##           % Var explained: 72.33
## [1] 3
##      |      Out-of-bag      |

```

```

## Tree |      MSE %Var(y) |
## 100 |    4.881  19.74 |
## 200 |    4.64   18.76 |
## 300 |    4.513  18.25 |
## 400 |    4.496  18.18 |
## 500 |    4.58   18.52 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 2
##
##           Mean of squared residuals: 4.580476
##           % Var explained: 81.48
## [1] 4
##           |      Out-of-bag |
## Tree |      MSE %Var(y) |
## 100 |    0.3976  12.71 |
## 200 |    0.3794  12.13 |
## 300 |    0.3728  11.91 |
## 400 |    0.3854  12.32 |
## 500 |    0.4034  12.89 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 2
##
##           Mean of squared residuals: 0.4033807
##           % Var explained: 87.11
##           |      Out-of-bag |
## Tree |      MSE %Var(y) |
## 100 | 0.0005126  88.06 |
## 200 | 0.0005079  87.27 |
## 300 | 0.0005077  87.23 |
## 400 | 0.0005067  87.06 |
## 500 | 0.0005076  87.20 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 1
##
##           Mean of squared residuals: 0.0005075655
##           % Var explained: 12.8
##           |      Out-of-bag |
## Tree |      MSE %Var(y) |
## 100 | 3.938e-06  75.57 |
## 200 | 3.89e-06  74.64 |
## 300 | 3.89e-06  74.63 |
## 400 | 3.887e-06  74.57 |

```

```

## 500 | 3.887e-06    74.58 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 1
##
##           Mean of squared residuals: 3.886671e-06
##           % Var explained: 25.42
##           |           Out-of-bag           |
## Tree |           MSE %Var(y) |
## 100 | 3.322e-05    21.61 |
## 200 | 3.292e-05    21.42 |
## 300 | 3.294e-05    21.43 |
## 400 | 3.289e-05    21.40 |
## 500 | 3.288e-05    21.40 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 1
##
##           Mean of squared residuals: 3.288301e-05
##           % Var explained: 78.6
##           |           Out-of-bag           |
## Tree |           MSE %Var(y) |
## 100 | 9.85e-06     44.48 |
## 200 | 9.804e-06    44.28 |
## 300 | 9.803e-06    44.27 |
## 400 | 9.798e-06    44.25 |
## 500 | 9.82e-06     44.35 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 1
##
##           Mean of squared residuals: 9.820011e-06
##           % Var explained: 55.65
## [1] 9
##           |           Out-of-bag           |
## Tree |           MSE %Var(y) |
## 100 | 0.0001465     9.23 |
## 200 | 0.0001413     8.90 |
## 300 | 0.0001462     9.21 |
## 400 | 0.0001403     8.84 |
## 500 | 0.0001389     8.75 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression

```



```

##               Number of trees: 500
## No. of variables tried at each split: 2
##
##           Mean of squared residuals: 0.0001388934
##               % Var explained: 91.25
##      |      Out-of-bag      |
## Tree |      MSE %Var(y) |
## 100 |      5.228  29.33 |
## 200 |      5.146  28.87 |
## 300 |      5.153  28.91 |
## 400 |      5.157  28.93 |
## 500 |      5.056  28.37 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 1
##
##           Mean of squared residuals: 5.056033
##               % Var explained: 71.63
##      |      Out-of-bag      |
## Tree |      MSE %Var(y) |
## 100 |      0.8108  28.76 |
## 200 |      0.8375  29.71 |
## 300 |      0.7931  28.13 |
## 400 |      0.7732  27.43 |
## 500 |      0.7701  27.32 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 1
##
##           Mean of squared residuals: 0.7700665
##               % Var explained: 72.68
## [1] 3
##      |      Out-of-bag      |
## Tree |      MSE %Var(y) |
## 100 |      4.175  17.12 |
## 200 |      4.292  17.60 |
## 300 |      4.384  17.98 |
## 400 |      4.4    18.05 |
## 500 |      4.373  17.93 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 2
##
##           Mean of squared residuals: 4.372666
##               % Var explained: 82.07

```

```

## [1] 4
##      |      Out-of-bag      |
## Tree |      MSE %Var(y) |
## 100 |    0.3615    11.77 |
## 200 |    0.3842    12.51 |
## 300 |    0.3919    12.76 |
## 400 |    0.4029    13.12 |
## 500 |    0.4008    13.05 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##      Type of random forest: regression
##      Number of trees: 500
## No. of variables tried at each split: 2
##
##      Mean of squared residuals: 0.4008003
##      % Var explained: 86.95
##      |      Out-of-bag      |
## Tree |      MSE %Var(y) |
## 100 | 0.0005152    88.58 |
## 200 | 0.0005096    87.62 |
## 300 | 0.0005065    87.08 |
## 400 | 0.0005051    86.84 |
## 500 | 0.0005043    86.70 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##      Type of random forest: regression
##      Number of trees: 500
## No. of variables tried at each split: 1
##
##      Mean of squared residuals: 0.000504269
##      % Var explained: 13.3
##      |      Out-of-bag      |
## Tree |      MSE %Var(y) |
## 100 | 3.854e-06    75.40 |
## 200 | 3.83e-06    74.95 |
## 300 | 3.83e-06    74.94 |
## 400 | 3.837e-06    75.08 |
## 500 | 3.845e-06    75.23 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##      Type of random forest: regression
##      Number of trees: 500
## No. of variables tried at each split: 1
##
##      Mean of squared residuals: 3.844829e-06
##      % Var explained: 24.77
##      |      Out-of-bag      |
## Tree |      MSE %Var(y) |
## 100 | 2.86e-05    21.01 |
## 200 | 2.832e-05    20.80 |
## 300 | 2.811e-05    20.64 |

```

```

## 400 | 2.806e-05    20.61 |
## 500 | 2.8e-05     20.56 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 1
##
##           Mean of squared residuals: 2.79959e-05
##           % Var explained: 79.44
##           |           Out-of-bag           |
## Tree |           MSE %Var(y) |
## 100 | 7.523e-06    39.75 |
## 200 | 7.528e-06    39.77 |
## 300 | 7.517e-06    39.71 |
## 400 | 7.495e-06    39.60 |
## 500 | 7.505e-06    39.65 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 1
##
##           Mean of squared residuals: 7.504973e-06
##           % Var explained: 60.35
## [1] 9
##           |           Out-of-bag           |
## Tree |           MSE %Var(y) |
## 100 | 8.103e-05     5.56 |
## 200 | 8.329e-05     5.72 |
## 300 | 8.212e-05     5.64 |
## 400 | 8.374e-05     5.75 |
## 500 | 8.514e-05     5.85 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 2
##
##           Mean of squared residuals: 8.51422e-05
##           % Var explained: 94.15
##           |           Out-of-bag           |
## Tree |           MSE %Var(y) |
## 100 | 4.977    24.09 |
## 200 | 4.72     22.84 |
## 300 | 4.719    22.84 |
## 400 | 4.702    22.75 |
## 500 | 4.747    22.97 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)

```

```

##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 1
##
##           Mean of squared residuals: 4.746846
##           % Var explained: 77.03
##           |           Out-of-bag           |
## Tree |           MSE %Var(y) |
## 100 |           0.6219 21.17 |
## 200 |           0.5992 20.40 |
## 300 |           0.6128 20.86 |
## 400 |           0.6393 21.76 |
## 500 |           0.6497 22.12 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 1
##
##           Mean of squared residuals: 0.6497476
##           % Var explained: 77.88
## [1] 3
##           |           Out-of-bag           |
## Tree |           MSE %Var(y) |
## 100 |           4.538 18.49 |
## 200 |           4.532 18.46 |
## 300 |           4.061 16.54 |
## 400 |           4.324 17.62 |
## 500 |           4.141 16.87 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 2
##
##           Mean of squared residuals: 4.140968
##           % Var explained: 83.13
## [1] 4
##           |           Out-of-bag           |
## Tree |           MSE %Var(y) |
## 100 |           0.4211 13.57 |
## 200 |           0.3851 12.41 |
## 300 |           0.3789 12.21 |
## 400 |           0.3756 12.10 |
## 500 |           0.3669 11.82 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 2
##

```

```

##           Mean of squared residuals: 0.3669106
##           % Var explained: 88.18
##           |           Out-of-bag           |
## Tree |           MSE %Var(y) |
## 100 | 0.0005066      87.38 |
## 200 | 0.0005015      86.51 |
## 300 | 0.0005017      86.54 |
## 400 | 0.0005017      86.54 |
## 500 | 0.0005009      86.41 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 1
##
##           Mean of squared residuals: 0.0005009194
##           % Var explained: 13.59
##           |           Out-of-bag           |
## Tree |           MSE %Var(y) |
## 100 | 3.738e-06      72.29 |
## 200 | 3.748e-06      72.49 |
## 300 | 3.737e-06      72.28 |
## 400 | 3.758e-06      72.70 |
## 500 | 3.757e-06      72.66 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 1
##
##           Mean of squared residuals: 3.756727e-06
##           % Var explained: 27.34
##           |           Out-of-bag           |
## Tree |           MSE %Var(y) |
## 100 | 2.224e-05      13.87 |
## 200 | 2.196e-05      13.70 |
## 300 | 2.203e-05      13.74 |
## 400 | 2.204e-05      13.75 |
## 500 | 2.206e-05      13.76 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 1
##
##           Mean of squared residuals: 2.2063e-05
##           % Var explained: 86.24
##           |           Out-of-bag           |
## Tree |           MSE %Var(y) |
## 100 | 8.037e-06      35.65 |
## 200 | 7.99e-06       35.44 |

```

```

## 300 | 7.971e-06    35.36 |
## 400 | 7.988e-06    35.43 |
## 500 | 7.981e-06    35.40 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 1
##
##           Mean of squared residuals: 7.981421e-06
##           % Var explained: 64.6
## [1] 9
##           |           Out-of-bag           |
## Tree |           MSE %Var(y) |
## 100 | 0.0001155    6.79 |
## 200 | 0.0001107    6.51 |
## 300 | 0.0001124    6.61 |
## 400 | 0.0001148    6.75 |
## 500 | 0.0001128    6.63 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 2
##
##           Mean of squared residuals: 0.0001127875
##           % Var explained: 93.37
##           |           Out-of-bag           |
## Tree |           MSE %Var(y) |
## 100 |      6.69    35.09 |
## 200 |     6.526    34.22 |
## 300 |     6.525    34.22 |
## 400 |     6.524    34.21 |
## 500 |     6.505    34.12 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 1
##
##           Mean of squared residuals: 6.505455
##           % Var explained: 65.88
##           |           Out-of-bag           |
## Tree |           MSE %Var(y) |
## 100 |    0.8397    28.87 |
## 200 |    0.8311    28.57 |
## 300 |    0.8468    29.12 |
## 400 |    0.8332    28.65 |
## 500 |    0.8475    29.14 |
##
## Call:

```

```

## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace=TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 1
##
##           Mean of squared residuals: 0.8475134
##           % Var explained: 70.86
## [1] 3
##      |      Out-of-bag      |
## Tree |      MSE %Var(y) |
## 100 |      4.363  17.64 |
## 200 |      4.518  18.26 |
## 300 |      4.583  18.53 |
## 400 |      4.699  19.00 |
## 500 |      4.725  19.10 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace=TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 2
##
##           Mean of squared residuals: 4.724535
##           % Var explained: 80.9
## [1] 4
##      |      Out-of-bag      |
## Tree |      MSE %Var(y) |
## 100 |      0.4528  14.45 |
## 200 |      0.4206  13.42 |
## 300 |      0.42    13.41 |
## 400 |      0.421   13.44 |
## 500 |      0.4302  13.73 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace=TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 2
##
##           Mean of squared residuals: 0.4302369
##           % Var explained: 86.27
##      |      Out-of-bag      |
## Tree |      MSE %Var(y) |
## 100 | 0.0005018  86.50 |
## 200 | 0.0005045  86.97 |
## 300 | 0.0005031  86.73 |
## 400 | 0.0005028  86.68 |
## 500 | 0.0005026  86.63 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace=TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 1

```

```

##
##           Mean of squared residuals: 0.0005025538
##           % Var explained: 13.37
##           |           Out-of-bag           |
## Tree |           MSE %Var(y) |
## 100 | 3.788e-06    72.97 |
## 200 | 3.793e-06    73.06 |
## 300 | 3.79e-06     73.00 |
## 400 | 3.795e-06    73.10 |
## 500 | 3.796e-06    73.11 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 1
##
##           Mean of squared residuals: 3.795982e-06
##           % Var explained: 26.89
##           |           Out-of-bag           |
## Tree |           MSE %Var(y) |
## 100 | 3.245e-05    20.97 |
## 200 | 3.225e-05    20.84 |
## 300 | 3.211e-05    20.75 |
## 400 | 3.207e-05    20.72 |
## 500 | 3.208e-05    20.73 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 1
##
##           Mean of squared residuals: 3.207868e-05
##           % Var explained: 79.27
##           |           Out-of-bag           |
## Tree |           MSE %Var(y) |
## 100 | 8.674e-06    38.49 |
## 200 | 8.718e-06    38.68 |
## 300 | 8.697e-06    38.59 |
## 400 | 8.707e-06    38.64 |
## 500 | 8.717e-06    38.68 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 1
##
##           Mean of squared residuals: 8.716979e-06
##           % Var explained: 61.32
## [1] 9
##           |           Out-of-bag           |
## Tree |           MSE %Var(y) |

```



```

## 100 | 0.0001367      8.14 |
## 200 | 0.0001312      7.82 |
## 300 | 0.000134      7.98 |
## 400 | 0.0001327      7.90 |
## 500 | 0.0001343      8.00 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 2
##
##           Mean of squared residuals: 0.000134324
##           % Var explained: 92
##           |           Out-of-bag           |
## Tree |           MSE %Var(y) |
## 100 |           5.559 27.87 |
## 200 |           5.534 27.75 |
## 300 |           5.503 27.59 |
## 400 |           5.659 28.37 |
## 500 |           5.703 28.59 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 1
##
##           Mean of squared residuals: 5.703179
##           % Var explained: 71.41
##           |           Out-of-bag           |
## Tree |           MSE %Var(y) |
## 100 |           0.7267 25.15 |
## 200 |           0.7532 26.06 |
## 300 |           0.7677 26.57 |
## 400 |           0.7715 26.70 |
## 500 |           0.7854 27.18 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 1
##
##           Mean of squared residuals: 0.7853514
##           % Var explained: 72.82
## [1] 3
##           |           Out-of-bag           |
## Tree |           MSE %Var(y) |
## 100 |           4.712 19.80 |
## 200 |           4.294 18.04 |
## 300 |           4.381 18.41 |
## 400 |           4.375 18.38 |
## 500 |           4.409 18.53 |

```

```

##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 2
##
##           Mean of squared residuals: 4.408619
##           % Var explained: 81.47
## [1] 4
##           |           Out-of-bag           |
## Tree |           MSE %Var(y) |
## 100 |    0.4835    16.02 |
## 200 |    0.4495    14.89 |
## 300 |    0.4328    14.34 |
## 400 |    0.4239    14.04 |
## 500 |    0.4299    14.24 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 2
##
##           Mean of squared residuals: 0.4299408
##           % Var explained: 85.76
##           |           Out-of-bag           |
## Tree |           MSE %Var(y) |
## 100 | 0.0004121    96.05 |
## 200 | 0.0004036    94.07 |
## 300 | 0.0004036    94.08 |
## 400 | 0.0004036    94.09 |
## 500 | 0.0004038    94.12 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 1
##
##           Mean of squared residuals: 0.000403779
##           % Var explained: 5.88
##           |           Out-of-bag           |
## Tree |           MSE %Var(y) |
## 100 | 4.038e-06    76.54 |
## 200 | 4.02e-06    76.21 |
## 300 | 4.012e-06    76.05 |
## 400 | 4.018e-06    76.17 |
## 500 | 4.013e-06    76.07 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500

```

```

## No. of variables tried at each split: 1
##
##           Mean of squared residuals: 4.012603e-06
##           % Var explained: 23.93
##           |           Out-of-bag           |
## Tree |           MSE %Var(y) |
## 100 | 3.368e-05    22.23 |
## 200 | 3.339e-05    22.04 |
## 300 | 3.34e-05     22.05 |
## 400 | 3.331e-05    21.99 |
## 500 | 3.33e-05     21.98 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 1
##
##           Mean of squared residuals: 3.329671e-05
##           % Var explained: 78.02
##           |           Out-of-bag           |
## Tree |           MSE %Var(y) |
## 100 | 9.064e-06    43.11 |
## 200 | 9.004e-06    42.82 |
## 300 | 9.046e-06    43.02 |
## 400 | 9.036e-06    42.97 |
## 500 | 9.019e-06    42.89 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 1
##
##           Mean of squared residuals: 9.019213e-06
##           % Var explained: 57.11
## [1] 9
##           |           Out-of-bag           |
## Tree |           MSE %Var(y) |
## 100 | 0.0001164     7.45 |
## 200 | 0.0001194     7.64 |
## 300 | 0.000121     7.74 |
## 400 | 0.0001228     7.85 |
## 500 | 0.0001218     7.79 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 2
##
##           Mean of squared residuals: 0.0001218271
##           % Var explained: 92.21
##           |           Out-of-bag           |

```

```

## Tree |      MSE  %Var(y) |
## 100 |    4.877   25.27 |
## 200 |    4.905   25.41 |
## 300 |    4.859   25.17 |
## 400 |    4.858   25.17 |
## 500 |    4.899   25.38 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 1
##
##           Mean of squared residuals: 4.898909
##           % Var explained: 74.62
##           |      Out-of-bag      |
## Tree |      MSE  %Var(y) |
## 100 |    0.3308   16.60 |
## 200 |    0.3312   16.62 |
## 300 |    0.3286   16.49 |
## 400 |    0.3287   16.49 |
## 500 |    0.3251   16.31 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 1
##
##           Mean of squared residuals: 0.3251026
##           % Var explained: 83.69
## [1] 3
##           |      Out-of-bag      |
## Tree |      MSE  %Var(y) |
## 100 |    0.65   18.59 |
## 200 |   0.6433   18.40 |
## 300 |   0.6407   18.33 |
## 400 |   0.6444   18.43 |
## 500 |   0.646   18.48 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 2
##
##           Mean of squared residuals: 0.6460076
##           % Var explained: 81.52
## [1] 4
##           |      Out-of-bag      |
## Tree |      MSE  %Var(y) |
## 100 |   0.07357    8.86 |
## 200 |   0.0727    8.76 |
## 300 |   0.07315    8.81 |

```

```

## 400 | 0.07272      8.76 |
## 500 | 0.07245      8.73 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 2
##
##           Mean of squared residuals: 0.07244703
##           % Var explained: 91.27
##           |           Out-of-bag           |
## Tree |           MSE %Var(y) |
## 100 | 0.0004631      87.25 |
## 200 | 0.0004647      87.54 |
## 300 | 0.0004635      87.32 |
## 400 | 0.0004644      87.48 |
## 500 | 0.0004636      87.33 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 1
##
##           Mean of squared residuals: 0.000463565
##           % Var explained: 12.67
##           |           Out-of-bag           |
## Tree |           MSE %Var(y) |
## 100 | 6.479e-07      27.84 |
## 200 | 6.459e-07      27.76 |
## 300 | 6.414e-07      27.56 |
## 400 | 6.406e-07      27.53 |
## 500 | 6.397e-07      27.49 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 1
##
##           Mean of squared residuals: 6.396504e-07
##           % Var explained: 72.51
##           |           Out-of-bag           |
## Tree |           MSE %Var(y) |
## 100 | 2.902e-05      19.22 |
## 200 | 2.884e-05      19.10 |
## 300 | 2.896e-05      19.18 |
## 400 | 2.898e-05      19.20 |
## 500 | 2.898e-05      19.20 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression

```

```

##                               Number of trees: 500
## No. of variables tried at each split: 1
##
##           Mean of squared residuals: 2.898286e-05
##           % Var explained: 80.8
##      |      Out-of-bag      |
## Tree |      MSE %Var(y) |
## 100 | 8.084e-06   36.87 |
## 200 | 8.033e-06   36.64 |
## 300 | 8.018e-06   36.57 |
## 400 | 8.001e-06   36.50 |
## 500 | 7.986e-06   36.43 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 1
##
##           Mean of squared residuals: 7.986045e-06
##           % Var explained: 63.57
## [1] 9
##      |      Out-of-bag      |
## Tree |      MSE %Var(y) |
## 100 | 0.0001253    7.23 |
## 200 | 0.0001245    7.18 |
## 300 | 0.0001236    7.13 |
## 400 | 0.0001223    7.06 |
## 500 | 0.0001226    7.07 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 2
##
##           Mean of squared residuals: 0.0001225953
##           % Var explained: 92.93
##      |      Out-of-bag      |
## Tree |      MSE %Var(y) |
## 100 |   5.874   32.14 |
## 200 |   5.691   31.14 |
## 300 |   5.691   31.14 |
## 400 |   5.721   31.30 |
## 500 |   5.792   31.69 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 1
##
##           Mean of squared residuals: 5.792012
##           % Var explained: 68.31

```

```

##      |      Out-of-bag      |
## Tree |      MSE %Var(y) |
## 100 |    0.7856    27.36 |
## 200 |    0.7992    27.83 |
## 300 |    0.8054    28.05 |
## 400 |    0.809     28.17 |
## 500 |    0.7889    27.47 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 1
##
##           Mean of squared residuals: 0.7889336
##           % Var explained: 72.53
## [1] 3
##      |      Out-of-bag      |
## Tree |      MSE %Var(y) |
## 100 |    5.304    20.49 |
## 200 |    5.333    20.60 |
## 300 |    5.101    19.70 |
## 400 |    5.021    19.40 |
## 500 |    4.999    19.31 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 2
##
##           Mean of squared residuals: 4.998727
##           % Var explained: 80.69
## [1] 4
##      |      Out-of-bag      |
## Tree |      MSE %Var(y) |
## 100 |    0.537    16.78 |
## 200 |    0.5046    15.77 |
## 300 |    0.5103    15.95 |
## 400 |    0.5038    15.74 |
## 500 |    0.5086    15.90 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 2
##
##           Mean of squared residuals: 0.5086253
##           % Var explained: 84.1
##      |      Out-of-bag      |
## Tree |      MSE %Var(y) |
## 100 | 0.0005046    86.33 |
## 200 | 0.0005067    86.69 |

```

```

## 300 | 0.0005094    87.16 |
## 400 | 0.0005111    87.45 |
## 500 | 0.0005107    87.37 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##              Type of random forest: regression
##              Number of trees: 500
## No. of variables tried at each split: 1
##
##              Mean of squared residuals: 0.0005106515
##              % Var explained: 12.63
##              |      Out-of-bag      |
## Tree |      MSE %Var(y) |
## 100 | 3.3e-06  64.76 |
## 200 | 3.285e-06  64.46 |
## 300 | 3.29e-06  64.55 |
## 400 | 3.282e-06  64.40 |
## 500 | 3.281e-06  64.38 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##              Type of random forest: regression
##              Number of trees: 500
## No. of variables tried at each split: 1
##
##              Mean of squared residuals: 3.28126e-06
##              % Var explained: 35.62
##              |      Out-of-bag      |
## Tree |      MSE %Var(y) |
## 100 | 2.76e-05  18.74 |
## 200 | 2.735e-05  18.57 |
## 300 | 2.735e-05  18.57 |
## 400 | 2.732e-05  18.55 |
## 500 | 2.736e-05  18.58 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##              Type of random forest: regression
##              Number of trees: 500
## No. of variables tried at each split: 1
##
##              Mean of squared residuals: 2.736238e-05
##              % Var explained: 81.42
##              |      Out-of-bag      |
## Tree |      MSE %Var(y) |
## 100 | 5.762e-06  27.08 |
## 200 | 5.722e-06  26.90 |
## 300 | 5.733e-06  26.95 |
## 400 | 5.737e-06  26.97 |
## 500 | 5.741e-06  26.99 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)

```



```

##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 1
##
##           Mean of squared residuals: 5.74112e-06
##           % Var explained: 73.01
## [1] 9
##      |      Out-of-bag      |
## Tree |      MSE %Var(y) |
## 100 | 0.0001319      8.34 |
## 200 | 0.0001362      8.62 |
## 300 | 0.0001359      8.60 |
## 400 | 0.0001334      8.44 |
## 500 | 0.0001343      8.50 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 2
##
##           Mean of squared residuals: 0.0001343171
##           % Var explained: 91.5
##      |      Out-of-bag      |
## Tree |      MSE %Var(y) |
## 100 |      6.497      33.37 |
## 200 |      6.723      34.54 |
## 300 |      6.557      33.68 |
## 400 |      6.547      33.63 |
## 500 |      6.513      33.45 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 1
##
##           Mean of squared residuals: 6.512585
##           % Var explained: 66.55
##      |      Out-of-bag      |
## Tree |      MSE %Var(y) |
## 100 |      0.9089      30.75 |
## 200 |      0.9292      31.43 |
## 300 |      0.9247      31.28 |
## 400 |      0.9144      30.93 |
## 500 |      0.9087      30.74 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 1
##
##           Mean of squared residuals: 0.9086776

```

```

##                               % Var explained: 69.26
## [1] 3
##      |      Out-of-bag      |
## Tree |      MSE %Var(y) |
## 100 |      5.848  22.54 |
## 200 |      5.481  21.12 |
## 300 |      5.275  20.33 |
## 400 |      5.148  19.84 |
## 500 |      5.182  19.97 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##              Type of random forest: regression
##              Number of trees: 500
## No. of variables tried at each split: 2
##
##              Mean of squared residuals: 5.182241
##              % Var explained: 80.03
## [1] 4
##      |      Out-of-bag      |
## Tree |      MSE %Var(y) |
## 100 |      0.4578  14.12 |
## 200 |      0.4581  14.13 |
## 300 |      0.4635  14.30 |
## 400 |      0.4699  14.50 |
## 500 |      0.4701  14.50 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##              Type of random forest: regression
##              Number of trees: 500
## No. of variables tried at each split: 2
##
##              Mean of squared residuals: 0.4700666
##              % Var explained: 85.5
##      |      Out-of-bag      |
## Tree |      MSE %Var(y) |
## 100 | 0.0005043  87.56 |
## 200 | 0.0005024  87.22 |
## 300 | 0.000502  87.16 |
## 400 | 0.0005031  87.34 |
## 500 | 0.0005023  87.21 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##              Type of random forest: regression
##              Number of trees: 500
## No. of variables tried at each split: 1
##
##              Mean of squared residuals: 0.0005023204
##              % Var explained: 12.79
##      |      Out-of-bag      |
## Tree |      MSE %Var(y) |
## 100 | 3.912e-06  74.84 |

```

```

## 200 | 3.911e-06    74.81 |
## 300 | 3.892e-06    74.45 |
## 400 | 3.906e-06    74.71 |
## 500 | 3.901e-06    74.62 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##              Type of random forest: regression
##              Number of trees: 500
## No. of variables tried at each split: 1
##
##              Mean of squared residuals: 3.900837e-06
##              % Var explained: 25.38
##      |      Out-of-bag      |
## Tree |      MSE %Var(y) |
## 100 | 3.318e-05    21.54 |
## 200 | 3.3e-05    21.42 |
## 300 | 3.306e-05    21.46 |
## 400 | 3.308e-05    21.48 |
## 500 | 3.306e-05    21.47 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##              Type of random forest: regression
##              Number of trees: 500
## No. of variables tried at each split: 1
##
##              Mean of squared residuals: 3.306354e-05
##              % Var explained: 78.53
##      |      Out-of-bag      |
## Tree |      MSE %Var(y) |
## 100 | 9.9e-06    43.98 |
## 200 | 9.922e-06    44.08 |
## 300 | 9.943e-06    44.18 |
## 400 | 9.958e-06    44.24 |
## 500 | 9.947e-06    44.19 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##              Type of random forest: regression
##              Number of trees: 500
## No. of variables tried at each split: 1
##
##              Mean of squared residuals: 9.947158e-06
##              % Var explained: 55.81
## [1] 9
##      |      Out-of-bag      |
## Tree |      MSE %Var(y) |
## 100 | 0.0001395    9.14 |
## 200 | 0.0001406    9.21 |
## 300 | 0.0001369    8.97 |
## 400 | 0.0001329    8.71 |
## 500 | 0.0001319    8.64 |
##

```

```

## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 2
##
##           Mean of squared residuals: 0.00013194
##           % Var explained: 91.36
##           |           Out-of-bag           |
## Tree |           MSE %Var(y) |
## 100 |           6.084 31.92 |
## 200 |           6.075 31.88 |
## 300 |           6.044 31.71 |
## 400 |           6.018 31.58 |
## 500 |           5.981 31.38 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 1
##
##           Mean of squared residuals: 5.980654
##           % Var explained: 68.62
##           |           Out-of-bag           |
## Tree |           MSE %Var(y) |
## 100 |           0.8028 26.95 |
## 200 |           0.8219 27.60 |
## 300 |           0.8353 28.04 |
## 400 |           0.8278 27.79 |
## 500 |           0.8269 27.76 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 1
##
##           Mean of squared residuals: 0.8269098
##           % Var explained: 72.24
## [1] 3
##           |           Out-of-bag           |
## Tree |           MSE %Var(y) |
## 100 |           6.457 24.87 |
## 200 |           5.786 22.28 |
## 300 |           5.543 21.35 |
## 400 |           5.409 20.83 |
## 500 |           5.297 20.40 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 2

```

```

##
##           Mean of squared residuals: 5.297189
##           % Var explained: 79.6
## [1] 4
##      |      Out-of-bag      |
## Tree |      MSE %Var(y) |
## 100 |      0.516   16.12 |
## 200 |      0.5069   15.84 |
## 300 |      0.5005   15.64 |
## 400 |      0.4758   14.87 |
## 500 |      0.4847   15.15 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 2
##
##           Mean of squared residuals: 0.4847101
##           % Var explained: 84.85
##      |      Out-of-bag      |
## Tree |      MSE %Var(y) |
## 100 | 0.0005129   88.61 |
## 200 | 0.0005108   88.26 |
## 300 | 0.0005107   88.25 |
## 400 | 0.0005074   87.67 |
## 500 | 0.000506    87.42 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 1
##
##           Mean of squared residuals: 0.0005059814
##           % Var explained: 12.58
##      |      Out-of-bag      |
## Tree |      MSE %Var(y) |
## 100 | 4.101e-06   78.40 |
## 200 | 4.107e-06   78.52 |
## 300 | 4.094e-06   78.26 |
## 400 | 4.071e-06   77.82 |
## 500 | 4.077e-06   77.94 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet, importance = TRUE, do.trace = TRUE)
##           Type of random forest: regression
##           Number of trees: 500
## No. of variables tried at each split: 1
##
##           Mean of squared residuals: 4.076772e-06
##           % Var explained: 22.06
##      |      Out-of-bag      |
## Tree |      MSE %Var(y) |

```

```

## 100 | 2.639e-05    16.48 |
## 200 | 2.614e-05    16.32 |
## 300 | 2.605e-05    16.27 |
## 400 | 2.601e-05    16.24 |
## 500 | 2.597e-05    16.22 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet,      importance = TRUE, do.trace = TRUE)
##              Type of random forest: regression
##              Number of trees: 500
## No. of variables tried at each split: 1
##
##              Mean of squared residuals: 2.596981e-05
##              % Var explained: 83.78
##      |      Out-of-bag      |
## Tree |      MSE %Var(y) |
## 100 | 8.164e-06    37.85 |
## 200 | 8.145e-06    37.76 |
## 300 | 8.122e-06    37.65 |
## 400 | 8.118e-06    37.64 |
## 500 | 8.112e-06    37.61 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet,      importance = TRUE, do.trace = TRUE)
##              Type of random forest: regression
##              Number of trees: 500
## No. of variables tried at each split: 1
##
##              Mean of squared residuals: 8.111995e-06
##              % Var explained: 62.39
## [1] 9
##      |      Out-of-bag      |
## Tree |      MSE %Var(y) |
## 100 | 0.0001456     9.45 |
## 200 | 0.0001483     9.63 |
## 300 | 0.0001464     9.50 |
## 400 | 0.0001445     9.38 |
## 500 | 0.0001411     9.16 |
##
## Call:
## randomForest(formula = trainingSet$Duration ~ ., data = trainingSet,      importance = TRUE, do.trace = TRUE)
##              Type of random forest: regression
##              Number of trees: 500
## No. of variables tried at each split: 2
##
##              Mean of squared residuals: 0.0001411076
##              % Var explained: 90.84

# result
colnames(result) <-c("Gpus", "Apps", "InputSize", "ThreadBlock" , "Measured", "Predicted", "accuracy",

Tempresult <- data.frame(Gpu, App, Size, Block, TestDuration, predictions, Acc, AccMin, AccMax, AccMean

```

```

result$Apps <- factor(result$Apps, levels = c("matMul_gpu_uncoalesced", "matMul_gpu", "matMul_gpu_shared",
      "matrix_sum_normal", "matrix_sum_coalesced",
      "dotProd", "vectorAdd", "subSeqMax"))

# result[result$Apps %in% "matrix_sum_normal" & result$Gpus %in% c("Quadro", "TitanX"),]

Graph <- ggplot(data=result, aes(x=Gpus, y=accuracy, group=Gpus, col=Gpus)) +
  geom_boxplot( size=1.5, outlier.size = 2.5) + # scale_y_continuous(limits = c(0, 6)) +
  stat_boxplot(geom = 'errorbar') +
  xlab("GPUs") +
  ggtitle("Random Forest") +
  ylab(expression(paste("Accuracy ", T[k]/T[m] ))) +
  theme(plot.title = element_text(family = "Times", face="bold", size=40)) +
  theme(axis.title = element_text(family = "Times", face="bold", size=30)) +
  theme(axis.text = element_text(family = "Times", face="bold", size=20, colour = "Black")) +
  theme(axis.text.x=element_blank()) +
  theme(legend.title = element_text(family = "Times", face="bold", size=0)) +
  theme(legend.text = element_text(family = "Times", face="bold", size=20)) +
  theme(legend.direction = "horizontal",
        legend.position = "bottom",
        legend.key=element_rect(size=5),
        legend.key.size = unit(5, "lines")) +
  # facet_grid(~Apps, scales="fixed")
  facet_wrap(~Apps, ncol=3, scales="free_y") +
  theme(strip.text = element_text(size=20))+
  scale_colour_grey()

ggsave(paste("./images/ResultsLearning/ResultRandomForest.pdf", sep=""), Graph, device = pdf, height=10,
write.csv(result, file = "./R-code/Results/RandomForest-MSCoalesced.csv")
# ggsave(paste("./images/ResultsLearning/ResultLinearRegression.png", sep=""), Graph, height=10, width=10)

```