

Knor Benchmark

Benchmarking Results From Knor package

Load Data

```
# Load libraries
library(knor)

## Loading required package: Rcpp
## Installed packages
installed.packages()[names(sessionInfo())$otherPkgs], "Version"]

##      knor      Rcpp
## "0.0-7" "1.0.6"
## R Version
version[['version.string']]

## [1] "R version 4.0.5 (2021-03-31)"
X_1m = as.matrix(read.csv("data_1m.csv", header = FALSE))
X_100k = as.matrix(read.csv("data_100k.csv", header = FALSE))
X_10k = as.matrix(read.csv("data_10k.csv", header = FALSE))
X_1k = as.matrix(read.csv("data_1k.csv", header = FALSE))
```

Elbow Method Speed Test

```
# Just a convenient function to test execution speed of the 'Elbow method'
test_multicore_speed <- function(x){

  rng <- 2:10
  iters <- rng * 0

  for (i in rng) {
    m <- knor::Kmeans(data = x, centers = i,
                      iter.max = 100000, nthread = -1,
                      init = "kmeanspp",
                      tolerance = 1e-06,
                      dist.type = "eucl")
    iters[i] = m$iters
  }

  return(iters)
}
```

```

check_run_time_speed <- function(x){
  # Replicate the number samples to get a mean time
  time.taken <- c(1:7)*0

  for (i in 1:7) {
    start.time <- Sys.time()

    test_multicore_speed(x)

    end.time <- Sys.time()

    time.taken[i] <- (as.numeric(end.time) - as.numeric(start.time))
  }
  return(summary(time.taken))
}

```

Check Speed For Each Sample

```
check_run_time_speed(X_1m)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  276.9   278.4   281.5   282.7   285.0   293.3
```

```
check_run_time_speed(X_100k)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   15.03   15.15   15.27   15.38   15.38   16.28
```

```
check_run_time_speed(X_10k)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.6947   0.7149   0.7280   0.7324   0.7521   0.7704
```

```
check_run_time_speed(X_1k)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.01558 0.01613 0.01619 0.01682 0.01671 0.02024
```