

Test

2023-04-30

```
# load required libraries
library(parallel)
library(dplyr)

##
## Attaching package: 'dplyr'
## The following objects are masked from 'package:stats':
##
##   filter, lag
## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union
library(stats)
library(knitr)
library(caret)

## Loading required package: ggplot2
## Loading required package: lattice
library(ggplot2)
library(foreach)
```

Load RDAs

```
#load(file = file.path(root_dir, "R_data", "X_wm.Rda"))
load(file = file.path(root_dir, "R_data", "X_gm.Rda"))
#load(file = file.path(root_dir, "R_data", "X_cb.Rda"))
```

PCA

```
# fnto perform PCA and save output
pca_gm <- perform_pca(X_gm)

## [1] "n components: 144"
save(pca_gm, file = file.path(root_dir, "R_data", "pca_gm.Rda"))

head(pca_gm$X_train_pca[,1:5])

##           PC1           PC2           PC3           PC4           PC5
## 1  -896945.8 1899996.5  979911.4 -836077.0 -535783.3
## 2 -1050540.5 1154776.8 1327196.5 -492493.3 -536841.4
## 4  -889088.5  509951.6  992029.1 -447768.1 -492607.2
```

```

## 5 -1001088.3 1131804.6 1195242.3 -402205.7 -477130.2
## 6 -133030.4 1427058.9 -737510.9 -569166.0 557344.5
## 7 -1135154.9 1215307.7 1242521.1 -492077.7 -487582.4
dim(pca_gm$X_train_pca)

## [1] 348 144
head(pca_gm$y_train)

## [1] 1 2 1 1 1 1
length(pca_gm$y_train)

## [1] 348
length(pca_gm$y_test)

## [1] 86
head(pca_gm$X_test_pca[,1:5])

##          PC1          PC2          PC3          PC4          PC5
## 3   -971569.6   803974.6   514470.4 -590613.3  -89693.94
## 10  1831789.6 -7494945.5 -2274209.1 3849322.7 -352171.89
## 15 -1008582.7  1686231.9  1934211.0 -465722.6 -232408.62
## 17 -1115262.1  1249268.3  1216732.9 -542212.1 -449832.03
## 18 -1062398.4  1153522.0  1231387.8 -518792.3 -401405.75
## 19 -1130824.0  1297966.8  1261840.8 -514473.4 -533188.94
dim(pca_gm$X_test_pca)

## [1] 86 144

```