CRASSMAT.R

nickkunz

2019-06-21

```r
## title: CRASSMAT
## description: conditional random sampling for sprase matrices
## author: nick kunz <nick.kunz@columbia.edu>
## created: 04/22/2019
## updated: 06/20/2019
## dependencies: svMisc

## create crassmat
crassmat <- function(sparse_matrix, test_split_thres, conditional) {

    ## duplicate original matrix (test set)
    sparse_matrix_copy <- sparse_matrix

    ## install / load progress tracking dependent library
    if (!require("svMisc"))
        install.packages("svMisc")

    ## conditional repeat
    repeat {

        ## conditional sampling for loop
        for (i in 0:nrow(sparse_matrix_copy)) {  ## loop through rows not na
            if (length(which(!is.na(sparse_matrix_copy[i, ]))) > conditional) {  ## conditional stateme

                ## conduct single sample removal for rows that meet conditional
                sparse_matrix_copy[i, ][sample(which(!is.na(sparse_matrix_copy[i, ])), size = 1)] <- NA

                ## conduct progress tracking by percentage
                svMisc::progress(i / nrow(sparse_matrix_copy) * 100)

                ## print prompt when completed
                if(i == nrow(sparse_matrix_copy))
                    cat("Complete!")

            } ## close matrix for loop conditional statement
        } ## close matrix for loop through rows not na

        ## stop repeat when sampling threshold is met
        if ((mean(is.na(sparse_matrix_copy)) - mean(is.na(sparse_matrix))) /
            mean(!is.na(sparse_matrix)) > test_split_thres) break

    } ## close conditional repeat

    ## return sampled matrix
    return(sparse_matrix_copy)

} ## close function
```

```
## end
```