Package 'MSML'

December 13, 2023

Title Model Selection Based on Machine Learning (ML)

Version 1.0.0.0

Description Models evaluation based on a modified version of the recursive feature elimination algorithm. This package is designed to determine the optimal model(s) by leveraging all available features. (Olkin and Finn, 1995) <doi:10.1037/0033-2909.118.1.155>, (DeLong et al., 1998) <doi:10.2307/2531595>, (Guyon et al., 2002) <doi:10.1023/A:1012487302797> and (Momin et al., 2

License GPL (>=3)

URL https://github.com/mommy003/MSML

Encoding UTF-8

Roxygen list(markdown = TRUE)

RoxygenNote 7.1.2

Depends R (>= 2.10)

Imports r2redux, R2ROC

LazyData true

R topics documented:

`data_test` *7 sets of PRSs for test dataset and target phenotype*

Description

A dataset containing 7 sets of PRSs for test dataset and target phenotype

Usage

```
data_test
```

Format

A data frame for test dataset:

V1 PRS1, for bin1

V2 PRS2, for bin1

V3 PRS3, for bin1

V4 PRS4, for bin1

V5 PRS5, for bin1

V6 PRS6, for bin1

V7 PRS7, for bin1

phenotype Target Phenotype, value

---

data_train *7 sets of PRSs for training data set and target phenotype*

Description

A dataset containing 7 sets of PRSs for training data set and target phenotype

Usage

```
data_train
```

Format

A data frame for training dataset:

V1 PRS1, for bin1

V2 PRS2, for bin1

V3 PRS3, for bin1

V4 PRS4, for bin1

V5 PRS5, for bin1

V6 PRS6, for bin1

V7 PRS7, for bin1

phenotype Target Phenotype, value

data_valid *7 sets of PRSs for validation dataset and target phenotype*

Description

A dataset containing 7 sets of PRSs for validation dataset and target phenotype

Usage

```
data_valid
```

Format

A data frame for validation dataset:

V1 PRS1, for bin1

V2 PRS2, for bin1

V3 PRS3, for bin1

V4 PRS4, for bin1

V5 PRS5, for bin1

V6 PRS6, for bin1

V7 PRS7, for bin1

phenotype Target Phenotype, value

model_configuration *model_configuration function*

Description

This function will generate features (e.g. PRSs) based on all possible combinations of model. The total number of models required to explore the combinations of these 'n' features can be calculated by summing the combinations for each possible number of features, ranging from 1 to 'n' (C(n,i)). where C(n,k) represents the binomial coefficient or "n choose k," with n denoting the total number of features and k indicating the number of features to include in each model.

Usage

```
model_configuration(data_train, data_valid, mv)
```

Arguments

data_train	This is the matrix for the training dataset
data_valid	This is the matrix for the validation dataset
mv	The total number of columns in data_train/data_valid

Value

This function will generate all possible model outcomes for validation and test dataset

Examples

```
## Not run:
data_train <- data_train
data_valid  <- data_valid
mv=8
out=model_configuration(data_train,data_valid,mv)
#This process will produce predicted values for the validation datasets,
#corresponding to each model configuration trained on the training dataset.
#The outcome of this function will yield variables named 'predict_validation'
#and 'total_model_configurations.
#To print the outcomes run out$predict_validation and out$total_model_configurations.
#For details (see https://github.com/mommy003/MSML).

## End(Not run)
```

---

model_evaluation *model_evaluation function*

Description

This function will identify the best model in the validation and test dataset.

Usage

```
model_evaluation(dat, mv, tn, prev, pthreshold = 0.05, method = "R2ROC")
```

Arguments

dat	This is the matrix for all the combinations of the model
mv	The total number of columns in data_train/data_valid
tn	The total no of best models to be identified
prev	The prevalence of disease in the data
pthreshold	The P value threshold for the significance level
method	The methods to be used to evaluate models

Value

This function will generate all possible model outcomes for validation and test dataset

Examples

```
## Not run:
dat <- predict_validation
mv=8
tn=15
prev=0.047
model_evaluation(dat,mv,tn,prev)
```

```
#This process will generate three distinct output files in the working directory
#named evaluation1.out, evaluation2.out and evaluation3.out.
#For details (see https://github.com/mommy003/MSML).

## End(Not run)
```

`predict_validation` *target phenotype and 127 sets of model configurations based on validation dataset*

Description

A dataset containing target phenotype and 127 sets of model configurations based on validation dataset

Usage

```
predict_validation
```

Format

A data frame for models_test:

V1 target, phenotype

V2 model1, based on configurations

V3 model2, based on configurations

V4 model3, based on configurations

V5 model4, based on configurations

V6 model5, based on configurations

V7 model6, based on configurations

V8 model7, based on configurations

V9 model8, based on configurations

V10 model9, based on configurations

V11 model10, based on configurations

V12 model11, based on configurations

V13 model12, based on configurations

V14 model13, based on configurations

V15 model14, based on configurations

V16 model15, based on configurations

V17 model16, based on configurations

V18 model17, based on configurations

V19 model18, based on configurations

V20 model19, based on configurations

V21 model10, based on configurations

V22 model21, based on configurations

V23 model22, based on configurations
V24 model23, based on configurations
V25 model24, based on configurations
V26 model25, based on configurations
V27 model26, based on configurations
V28 model27, based on configurations
V29 model28, based on configurations
V30 model29, based on configurations
V31 model30, based on configurations
V32 model31, based on configurations
V33 model32, based on configurations
V34 model33, based on configurations
V35 model34, based on configurations
V36 model35, based on configurations
V37 model36, based on configurations
V38 model37, based on configurations
V39 model38, based on configurations
V40 model39, based on configurations
V41 model40, based on configurations
V42 model41, based on configurations
V43 model42, based on configurations
V44 model43, based on configurations
V45 model44, based on configurations
V46 model45, based on configurations
V47 model46, based on configurations
V48 model47, based on configurations
V49 model48, based on configurations
V50 model49, based on configurations
V51 model50, based on configurations
V52 model51, based on configurations
V53 model52, based on configurations
V54 model53, based on configurations
V55 model54, based on configurations
V56 model55, based on configurations
V57 model56, based on configurations
V58 model57, based on configurations
V59 model58, based on configurations
V60 model59, based on configurations
V61 model60, based on configurations
V62 model61, based on configurations

V63 model62, based on configurations
V64 model63, based on configurations
V65 model64, based on configurations
V66 model65, based on configurations
V67 model66, based on configurations
V68 model67, based on configurations
V69 model68, based on configurations
V70 model69, based on configurations
V71 model70, based on configurations
V72 model71, based on configurations
V73 model72, based on configurations
V74 model73, based on configurations
V75 model74, based on configurations
V76 model75, based on configurations
V77 model76, based on configurations
V78 model77, based on configurations
V79 model78, based on configurations
V80 model79, based on configurations
V81 model80, based on configurations
V82 model81, based on configurations
V83 model82, based on configurations
V84 model83, based on configurations
V85 model84, based on configurations
V86 model85, based on configurations
V87 model86, based on configurations
V88 model87, based on configurations
V89 model88, based on configurations
V90 model89, based on configurations
V91 model90, based on configurations
V92 model91, based on configurations
V93 model92, based on configurations
V94 model93, based on configurations
V95 model94, based on configurations
V96 model95, based on configurations
V97 model96, based on configurations
V98 model97, based on configurations
V99 model98, based on configurations
V100 model99, based on configurations
V101 model100, based on configurations
V102 model101, based on configurations

V103 model102, based on configurations
V104 model103, based on configurations
V105 model104, based on configurations
V106 model105, based on configurations
V107 model106, based on configurations
V108 model107, based on configurations
V109 model108, based on configurations
V110 model109, based on configurations
V111 model110, based on configurations
V112 model111, based on configurations
V113 model112, based on configurations
V114 model113, based on configurations
V115 model114, based on configurations
V116 model115, based on configurations
V117 model116, based on configurations
V118 model117, based on configurations
V119 model118, based on configurations
V120 model119, based on configurations
V121 model120, based on configurations
V122 model121, based on configurations
V123 model122, based on configurations
V124 model123, based on configurations
V125 model124, based on configurations
V126 model125, based on configurations
V127 model126, based on configurations
V128 model127, based on configurations

Index