Package 'batl'

June 1, 2021

Type Package

Title Bayesian annotations for targeted lipidomics

Version 2.05.6

Imports data.table,
doFuture,
foreach,
future,
graphics,
igraph,
openxlsx,
parallel,
progressr,
stats,
utils

Depends R (>= 3.6.0)

Suggests knitr,
rmarkdown

Description Annotates targeted lipidomics data (SRM/MRM/NL/PIS) using a naive Bayes model and a local, user-defined training dataset.

License GPL-3

Encoding UTF-8

LazyData true

biocViews Bayesian,
Lipidomics,
MassSpectrometry,
Metabolomics,
Software

RoxygenNote 7.1.1

VignetteBuilder knitr

R topics documented:

`deartifact_export_peaks` *deartifact_export_peaks*

Description

Wrapper function to export a peak file from `deartifact_peaks`.

Usage

```
deartifact_export_peaks(file, filename)
```

Arguments

`file`	Data table of peak file with the Insource_annotation column.
`filename`	Name and path of output text file (ending in *.txt)

Examples

```
## Not run:
See batl-Introduction vignette for details

## End(Not run)
```

deartifact_import_peaks *deartifact_import_peaks*

Description

Wrapper function to import a peak file from deartifact_peaks.

Usage

```
deartifact_import_peaks(filename)
```

Arguments

filename	Name and path of output text file (ending in *.txt).

Examples

```
## Not run:
See batl-Introduction vignette for details.

## End(Not run)
```

deartifact_peaks *deartifact_peaks*

Description

Annotate a vector of targeted lipidomic peak files with in-source sphingolipid lipid artifacts (dehydrations, multimerization, deglycosylations, etc.) and isotopes.

The following columns must be present in the peak file: "Index", "Sample.Index", "Sample.Name", "Mass.Info", QSTANDARD_COL (column which refers to the component names of the MS method and includes the component name corresponding to the internal standard), and FEATURE_COLS (features describing each peak).

A "Retention Time" or "Retention.Time" feature column must be present to work.

Usage

```
deartifact_peaks(
    filenames,
    category,
    Q1_tolerance,
    retention_tolerance,
    exceptions,
    ...)
```

Arguments

filenames	Vector of peak files, peak file data frame, or list of peak file data frames.
category	LIPID MAPS 2-letter string denoting the lipid category. Currently supports "FA", "GL", "GP", "SP", "ST", "PR", "SL", "PK".
Q1_tolerance	m/z tolerance of mass spectrometer. '0.5' means +/- 0.5 m/z units above and below the target m/z.
retention_tolerance	Number of significant digits to match peaks by shared retention time. 2 is the default.
exceptions	Keyword exceptions in the "Mass Info" column indicating rows to exclude from labelling.
...	Advanced parameter to customize the 'annotation_table'.

Details

... is a 3 column data.table named "Delta_q1", "Insource_annotation", and "Type". "Delta_q1" is the expected difference in parent ion m/z of the annotation named under the column "Insource_annotation". Example: +1 isotope would have a value of +1 while a dehydration would have a value of -18. Type is an integer used to compute all compound annotations. Compound annotations are only generated from annotations with unique integer types.

Value

List of output files with an additional column "Insource_annotation" indicating any in-source artifacts/isotopes.

See Also

deartifact_import _peaks for loading these files into a data table and deartifact_export_peaks for exporting these peak files to a tab-delimited file.

Examples

```
## Not run:
See batl-Introduction vignette for details

## End(Not run)
```

---

nb_append_library *nb_append_library*

Description

Reshape and concatenate peak files to the training set.

Usage

```
nb_append_library(
    library,
    labelled_files,
    filenames,
    qstandard,
    qstandard_col,
    subtract_constant,
    barcode_decision,
    exceptions,
    ...)
```

Arguments

library	Labelled peak list library.
labelled_files	List of labelled peak files.
filenames	Vector of peak file names (to create the Filename column).
qstandard	Internal standard for feature normalization.
qstandard_col	Column name where the internal standard is indicated.
subtract_constant	Constant for subtracted retention time feature
barcode_decision	Barcode_<decision_rule> column from BATL
exceptions	Keyword exceptions in the "Mass Info" column indicating rows to exclude from labelling.
...	Advanced argument to specify which features should be normalized to the internal standard. Default setting normalizes "Area", "Height", "Relative.RT", and "Subtracted.RT" using the internal standard.

Details

... is a 2 column matrix with column names "Feature" and "Normalize". The 1st column contains the feature names and 2nd column is TRUE/FALSE, where TRUE indicates that feature should be relativized to the corresponding feature value of the "qstandard".

Examples

```
## Not run:
See batl-Introduction vignette for details

## End(Not run)
```

nb_build_model *nb_build_model*

Description

Construct the naive Bayes model using the specified combination of features or all feature subsets.

Usage

```
nb_build_model(
    library,
    folds,
    features,
    exhaustive,
    decision,
    tolerance,
    parallel_cores,
    pseudocount,
    full_output)
```

Arguments

library	Data table of labelled SRM data. See *Details* for formatting.
folds	Number of folds in *k*-fold cross validation.
features	Vector of feature names shared in the peak file/library.
exhaustive	Build and train models for all combinations of features. Default is FALSE.
decision	Decision algorithm. Default is "MWBM". Other options include "MAP_dup", and "MAP_nodup". See *Details* for more information.
tolerance	Q1 and Q3 m/z tolerance for matching peaks to library barcodes. m/z +/- 0.5 is the default.
parallel_cores	Number of cores to enable parallel computing of cross validation.
pseudocount	Default is 0. For prior estimation.
full_output	TRUE will return a data table of all possible assignments during cross validation for model selection. Default is FALSE.

Details

"library" (also known as the training set) must contain the following columns in this order: Filename, Sample.Index, Sample.Name, Index, Q1, Q3, Barcode, and all desired peak features.

decision = MWBM (maximum weighted bipartite matching) assigns lipid labels by maximizing the joint log posterior probability of all peak assignments within a transition.

decision = MAP_dup assigns lipid labels based on the greatest posterior probability. It is possible that two or more peaks within the same sample will be assigned the same lipid identity, however.

decision = MAP_nodup assigns lipid labels based on the greatest posterior probability constrained such that the same lipid label will not be assigned more than once within a sample.

full_output = TRUE if you would like to analyze your cross validation results from your personal training set using *nb_model_scores*.

Value

List of data tables for each feature combination, containing the priors, likelihood parameters, posterior probability cutoffs, and assigned lipid identity. Also includes relevant meta information. When *full_output* is TRUE, a list of all peak assignments is returned and a list of all posterior cutoffs across all cross validation folds.

Examples

```
## Not run:
See batl-Introduction vignette

## End(Not run)
```

`nb_export_labelled_file` *nb_export_labelled_file*

Description

Wrapper function to export a labelled peak file to a text file from `nb_label_peaks`.

Usage

```
nb_export_labelled_file(labelled_file, filename)
```

Arguments

`labelled_file`	Labelled peak file.
`filename`	Name and path of output text file (ending in *.txt)

Examples

```
## Not run:
See batl-Introduction vignette for details

## End(Not run)
```

nb_export_library *nb_export_library*

Description

Wrapper function to export a data.table library/training set generated from nb_append_library to a tab-delimited text file.

Usage

```
nb_export_library(library, filename)
```

Arguments

library	Data.table library of labelled peaks.
filename	Name and path of output text file (ending in *.txt)

Examples

```
## Not run:
See batl-Introduction vignette for details

## End(Not run)
```

nb_export_model *nb_export_model*

Description

Wrapper function to export a BATL model to a tab-delimited text file from nb_build_model.

Usage

```
nb_export_model(nb_model, filename)
```

Arguments

nb_model	List containing data table model and meta information.
filename	Name and path of output text file (ending in *.txt)

Examples

```
## Not run:
See batl-Introduction vignette for details

## End(Not run)
```

nb_import_labelled_file *nb_import_labelled_file*

Description

Wrapper function to import a tab-delimited, labelled peak file.

Usage

```
nb_import_labelled_file(labelled_file)
```

Arguments

labelled_file	Name and path of labelled peak file (ending in *.txt).

Examples

```
## Not run:
See batl-Introduction vignette for details

## End(Not run)
```

nb_import_labelled_file_list *nb_import_labelled_file_list*

Description

Wrapper function to import one or more tab-delimited, labelled peak files as a list.

Usage

```
nb_import_labelled_file_list(labelled_files)
```

Arguments

labelled_files	Name and path of one or more labelled peak files (ending in *.txt).

Examples

```
## Not run:
See batl-Introduction vignette for details

## End(Not run)
```

nb_import_library *nb_import_library*

Description

Wrapper function to import a BATL training set from a tab-delimited text file.

Usage

```
nb_import_library(library)
```

Arguments

library	Name and path of the tab-delimited targeted lipidomics library/training set (ending in *.txt).

Examples

```
## Not run:
See batl-Introduction vignette for details

## End(Not run)
```

nb_import_model *nb_import_model*

Description

Wrapper function to import a BATL model from a tab-delimited text file.

Usage

```
nb_import_model(model_name)
```

Arguments

model_name	Name and path of the BATL model (ending in *.txt).

Examples

```
## Not run:
See batl-Introduction vignette

## End(Not run)
```

nb_label_peaks *nb_label_peaks*

Description

Label a vector of peak files using a trained BATL model from nb_build_model.

Usage

```
nb_label_peaks(
    filenames,
    model_filename,
    qstandard,
    qstandard_col,
    subtract_constant,
    exceptions,
    ...)
```

Arguments

filenames	Vector of peak files to label.
model_filename	Name of trained BATL model.
qstandard	Internal standard for feature normalization.
qstandard_col	Column name where the internal standard is indicated.
subtract_constant	Constant for subtracted retention time feature
exceptions	Keyword exceptions in the "Mass Info" column indicating rows to exclude from labelling.
...	Advanced argument to specify which features should be normalized to the internal standard. Default setting normalizes "Area", "Height", "Relative.RT", and "Subtracted.RT" using the internal standard.

Details

... is a 2 column matrix with column names "Feature" and "Normalize". The 1st column contains the feature names and the 2nd column is TRUE/FALSE, where TRUE indicates that feature should be relativized to the corresponding feature value of the "qstandard".

Examples

```
## Not run:
See batl-Introduction vignette

## End(Not run)
```

nb_model_scores *nb_model_scores*

Description

Flatten all *k*-fold cross validation list results and assess model performance from nb_build_model.

Usage

```
nb_model_scores(model_assignments)
```

Arguments

model_assignments	Cross validation assignments outputted from *nb_label_peaks* with full_output set to TRUE.

Examples

```
## Not run:
See batl-Introduction vignette for details

## End(Not run)
```

srm_tc *Targeted lipidomics dataset of mouse temporcal cortex*

Description

Toy SRM dataset.

Usage

```
data(srm_tc)
```

Format

An object of class data.table (inherits from data.frame) with 861 rows and 14 columns.

Examples

```
data(srm_tc)
```

Index