

Package ‘mleap’

January 10, 2020

Type Package

Title Interface to 'MLeap'

Version 1.0.0

Description A 'sparklyr' <<https://spark.rstudio.com>> extension that provides an interface to 'MLeap' <<https://github.com/combust/mleap>>, an open source library that enables exporting and serving of 'Apache Spark' pipelines.

URL <https://github.com/rstudio/mleap>

BugReports <https://github.com/rstudio/mleap/issues>

License Apache License (>= 2.0)

SystemRequirements Apache Spark 2.0+, Apache Maven 3.5+, Java JDK 8, MLeap Runtime 0.10.1+

Encoding UTF-8

LazyData true

Depends R (>= 3.1.2)

Imports sparklyr, jsonlite, purrr, fs, tibble, rJava, digest

RoxygenNote 6.1.1

Suggests testthat, covr, ggplot2, dplyr

NeedsCompilation no

Author Kevin Kuo [aut, cre] (<<https://orcid.org/0000-0001-7803-7901>>)

Maintainer Kevin Kuo <kevin.kuo@rstudio.com>

Repository CRAN

Date/Publication 2020-01-10 22:20:02 UTC

R topics documented:

install_maven	2
install_mleap	2
mleap_installed_versions	3
mleap_load_bundle	3

mleap_model_schema	4
mleap_transform	4
ml_write_bundle	5

Index	6
--------------	----------

install_maven	<i>Install Maven</i>
---------------	----------------------

Description

This function installs Apache Maven.

Usage

```
install_maven(dir = NULL, version = NULL)
```

Arguments

dir	(Optional) Directory to install maven in. Defaults to maven/ under user's home directory.
version	Version of Maven to install, defaults to the latest version tested with this package.

Examples

```
## Not run:
install_maven()

## End(Not run)
```

install_mleap	<i>Install MLeap runtime</i>
---------------	------------------------------

Description

Install MLeap runtime

Usage

```
install_mleap(dir = NULL, version = NULL, use_temp_cache = TRUE)
```

Arguments

dir	(Optional) Directory to save the jars
version	Version of MLeap to install, defaults to the latest version tested with this package.
use_temp_cache	Whether to use a temporary Maven cache directory for downloading. Setting this to TRUE prevents Maven from creating a persistent .m2/ directory. Defaults to TRUE.

Examples

```
## Not run:  
install_mleap()  
  
## End(Not run)
```

```
mleap_installed_versions  
      Find existing MLeap installations
```

Description

Find existing MLeap installations

Usage

```
mleap_installed_versions()
```

Value

A data frame of MLeap Runtime installation versions and their locations.

```
mleap_load_bundle      Loads an MLeap bundle
```

Description

Loads an MLeap bundle

Usage

```
mleap_load_bundle(path)
```

Arguments

path	Path to the exported bundle zip file.
------	---------------------------------------

Value

An MLeap model object.

mleap_model_schema	<i>MLeap model schema</i>
--------------------	---------------------------

Description

Returns the schema of an MLeap transformer.

Usage

```
mleap_model_schema(x)
```

Arguments

x An MLeap model object.

Value

A data frame of the model schema.

mleap_transform	<i>Transform data using an MLeap model</i>
-----------------	--

Description

This functions transforms an R data frame using an MLeap model.

Usage

```
mleap_transform(model, data)
```

Arguments

model An MLeap model object, obtained by mleap_load_bundle().
 data An R data frame.

Value

A transformed data frame.

See Also

[mleap_load_bundle()]

ml_write_bundle	<i>Export a Spark pipeline for serving</i>
-----------------	--

Description

This functions serializes a Spark pipeline model into an MLeap bundle.

Usage

```
ml_write_bundle(x, sample_input, path, overwrite = FALSE)
```

Arguments

x	A Spark pipeline model object.
sample_input	A sample input Spark DataFrame with the expected schema.
path	Where to save the bundle.
overwrite	Whether to overwrite an existing file, defaults to FALSE.

Examples

```
## Not run:
library(sparklyr)
sc <- spark_connect(master = "local")
mtcars_tbl <- sdf_copy_to(sc, mtcars, overwrite = TRUE)
pipeline <- ml_pipeline(sc) %>%
  ft_binarizer("hp", "big_hp", threshold = 100) %>%
  ft_vector_assembler(c("big_hp", "wt", "qsec"), "features") %>%
  ml_gbt_regressor(label_col = "mpg")
pipeline_model <- ml_fit(pipeline, mtcars_tbl)
model_path <- file.path(tempdir(), "mtcars_model.zip")
ml_write_bundle(pipeline_model,
                mtcars_tbl,
                model_path,
                overwrite = TRUE)

## End(Not run)
```

Index

`install_maven`, 2

`install_mleap`, 2

`ml_write_bundle`, 5

`mleap_installed_versions`, 3

`mleap_load_bundle`, 3

`mleap_model_schema`, 4

`mleap_transform`, 4