Package 'viralmodels'

January 21, 2024

Title Viral Load and CD4 Lymphocytes Regression Models

Version 1.2.0

Description

Provides a comprehensive framework for building, evaluating, and visualizing regression models for analyzing viral load and CD4 (Cluster of Differentiation 4) lymphocytes data. It leverages the principles of the tidymodels ecosystem of Max Kuhn and Hadley Wickham (2020) <https://www.tidymodels.org> to offer a user-friendly experience in model development. This package includes functions for data preprocessing, feature engineering, model training, tuning, and evaluation, along with visualization tools to enhance the interpretation of model results. It is specifically designed for researchers in biostatistics, computational biology, and HIV research who aim to perform reproducible and rigorous analyses to gain insights into disease dynamics. The main focus is on improving the understanding of the relationships between viral load, CD4 lymphocytes, and other relevant covariates to contribute to HIV research and the visibility of vulnerable seropositive populations.

License GPL (>= 3)

Encoding UTF-8

RoxygenNote 7.2.3

Imports dplyr, earth, kknn, nnet, parsnip, recipes, rsample, tidyselect, tune, vdiffr, workflows, workflowsets

Suggests testthat (>= 3.0.0)

Config/testthat/edition 3

NeedsCompilation no

Author Juan Pablo Acuña González [aut, cre] (<https://orcid.org/0009-0003-6029-6560>)

Maintainer Juan Pablo Acuña González <22253567@uagro.mx>

Repository CRAN

Date/Publication 2024-01-21 13:00:05 UTC

R topics documented:

---

viralmodel	*Select best model*

Description

viralmodel returns metrics for a selected model

Usage

```
viralmodel(x, semilla, target, pliegues, repeticiones, rejilla, modelo)
```

Arguments

x	A data frame
semilla	A numeric value
target	A character value
pliegues	A numeric value
repeticiones	A numeric value
rejilla	A numeric value
modelo	A character value

Value

A table with a single model hyperparameters

Examples

```
cd_2019 <- c(824, 169, 342, 423, 441, 507, 559,
             173, 764, 780, 244, 527, 417, 800,
             602, 494, 345, 780, 780, 527, 556,
             559, 238, 288, 244, 353, 169, 556,
             824, 169, 342, 423, 441, 507, 559)
vl_2019 <- c(40, 11388, 38961, 40, 75, 4095, 103,
             11388, 46, 103, 11388, 40, 0, 11388,
             0,  4095,  40,  93,  49,  49,  49,
             4095,  6837, 38961, 38961, 0, 0, 93,
             40, 11388, 38961, 40, 75, 4095, 103)
cd_2021 <- c(992, 275, 331, 454, 479, 553,  496,
             230, 605, 432, 170, 670, 238,  238,
             634, 422, 429, 513, 327, 465,  479,
             661, 382, 364, 109, 398, 209, 1960,
             992, 275, 331, 454, 479, 553,  496)
vl_2021 <- c(80, 1690,  5113,  71,  289,  3063,  0,
             262,  0,  15089,  13016, 1513, 60, 60,
             49248, 159308, 56, 0, 516675, 49, 237,
             84,  292,  414, 26176,  62,  126,  93,
```

```
                80, 1690, 5113,    71, 289, 3063,    0)
cd_2022 <- c(700, 127, 127, 547, 547, 547, 777,
                149, 628, 614, 253, 918, 326, 326,
                574, 361, 253, 726, 659, 596, 427,
                447, 326, 253, 248, 326, 260, 918,
                700, 127, 127, 547, 547, 547, 777)
vl_2022 <- c(0,    0,   53250,   0,    40,   1901, 0,
                955,    0,    0,    0,    0,   40,    0,
                49248, 159308, 56, 0, 516675, 49, 237,
                0,    23601,   0,   40,    0,    0,    0,
                0,    0,    0,    0,    0,    0, 0)
x <- cbind(cd_2019, vl_2019, cd_2021, vl_2021, cd_2022, vl_2022) |> as.data.frame()
semilla <- 123
target <- "cd_2022"
pliegues <- 2
repeticiones <- 1
rejilla <- 1
modelo <- "simple_MARS"
viralmodel(x, semilla, target, pliegues, repeticiones, rejilla, modelo)
```

---

viraltab	*Competing models table*

Description

viraltab trains and optimizes a series of regression models for viral load or cd4 counts

Usage

```
viraltab(x, semilla, target, pliegues, repeticiones, rejilla)
```

Arguments

x	A data frame
semilla	A numeric value
target	A character value
pliegues	A numeric value
repeticiones	A numeric value
rejilla	A numeric value

Value

A table of competing models

Examples

```
cd_2019 <- c(824, 169, 342, 423, 441, 507, 559,
             173, 764, 780, 244, 527, 417, 800,
             602, 494, 345, 780, 780, 527, 556,
             559, 238, 288, 244, 353, 169, 556,
             824, 169, 342, 423, 441, 507, 559)
vl_2019 <- c(40, 11388, 38961, 40, 75, 4095, 103,
             11388, 46, 103, 11388, 40, 0, 11388,
             0,  4095,  40,  93,  49,  49,  49,
             4095,  6837, 38961, 38961, 0, 0, 93,
             40, 11388, 38961, 40, 75, 4095, 103)
cd_2021 <- c(992, 275, 331, 454, 479, 553,  496,
             230, 605, 432, 170, 670, 238,  238,
             634, 422, 429, 513, 327, 465,  479,
             661, 382, 364, 109, 398, 209, 1960,
             992, 275, 331, 454, 479, 553,  496)
vl_2021 <- c(80, 1690,  5113,  71,  289,  3063,  0,
             262,  0,  15089,  13016, 1513, 60, 60,
             49248, 159308, 56, 0, 516675, 49, 237,
             84,  292,  414, 26176,  62,  126,  93,
             80, 1690, 5113,   71, 289, 3063,   0)
cd_2022 <- c(700, 127, 127, 547, 547, 547, 777,
             149, 628, 614, 253, 918, 326, 326,
             574, 361, 253, 726, 659, 596, 427,
             447, 326, 253, 248, 326, 260, 918,
             700, 127, 127, 547, 547, 547, 777)
vl_2022 <- c(0,   0,   53250,  0,   40,   1901, 0,
             955,  0,    0,    0,   0,    40,   0,
             49248, 159308, 56, 0, 516675, 49, 237,
             0,   23601,  0,   40,   0,    0,   0,
             0,   0,    0,    0,   0,    0,   0)
x <- cbind(cd_2019, vl_2019, cd_2021, vl_2021, cd_2022, vl_2022) |> as.data.frame()
semilla <- 123
target <- "cd_2022"
pliegues <- 2
repeticiones <- 1
rejilla <- 1
viraltab(x, semilla, target, pliegues, repeticiones, rejilla)
```

---

viralvis	*Competing models plot*

Description

viralvis plots the rankings of a series of regression models for viral load or cd4 counts

Usage

```
viralvis(x, semilla, target, pliegues, repeticiones, rejilla)
```

Arguments

`x`	A data frame
`semilla`	A numeric value
`target`	A character value
`pliegues`	A numeric value
`repeticiones`	A numeric value
`rejilla`	A numeric value

Value

A plot of ranking models

Examples

```
cd_2019 <- c(824, 169, 342, 423, 441, 507, 559,
             173, 764, 780, 244, 527, 417, 800,
             602, 494, 345, 780, 780, 527, 556,
             559, 238, 288, 244, 353, 169, 556,
             824, 169, 342, 423, 441, 507, 559)
vl_2019 <- c(40, 11388, 38961, 40, 75, 4095, 103,
             11388, 46, 103, 11388, 40, 0, 11388,
             0,  4095,  40, 93, 49,  49,  49,
             4095, 6837, 38961, 38961, 0, 0, 93,
             40, 11388, 38961, 40, 75, 4095, 103)
cd_2021 <- c(992, 275, 331, 454, 479, 553,  496,
             230, 605, 432, 170, 670, 238,  238,
             634, 422, 429, 513, 327, 465,  479,
             661, 382, 364, 109, 398, 209, 1960,
             992, 275, 331, 454, 479, 553,  496)
vl_2021 <- c(80, 1690,  5113,  71,  289,  3063,  0,
             262,  0, 15089,  13016, 1513, 60, 60,
             49248, 159308, 56, 0, 516675, 49, 237,
             84,  292,  414, 26176,  62,  126,  93,
             80, 1690, 5113,    71, 289, 3063,   0)
cd_2022 <- c(700, 127, 127, 547, 547, 547, 777,
             149, 628, 614, 253, 918, 326, 326,
             574, 361, 253, 726, 659, 596, 427,
             447, 326, 253, 248, 326, 260, 918,
             700, 127, 127, 547, 547, 547, 777)
vl_2022 <- c(0,   0,  53250,   0,   40,   1901, 0,
             955,  0,    0,    0,   0,    40,   0,
             49248, 159308, 56, 0, 516675, 49, 237,
             0,    23601,  0,   40,   0,   0,   0,
             0,    0,     0,    0,    0,    0,  0)
x <- cbind(cd_2019, vl_2019, cd_2021, vl_2021, cd_2022, vl_2022) |> as.data.frame()
semilla <- 123
target <- "cd_2022"
pliegues <- 2
repeticiones <- 1
rejilla <- 1
```

```
viralvis(x, semilla, target, pliegues, repeticiones, rejilla)
```

Index