

# Package ‘tetragon’

November 22, 2021

**Type** Package

**Title** Automatic Sequence Prediction by Expansion of the Distance Matrix

**Version** 1.0.0

**Author** Giancarlo Vercellino

**Maintainer** Giancarlo Vercellino <giancarlo.vercellino@gmail.com>

**Description** Each sequence is predicted by expanding the distance matrix. The compact set of hyper-parameters is tuned via grid or random search.

**License** GPL-3

**Encoding** UTF-8

**LazyData** true

**RoxygenNote** 7.1.1

**Depends** R (>= 3.6)

**Imports** purrr (>= 0.3.4), abind (>= 1.4-5), ggplot2 (>= 3.3.5), readr (>= 2.0.1), stringr (>= 1.4.0), lubridate (>= 1.7.10), narray (>= 0.4.1.1), imputeTS (>= 3.2), scales (>= 1.1.1), tictoc (>= 1.0.1), modeest (>= 2.4.0), moments (>= 0.14), greybox (>= 1.0.1), philentropy (>= 0.5.0)

**URL** [https://rpubs.com/giancarlo\\_vercellino/tetragon](https://rpubs.com/giancarlo_vercellino/tetragon)

**NeedsCompilation** no

**Repository** CRAN

**Date/Publication** 2021-11-22 09:50:05 UTC

## R topics documented:

covid_in_europe	2
engine	2
tetragon	3

<b>Index</b>	<b>5</b>
--------------	----------

---

covid_in_europe	<i>covid_in_europe data set</i>
-----------------	---------------------------------

---

**Description**

A data frame with with daily and cumulative cases of Covid infections and deaths in Europe since March 2021.

**Usage**

```
covid_in_europe
```

**Format**

A data frame with 5 columns and 163 rows.

**Source**

[www.ecdc.europa.eu](http://www.ecdc.europa.eu)

---

engine	<i>support functions for tetragon</i>
--------	---------------------------------------

---

**Description**

support functions for tetragon

**Usage**

```
engine(df, seq_len, deriv, ci, method, distr, measure_error, n_sim, dates)
```

**Arguments**

df	A data frame with time features on columns
seq_len	Positive integer. Time-step number of the projected sequence
deriv	Integer vector. Number of differentiation operations to perform for each original time feature. 0 = no change; 1: one diff; 2: two diff.
ci	Confidence interval. Default: 0.8
method	Positive integer. Distance method for calculating neighbors. Possible options are: "euclidean", "manhattan", "chebyshev", "sorensen", "gower", "soergel", "kulczynski_d", "canberra", "lorentzian", "intersection", "non-intersection", "wave-hedges", "czekanowski", "motyka", "kulczynski_s", "tanimoto", "ruzicka", "inner_product", "harmonic_mean", "cosine", "hassebrook", "jaccard", "dice", "fidelity", "bhattacharyya", "squared_chord", "squared_euclidean", "pearson", "neyman", "squared_chi", "prob_symm", "divergence", "clark", "additive_symm", "taneja", "kumar-johnson", "avg".

distr	String. DIstribution used to calculate kernel densities. Possible options are: "norm", "cauchy", "logis", "t", "exp".
measure_error	Logical. TRUE for measuring validation error. FALSE otherwise.
n_sim	String. Sequencing method: deterministic ("segmented"), or non-deterministic ("sampled").
dates	Date. Vector with dates for time features.

**Author(s)**

Giancarlo Vercellino <giancarlo.vercellino@gmail.com>

---

tetragon	<i>tetragon</i>
----------	-----------------

---

**Description**

blablabla

**Usage**

```
tetragon(
  df,
  seq_len = NULL,
  deriv = NULL,
  ci = 0.8,
  method = NULL,
  distr = NULL,
  n_windows = 10,
  n_sample = 30,
  search = "random",
  fixed = FALSE,
  dates = NULL,
  seed = 42
)
```

**Arguments**

df	A data frame with time features on columns.
seq_len	Positive integer. Time-step number of the projected sequence.
deriv	Integer vector. Number of differentiation operations to perform for each original time feature. 0 = no change; 1: one diff; 2: two diff.
ci	Confidence interval. Default: 0.8.
method	String. Distance method for calculating distance matrix among sequences. Possible options are: "euclidean", "manhattan", "chebyshev", "gower", "lorentzian", "jaccard", "dice", "squared_euclidean", "divergence", "clark", "avg". For further information, please make reference to philentropy package. Default: NULL (random selection among all possible options).

distr	String. DIstribution used to expand the distance matrix. Possible options are: "norm", "cauchy", "logis", "t", "exp". Default: NULL (random selection among all possible options).
n_windows	Positive integer. Number of validation tests to measure/sample error. Default: 10.
n_sample	Positive integer. Number of samples for grid or random search. Default: 30.
search	String. Two option available: "grid", "random". Default: "random".
fixed	Logical. Setting to TRUE, calculate a single model (if the variables passed to the functions are completed). Default: FALSE.
dates	Date. Vector with dates for time features.
seed	Positive integer. Random seed. Default: 42.

### Value

This function returns a list including:

- exploration: list of all not-null models, complete with predictions, test metrics, prediction stats and plot
- history: a table with the sampled models, hyper-parameters, validation errors, weighted average rank
- best: results for the best model in term of weighted average rank, including:
  - wt\_avg\_best: hyper-parameters of the best model selected through grid/random search
  - predictions: min, max, q25, q50, q75, quantiles at selected ci, mean, sd, mode, kurtosis, skewness for each time feature
  - testing\_errors: testing errors for each time feature (ME, MAE, MSE, MPE, MAPE, sCE, MAPE, sMAE, sMSE, MASE, RMSSE, rMAE, rRMSE, rAME). For further information, make reference to greybox package.
  - pred\_stats: for each predicted time feature, IQR to range, KL-divergence, risk ratio, up-side probability, averaged across prediction time-points and at the terminal points.
- time\_log

### Author(s)

Giancarlo Vercellino <giancarlo.vercellino@gmail.com>

### See Also

Useful links:

- [https://rpubs.com/giancarlo\\_vercellino/tetragon](https://rpubs.com/giancarlo_vercellino/tetragon)

### Examples

```
tetragon(covid_in_europe[, c(2, 4)], seq_len = 40, n_sample = 2, deriv = c(1, 2))
```

# Index

## \* datasets

covid\_in\_europe, [2](#)

covid\_in\_europe, [2](#)

engine, [2](#)

tetragon, [3](#)

tetragon-package (tetragon), [3](#)