

Package: codez (via r-universe)

October 25, 2024

Type Package

Title Seq2Seq Encoder-Decoder Model for Time-Feature Analysis Based on Tensorflow

Version 1.0.0

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

Description Proposes Seq2seq Time-Feature Analysis using an Encoder-Decoder to project into latent space and a Forward Network to predict the next sequence.

License GPL-3

Encoding UTF-8

LazyData true

RoxygenNote 7.2.1

Depends R (>= 3.6)

Imports purrr (>= 0.3.4), abind (>= 1.4-5), ggplot2 (>= 3.3.6), readr (>= 2.1.2), fANCOVA (>= 0.6-1), imputeTS (>= 3.2), modeest (>= 2.4.0), scales (>= 1.1.1), tictoc (>= 1.0.1), tensorflow (>= 2.9.0), keras (>= 2.9.0), moments (>= 0.14), narray (>= 0.4.1.1), fastDummies (>= 1.6.3), entropy (>= 1.3.1), philentropy (>= 0.5.0), greybox (>= 1.0.1), lubridate (>= 1.7.10)

URL https://rpubs.com/giancarlo_vercellino/codez

Suggests testthat (>= 3.0.0), reticulate (>= 1.26)

Config/testthat/edition 3

NeedsCompilation no

Author Giancarlo Vercellino [aut, cre, cph]

Date/Publication 2022-09-23 15:10:02 UTC

Repository <https://pigan.r-universe.dev>

RemoteUrl <https://github.com/cran/codez>

RemoteRef HEAD

RemoteSha d69abaf24b176b0f3ecfe063bcc7e756c8f6ed8

Contents

amzn_aapl_fb	2
codez	2
Index	6

amzn_aapl_fb	<i>amzn_aapl_fb data set</i>
--------------	------------------------------

Description

A data frame with the close prices for Amazon, Google and Facebook.

Usage

```
amzn_aapl_fb
```

Format

A data frame with 4 columns and 1798 rows.

Source

Yahoo Finance

codez	<i>codez</i>
-------	--------------

Description

Seq2seq Time-Feature Analysis using an Encoder-Decoder to project into latent space and a Forward Network to predict the next sequence.

Usage

```
codez(
  df,
  seq_len = NULL,
  n_windows = 10,
  latent = NULL,
  smoother = FALSE,
  n_samp = 30,
  autoencoder_layers_n = NULL,
  autoencoder_layers_size = NULL,
  autoencoder_activ = NULL,
  forward_net_layers_n = NULL,
```

```

forward_net_layers_size = NULL,
forward_net_activ = NULL,
forward_net_reg_L1 = NULL,
forward_net_reg_L2 = NULL,
forward_net_drop = NULL,
loss_metric = "mae",
autoencoder_optimizer = NULL,
forward_net_optimizer = NULL,
epochs = 100,
patience = 10,
holdout = 0.5,
verbose = FALSE,
ci = 0.8,
error_scale = "naive",
error_benchmark = "naive",
dates = NULL,
seed = 42
)

```

Arguments

df	A data frame with time features on columns. They could be numeric variables or categorical, but not both.
seq_len	Positive integer. Time-step number of the forecasting sequence. Default: NULL (random selection within 2 to max preset boundary).
n_windows	Positive integer. Number of validation windows to test prediction error. Default: 10.
latent	Positive integer. Dimensions of the latent space for encoding-decoding operations. Default: NULL (random selection within preset boundaries)
smoother	Logical. Perform optimal smoothing using standard loess for each time feature. Default: FALSE
n_samp	Positive integer. Number of samples for random search. Default: 30.
autoencoder_layers_n	Positive integer. Number of layers for the encoder-decoder model. Default: NULL (random selection within preset boundaries)
autoencoder_layers_size	Positive integer. Numbers of nodes for the encoder-decoder model. Default: NULL (random selection within preset boundaries)
autoencoder_activ	String. Activation function to be used by the encoder-decoder model. Implemented functions are: "linear", "relu", "leaky_relu", "selu", "elu", "sigmoid", "tanh", "swish", "gelu". Default: NULL (random selection within standard activations)
forward_net_layers_n	Positive integer. Number of layers for the forward net model. Default: NULL (random selection within preset boundaries)

<code>forward_net_layers_size</code>	Positive integer. Numbers of nodes for the forward net model. Default: NULL (random selection within preset boundaries)
<code>forward_net_activation</code>	String. Activation function to be used by the forward net model. Implemented functions are: "linear", "relu", "leaky_relu", "selu", "elu", "sigmoid", "tanh", "swish", "gelu". Default: NULL (random selection within standard activations)
<code>forward_net_reg_L1</code>	Positive numeric between. Weights for L1 regularization. Default: NULL (random selection within preset boundaries).
<code>forward_net_reg_L2</code>	Positive numeric between. Weights for L2 regularization. Default: NULL (random selection within preset boundaries).
<code>forward_net_drop</code>	Positive numeric between 0 and 1. Value for the dropout parameter for each layer of the forward net model (for example, a neural net with 3 layers may have dropout = c(0, 0.5, 0.3)). Default: NULL (random selection within preset boundaries).
<code>loss_metric</code>	String. Loss function for both models. Available metrics: "mse", "mae", "mape". Default: "mae".
<code>autoencoder_optimizer</code>	String. Optimization method for autoencoder. Implemented methods are: "adam", "adadelta", "adagrad", "rmsprop", "sgd", "nadam", "adamax". Default: NULL (random selection within standard optimizers).
<code>forward_net_optimizer</code>	String. Optimization method for forward net. Implemented methods are: "adam", "adadelta", "adagrad", "rmsprop", "sgd", "nadam", "adamax". Default: NULL (random selection within standard optimizers).
<code>epochs</code>	Positive integer. Default: 100.
<code>patience</code>	Positive integer. Waiting time (in epochs) before evaluating the overfit performance. Default: 10.
<code>holdout</code>	Positive numeric between 0 and 1. Holdout sample for validation. Default: 0.5.
<code>verbose</code>	Logical. Default: FALSE.
<code>ci</code>	Positive numeric. Confidence interval. Default: 0.8
<code>error_scale</code>	String. Scale for the scaled error metrics (for continuous variables). Two options: "naive" (average of naive one-step absolute error for the historical series) or "deviation" (standard error of the historical series). Default: "naive".
<code>error_benchmark</code>	String. Benchmark for the relative error metrics (for continuous variables). Two options: "naive" (sequential extension of last value) or "average" (mean value of true sequence). Default: "naive".
<code>dates</code>	Date. Vector with dates for time features.
<code>seed</code>	Positive integer. Random seed. Default: 42.

Value

This function returns a list including:

- history: a table with the sampled models, hyper-parameters, validation errors
- best_model: results for the best selected model according to the weighted average rank, including:
 - predictions: for continuous variables, min, max, q25, q50, q75, quantiles at selected ci, mean, sd, mode, skewness, kurtosis, IQR to range, risk ratio, upside probability and divergence for each point fo predicted sequences; for factor variables, min, max, q25, q50, q75, quantiles at selected ci, proportions, difformity (deviation of proportions normalized over the maximum possible deviation), entropy, upgrade probability and divergence for each point fo predicted sequences
 - testing_errors: testing errors for each time feature for the best selected model (for continuous variables: me, mae, mse, rmsse, mpe, mape, rmae, rrmse, rame, mase, smse, sce, gmrae; for factor variables: czekanowski, tanimoto, cosine, hassebrook, jaccard, dice, canberra, gower, lorentzian, clark)
 - plots: standard plots with confidence interval for each time feature
- time_log

Author(s)

Maintainer: Giancarlo Vercellino <giancarlo.vercellino@gmail.com> [copyright holder]

Giancarlo Vercellino <giancarlo.vercellino@gmail.com>

See Also

Useful links:

- https://rpubs.com/giancarlo_vercellino/codez

Index

* **datasets**

amzn_aapl_fb, [2](#)

amzn_aapl_fb, [2](#)

codez, [2](#)

codez-package (codez), [2](#)