

Package: CompareMultipleModels (via r-universe)

September 4, 2024

Type Package

Title Finding the Best Model Using Eight Metrics Values

Version 0.1.0

Author Mr. Sandip Garai [aut, cre]

Maintainer Mr. Sandip Garai <sandipnicksandy@gmail.com>

Description In statistical modeling, multiple models need to be compared based on certain criteria. The method described here uses eight metrics from 'AllMetrics' package. 'input_df' is the data frame (at least two columns for comparison) containing metrics values in different rows of a column (which denotes a particular model's performance). First five metrics are expected to be minimum and last three metrics are expected to be maximum for a model to be considered good. Firstly, every metric value (among first five) is searched in every columns and minimum values are denoted as 'MIN' and other values are denoted as 'NA'. Secondly, every metric (among last three) is searched in every columns and maximum values are denoted as 'MAX' and other values are denoted as 'NA'. 'output_df' contains the similar number of rows (which is 8) and columns (which is number of models to be compared) as of 'input_df'. Values in 'output_df' are corresponding 'NA', 'MIN' or 'MAX'. Finally, the column containing minimum number of 'NA' values is denoted as the best column. 'min_NA_col' gives the name of the best column (model). 'min_NA_values' are the corresponding metrics values. 'BestColumn_metrics' is the data frame (dimension: 1*8) containing different metrics of the best column (model). 'best_column_results' is the final result (a list) containing all of these output elements. In special case, if two columns having equal 'NA', it will be checked among these two column which one is having least 'NA' in first five rows and will be inferred as the best. More details about 'AllMetrics' can be found in Garai (2023) <[doi:10.13140/RG.2.2.18688.30723](https://doi.org/10.13140/RG.2.2.18688.30723)>.

License GPL-3**Encoding** UTF-8**Imports** CEEMDANML**RoxygenNote** 7.2.1**NeedsCompilation** no**Date/Publication** 2023-04-12 11:50:06 UTC**Repository** <https://sandipgarai.r-universe.dev>**RemoteUrl** <https://github.com/cran/CompareMultipleModels>**RemoteRef** HEAD**RemoteSha** f971f314dc7b8b310e4c213013cdb5310e5e55b1

Contents

best_model_finder	2
Index	4

best_model_finder	<i>Finding the Best Model Using Eight Metrics Values</i>
-------------------	----------------------------------------------------------

Description

In statistical modeling, multiple models need to be compared based on certain criteria. The method described here uses eight metrics from "AllMetrics" package. 'input_df' is the data frame (at least two columns for comparison) containing metrics values in different rows of a column (which denotes a particular model's performance). First five metrics are expected to be minimum and last three metrics are expected to be maximum for a model to be considered good. Firstly, every metric value (among first five) is searched in every columns and minimum values are denoted as 'MIN' and other values are denoted as 'NA'. Secondly, every metric (among last three) is searched in every columns and maximum values are denoted as 'MAX' and other values are denoted as 'NA'. 'output_df' contains the similar number of rows (which is 8) and columns (which is number of models to be compared) as of 'input_df'. Values in 'output_df' are corresponding 'NA', 'MIN' or 'MAX'. Finally, the column containing minimum number of 'NA' values is denoted as the best column. 'min_NA_col' gives the name of the best column (model). 'min_NA_values' are the corresponding metrics values. 'BestColumn_metrics' is the data frame (dimension: 1*8) containing different metrics of the best column (model). 'best_column_results' is the final result (a list) containing all of these output elements. In special case, if two columns having equal 'NA', it has been checked among these two columns which one is having least 'NA' in first five rows and has been inferred as the best. More details about 'AllMetrics' can be found in Garai (2023) <doi:10.13140/RG.2.2.18688.30723>.

Usage

```
best_model_finder(input_df)
```

Arguments

`input_df` Data frame containing eight metrics Values in different columns indicating a particular model's performance

Value

- `best_column_results`: A list containing four data frames

References

- Paul, R. K., & Garai, S. (2021). Performance comparison of wavelets-based machine learning technique for forecasting agricultural commodity prices. *Soft Computing*, 25(20), 12857-12873.
- Paul, R. K., & Garai, S. (2022). Wavelets based artificial neural network technique for forecasting agricultural prices. *Journal of the Indian Society for Probability and Statistics*, 23(1), 47-61.
- Garai, S., & Paul, R. K. (2023). Development of MCS based-ensemble models using CEEM-DAN decomposition and machine intelligence. *Intelligent Systems with Applications*, 18, 200202.
- Garai, S., Paul, R. K., Rakshit, D., Yeasin, M., Paul, A. K., Roy, H. S., Barman, S. & Manjunatha, B. (2023). An MRA Based MLR Model for Forecasting Indian Annual Rainfall Using Large Scale Climate Indices. *International Journal of Environment and Climate Change*, 13(5), 137-150.
- Garai, S. (2023). Package 'AllMetrics' Type Package Title Calculating Multiple Performance Metrics of a Prediction Model Version 0.1.0, Repository: <https://cran.r-project.org/web/packages/AllMetrics/index.html>

Examples

```
set.seed(123) # for reproducibility
Y <- rnorm(100, 100, 10)
model_1 <- CEEMDANML::carigaan(Y, ratio = 0.8, n_lag = 4)
model_2 <- CEEMDANML::carigas(Y, ratio = 0.8, n_lag = 4)
model_1_metrics_values <- model_1$Accuracy[,2]
model_2_metrics_values <- model_2$Accuracy[,2]
combined_results <- data.frame(cbind(model_1_metrics_values, model_2_metrics_values))
colnames(combined_results) <- c("model_1", "model_2")
best_model <- best_model_finder(input_df = combined_results)
```

Index

`best_model_finder`, [2](#)