Package 'MARSGWR'

May 9, 2023

Type Package

Title A Hybrid Spatial Model for Capturing Spatially Varying Relationships Between Variables in the Data

Version 0.1.0

Depends R(>= 2.10)

Suggests knitr, rmarkdown, testthat (>= 3.0.0)

Description It is a hybrid spatial model that combines the strength of two widely used regression models, MARS (Multivariate Adaptive Regression Splines) and GWR (Geographically Weighted Regression) to provide an effective approach for predicting a response variable at unknown locations. The MARS model is used in the first step of the development of a hybrid model to identify the most important predictor variables that assist in predicting the response variable. For method details see, Friedman, J.H. (1991). <DOI:10.1214/aos/1176347963>.The GWR model is then used to predict the response variable at testing locations based on these selected variables that account for spatial variations in the relationships between the variables. This hybrid model can improve the accuracy of the predictions compared to using an individual model alone.This developed hybrid spatial model can be useful particularly in cases where the relationship between the response variable and predictor variables is complex and non-linear, and varies across locations.

License GPL (>= 2.0)

Encoding UTF-8

RoxygenNote 7.2.3

Imports stats, qpdf, numbers, earth

NeedsCompilation no

Author Nobin Chandra Paul [aut, cre, cph],
Anil Rai [aut],
Ankur Biswas [aut],
Tauqueer Ahmad [aut],
Dhananjay D. Nangare [aut],
Bhaskar B. Gaikwad [aut]

Maintainer Nobin Chandra Paul <nobin.paul@icar.gov.in>

Repository CRAN

Date/Publication 2023-05-09 09:20:05 UTC

R topics documented:

`MARSGWR_exponential` *MARSGWR: a hybrid model that uses the MARS model for important variable selection and the GWR model for prediction at an unknown location based on the selected variables.*

Description

MARSGWR: a hybrid model that uses the MARS model for important variable selection and the GWR model for prediction at an unknown location based on the selected variables.

Usage

```
MARSGWR_exponential(sp_data, bw, deg, sv, exponential_kernel)
```

Arguments

`sp_data`	A dataframe containing the response variables and the predictor variable, as well as the coordinates of the locations. In the dataframe, first column is the response variable (y), last two columns are coordinates i.e., Latitude and Longitudes and in between them is the set of predictor variables(X's).
`bw`	A numeric value specifying the bandwidth parameter for the GWR model
`deg`	The degree of interactions to be considered in the MARS model
`sv`	Splitting value for dividing the dataset into training and testing set, e.g. 0.8 or 0.7
`exponential_kernel`	Spatial weight function of the GWR model, e.g. exponential_kernel

Value

A list with the following components: - 'Selected_variables': The selected variables from the MARS model - 'GWR_y_pred_train': The GWR predictions at the training locations - 'GWR_y_pred_test': The GWR predictions at testing locations - 'In_sample_accuracy': In sample accuracy measures - 'Out_of_sample_accuracy': Out of sample accuracy measures

References

1. Friedman, J.H. (1991).Multivariate Adaptive Regression Splines. Ann. Statist. 19(1),1-67. <DOI:10.1214/aos/1176347963>. 2. Brunsdon, C., Fotheringham, A.S. and Charlton, M,E. (1996).Geographically weighted regression: a method for exploring spatial non-stationarity. Geogr Anal.28(4),281-298.<DOI:10.1111/j.1538-4632.1996.tb00936.x>.

Examples

```
n<- 100
p<- 5
m<-sqrt(n)
id<-seq(1:n)
x<-matrix(runif(n*p), ncol=p)
e<-rnorm(n, mean=0, sd=1)
xy_grid<-expand.grid(c(1:m),c(1:m))
Latitude<-xy_grid[,1]
Longitude<-xy_grid[,2]
B0<-(Latitude+Longitude)/6
B1<-(Latitude/3)
B2<-(Longitude/3)
B3<-(2*Longitude)
B4<-2*(Latitude+Longitude)/6
B5<-(4*Longitude/3)
y<-B0+(B1*x[,1])+(B2*x[,2])+(B3*x[,3])+(B4*x[,4])+(B5*x[,5])+ e
sp_data<-data.frame(y,x,Latitude,Longitude)
MARSGWR_exp<-MARSGWR_exponential(sp_data,5,3,0.7,exponential_kernel)
```

`MARSGWR_gaussian` *MARSGWR: a hybrid model that uses the MARS model for important variable selection and the GWR model for prediction at an unknown location based on the selected variables.*

Description

MARSGWR: a hybrid model that uses the MARS model for important variable selection and the GWR model for prediction at an unknown location based on the selected variables.

Usage

```
MARSGWR_gaussian(sp_data, bw, deg, sv, gaussian_kernel)
```

Arguments

`sp_data`	A dataframe containing the response variables and the predictor variable, as well as the coordinates of the locations. In the dataframe, first column is the response variable (y), last two columns are coordinates i.e., Latitude and Longitudes and in between them is the set of predictor variables(X's).

bw	A numeric value specifying the bandwidth parameter for the GWR model. It can be noted that, optimum bandwidth value can vary depending on the specific dataset and bandwidth parameter depends on the spatial pattern of the data
deg	The degree of interactions to be considered in the MARS model
sv	Splitting value for dividing the dataset into training and testing set, e.g. 0.8 or 0.7
gaussian_kernel	Spatial weight function of the GWR model, e.g. gaussian_kernel

Value

A list with the following components: - 'Selected_variables': The selected variables from the MARS model - 'GWR_y_pred_train': The GWR predictions at the training locations - 'GWR_y_pred_test': The GWR predictions at testing locations - 'In_sample_accuracy': In sample accuracy measures - 'Out_of_sample_accuracy': Out of sample accuracy measures

References

1. Friedman, J.H. (1991).Multivariate Adaptive Regression Splines. Ann. Statist. 19(1),1-67. <DOI:10.1214/aos/1176347963>. 2. Brunsdon, C., Fotheringham, A.S. and Charlton, M,E. (1996).Geographically weighted regression: a method for exploring spatial non-stationarity. Geogr Anal.28(4),281-298.<DOI:10.1111/j.1538-4632.1996.tb00936.x>.

Examples

```
n<- 100
p<- 5
m<-sqrt(n)
id<-seq(1:n)
x<-matrix(runif(n*p), ncol=p)
e<-rnorm(n, mean=0, sd=1)
xy_grid<-expand.grid(c(1:m),c(1:m))
Latitude<-xy_grid[,1]
Longitude<-xy_grid[,2]
B0<-(Latitude+Longitude)/6
B1<-(Latitude/3)
B2<-(Longitude/3)
B3<-(2*Longitude)
B4<-2*(Latitude+Longitude)/6
B5<-(4*Longitude/3)
y<-B0+(B1*x[,1])+(B2*x[,2])+(B3*x[,3])+(B4*x[,4])+(B5*x[,5])+ e
sp_data<-data.frame(y,x,Latitude,Longitude)
MARSGWR_gau<-MARSGWR_gaussian(sp_data,5,3,0.7,gaussian_kernel)
```

Index