p-ISSN 2338 – 3216
e-ISSN 2528 – 1070

JURNAL STATISTIKA
UNIVERSITAS MUHAMMADIYAH SEMARANG
12(2) 2024: 19-28

MODELING OF POVERTY INDICATORS IN EAST JAVA PROVINCE USING BOOTSTRAP AGGREGATING MULTIVARIATE ADAPTIVE REGRESSION SPLINE (BAGGING MARS)

Danu Priambodo[1*], Rochdi Wasono[2], M Al Haris[3]
[1,2,3] Department of Statistics, Universitas Muhammadiyah Semarang, Semarang, Indonesia

***e-mail**: *danupriambodo18@gmail.com*

Article Info:
Received: September 17th, 2024
Accepted: October 24th, 2024
Available Online: December, 31th 2024

Keywords:
Bagging, GCV, Poverty, MARS

Abstract: Poverty is defined as a situation in which a person falls below the minimum standard value line. Poverty is considered a multidimensional phenomenon with many indicators influencing it, necessitating modeling to identify these indicators. This research uses the Multivariate Adaptive Regression Spline (MARS) with Bootstrap Aggregating. MARS is a nonparametric regression method capable of handling high-dimensional data. The best model produced by MARS is a combination of BF=24, MI=1, and MO=0 with a GCV of 9.231184. Bagging was then applied to the initial dataset with 35, 45, 50, 75, and 100 bootstrap replications. The best model was produced by MARS Bagging with 45 replications, achieving a GCV of 3.84492. The GCV value obtained by bagging MARS is smaller than that of MARS alone. This demonstrates that bagging can reduce GCV and increase accuracy, making this method suitable for this research.

1. INTRODUCTION

Development is a step towards a change towards a better future in order to realize a society that has a sense of justice, competence, progress and prosperity. Development includes strategies in overcoming problems in each region. Therefore, one of the important parameters for the success of development is the depreciation of the rate of the number of poor people [1]. Poverty is a problem that should not be underestimated. Poverty is a social problem that arises from economic factors where a person is unable to meet his or her needs, especially [2].

Indonesia as a developing country certainly always faces this problem. Poverty is included in the routine agenda in Indonesia which must always be cared for, because poverty is definitely a problem faced in achieving national development goals, namely minimizing the number of poor people in order to create justice and prosperity for society. The view on poverty can be said that poverty is a multidimensional phenomenon where the indicators that are thought to influence poverty [3]. Almost all regions in Indonesia have poor people. Especially in areas that are centers of the economy and densely populated such as Java Island [4]. According to data from BPS, the province with the largest number of poor people in 2023 is located on the island of Java, in East Java Province with a population of 4.18 million poor people. With the still high level of poverty, there must be indicators that influence poverty itself. Therefore, it is suitable to do modeling to find out what indicators influence poverty [5].

[DOI: 10.14710/JSUNIMUS.12.2.2024.19-28]

There is a method in statistics that can be implemented to model indicators that influence poverty, namely the Multivariate Adaptive Regression Spline (MARS) method. MARS is a type of nonparametric regression approach which was previously a development of the Recursive Partition Regression (RPR) method combined with the spline method. Because MARS is included in nonparametric regression, it does not rely on a particular assumption and this method is suitable for handling problems with data that has high dimensions (predictor variables as many as $3 \leq i \leq 20$). GCV is the best criterion in choosing the best model. The selection of the best or optimum model is chosen if the GCV value in the model produces the smallest value compared to other models [6].

There is a technique to improve the accuracy of the MARS model, namely using resampling techniques, one of the commonly used resampling techniques is Bootstrap Aggregating (Bagging). This technique can be used to improve the stability and strength of predictions. By combining the MARS method with the Bagging technique, it can provide an increase in the predictive accuracy of the MARS model [7]. This can be strengthened by previously conducted research which explains that MARS Bagging can increase predictive accuracy in the MARS model [8].

Several studies using the Bagging MARS method have been carried out, including those conducted by Jamilah and her friends who used the Bagging MARS method to model Gross Regional Domestic Product (GRDP) in Central Java Province, producing the best MARS model in the combination of BF=6, MI=1, MO=0 with a GCV value of 5667.66 and Bagging MARS obtained the smallest GCV of 2258.61 using 55 replications [9]. Then the research conducted by Mijayanti and Helma which used the Bagging MARS method which was applied to model the Gross Regional Domestic Product (PDRB) in West Sumatra Province using a combination of BF=8, MI=3, MO=0 produced the best MARS model with a GCV of 7.3686 and Bagging MARS obtained the smallest GCV of 5.2562 using 33 replications [10]. And the research conducted by Hasyim and his friends who used the Bagging MARS method which was applied to analyze the research performance of lecturers at Private Universities produced the best MARS model in the combination of BF=66, MI=3, MO=2 with a GCV value of 0.018 and Bagging MARS obtained the smallest GCV of 0.048 using 35 replications [11].

Based on the description above, this research will be conducted to model poverty indicators using the MARS method to produce the best model and use the resampling method, namely Bagging MARS, to increase the predictive accuracy of the MARS model.

2. LITERATURE REVIEW

2.1. Regression Analysis

Regression analysis is one of the simple methods carried out by testing the relationship between several variables such as response variables and predictor variables. Regression analysis is divided into two types based on the pattern of data relationships between two variables, known as regression curves, namely parametric regression and nonparametric regression [12]. In the parametric approach, there is an assumption on the form of a predetermined model, where the form of the regression curve is already known or produces a pattern. While nonparametric regression is a statistical method used when the pattern of the relationship between the response variable and its predictor is not identified. If there is no information about the pattern and function of the regression, such as a data pattern that does not have a specific form [13].

2.2. Nonparametric Regression

Nonparametric regression is a technique in statistics that can be used when you want to know whether there is a relationship between the dependent variable and the independent

[DOI: 10.14710/JSUNIMUS.12.2.2024.19-28]

variable where the pattern of the regression curve is unknown or there is no complete historical information related to the data curve pattern. The model in nonparametric regression is generally written as follows: [14].

$$y_i = f_{(xi)} + \varepsilon_i, i = 1,2, \dots, n \tag{1}$$

with,

y_i : dependent variable on the *i*-th response
$f_{(xi)}$: unknown regression function
ε_i : error of random variables

In nonparametric regression. Commonly used models in nonparametric regression include Histogram, Neural Network (NN), Kernel, k-NN, Wavelets, MARS, Spline and others. [12].

2.3. Multivariate Adaptive Regression Spline (MARS)

MARS is a method that uses a nonparametric regression analysis approach which was first introduced in 1991 by Friedman. MARS has a model that can be used as a solution to data that has high-dimensional problems, where data with predictor variables totaling 3 ≤ i ≤ 20 variables. In general, the MARS model equation is written as follows [6]:

$$\hat{f}(x) = \alpha_0 + \sum_{m=1}^{M} \alpha_m \prod_{k=1}^{K_m} [S_{km}(x_{v(k,m)} - t_{km})]_+ \tag{2}$$

with,
α_0 = constant coefficient
α_m = coefficient of the *m*-th basis function
M = maximum basis function
K_m = degree of interaction
S_{km} = has a value of ±1
$x_{v(k,m)}$ = *v*-th explanatory variable, *k*-th interaction and *m*-th basis function
t_{km} = knot value of the explanatory variable $x_{v(k,m)}$

When building a MARS model, there are things that must be considered, namely [6]:

1. Knot, is the end point of a regression line and the starting point of the next regression line. At each knot point on the basis function between each region must be continuous (continuous) which is defined by the relationship between the response variable and the predictor. Minimum Observation (MO) is the minimum distance between knots recommended values of 0.1,2 and 3. If the MO value is more than 3, it can damage the accuracy and flexibility of the model.
2. Basis Function is called a parametric function that interprets each region to explain the relationship between the dependent variable and the independent variable. The recommended number of basis functions is two to four times the number of independent variables.
3. Maximum Interaction is the relationship between one variable and another. Maximum Interaction (MI) or the recommended number of interactions is 1,2 and 3. If the use of MI is more than 3, this will produce a more complex model and it is more difficult to interpret it and can result in an increase in the GCV value.

[DOI: 10.14710/JSUNIMUS.12.2.2024.19-28]

In the MARS method, Generalized Cross Validation is the best criterion in selecting the best model. The GCV function is written as follows [6]:

$$GCV(M) = \frac{\frac{1}{n}\sum_{i=1}^{n}\left[y_i - \hat{f}_M(x_i)\right]^2}{\left[1 - \frac{\tilde{C}(M)}{n}\right]^2} = \frac{\frac{1}{n}\,RSS}{\left[1 - \frac{\tilde{C}(M)}{n}\right]^2} \tag{3}$$

with,

n = amount of data
y_i = response variable, i = 1,2,..,n
x_i = explanatory variable, i = 1,2,..,n
M = maximum number of basis functions
$\hat{f}_M(x_i)$ = estimated value of the response on the i-th data
$\tilde{C}(M)$ = $C(M) + dM$; d = the best value of d is in the interval 2 to 4
$C(M)$ = $e[B(B^TB)^{-1}B^T]$; B is a matrix of M basis functions.

2.4. Bootstrap Aggregating

Bootstrap is a method with nonparametric estimation used to estimate parameters in a distribution. Bootstrap can provide the best estimate of the estimation results with little bias. [15]. In this technique, resampling is done on the data by returning the sample of observation results with bootstrap replication B times. Determining the size of the B value certainly varies because the size of the B value will produce different results at each step in the analysis. So the flow or steps taken when doing bootstrapping are as follows [15]:

1. Interpret the data sample *x*, where x is a data sample with size n containing the observation data vector $x_i = (x_1, x_2, \dots, x_n)$.
2. Take a sample of *x n* times to produce a new data sample called (x^*). The data sample (x^*) contains the original data set, some data may not appear at all or only appear once or twice.
3. Perform the second step by repeating B times, so that the bootstrap data set ($x^*_1, x^*_2, \dots, x^*_n$) is obtained. Each bootstrap sample is a random sample that is independent.

This technique is suitable for high-dimensional data that is difficult to model. The working principle of the bagging method lies in the diversity of predictor variables, a process known as predictor bagging. The initial concept of bagging involves bootstrap resampling to obtain several versions of the predictor. When combined, these versions should produce a better model than a single predictor built to solve an equivalent problem [16].

3. METHODOLOGY

This study uses secondary data on poverty in East Java Province in 38 districts/cities in East Java Province which comes from the BPS website which can be accessed via https://jatim.bps.go.id. The steps in this research were carried out as follows:

1. Prepare data.
2. Know the general picture of poverty in East Java in 2023 by conducting descriptive analysis.
3. Identify data patterns.

[DOI: 10.14710/JSUNIMUS.12.2.2024.19-28]

4. Combine the magnitude of the Basis Function (BF), Maximum Interaction (MI) and Minimum Observation (MO) as follows:
   a. The Basis Function (BF) that may be used is two to four times the number of independent variables. In this study, there are 8 independent variables used, so the maximum basis function is 16, 24 and 32.
   b. The Maximum Interaction (MI) used is 1, 2 and 3. Because if it exceeds 3 it will produce a very complex model which makes it difficult to interpret.
   c. Minimum Observation (MO) the minimum distance allowed is 0,1,2 and 3.
5. Determine the best MARS model based on the minimum GCV value.
6. Interpret the best MARS model and what variables affect poverty in the model.
7. Perform bagging on the original data set with 35, 45, 50, 75 and 100 bootstrap replications.
8. Perform MARS modeling on each B replication.
9. Obtain the bagging GCV value based on the average GCV value on each B replication.
10. Interpret the obtained MARS bagging GCV value.

4. RESULTS AND DISCUSSION

This analysis aims to understand the characteristics of poverty in East Java in 2023 based on the indicators it influences.

Table 1. Descriptive Statistics of Variables

Variable	Description	Min	Max	Mean	St.Dev
Y	Percentage of poor population	3,31	21,76	10,29	4,321
X_1	Percentage of population over 15 years old who have not completed elementary school	3,41	48,80	21,27	11,143
X_2	Literacy Rate of poor population aged 15-55 years	86,03	100	96,78	3,182
X_3	School Participation Rate at age 13-15 years	92,49	99,53	97,88	1,530
X_4	Percentage of population over 15 years old who are not working	28,14	58	39,95	5,779
X_5	Life Expectancy	67,60	74,91	72,42	1,982
X_6	Percentage of poor households using clean water	50,30	98,18	84,59	12,182
X_7	Percentage of poor households that have their own or shared defecation	79,29	100	96,12	4,811
X_8	Percentage of poor households that receive Jamkesmas services	22,83	83,33	45,54	12,622

Based on Table 1, The variable that has the lowest value among the other variables is variable Y (Percentage of poor population) with a value of 3.31%, while the variable that has a maximum value of 100% is in the variable X_2 (Literacy Rate of Poor Population 15-

[DOI: 10.14710/JSUNIMUS.12.2.2024.19-28]

55 years) and X_7 (Percentage of poor households that have their own or shared defecation. And variable X_8 has the highest standard deviation value among the other variables, namely 12.622%. This indicates that the diversity of data between regencies/cities in East Java in 2023 is in the variable percentage of poor households that receive Jamkesmas services (X_8).

4.1. Data Plot

The next step, the process carried out in this study is to look at the plot between the response variable and the eight predictor variables to identify whether there is a pattern of relationship between the response variable and the eight variables.

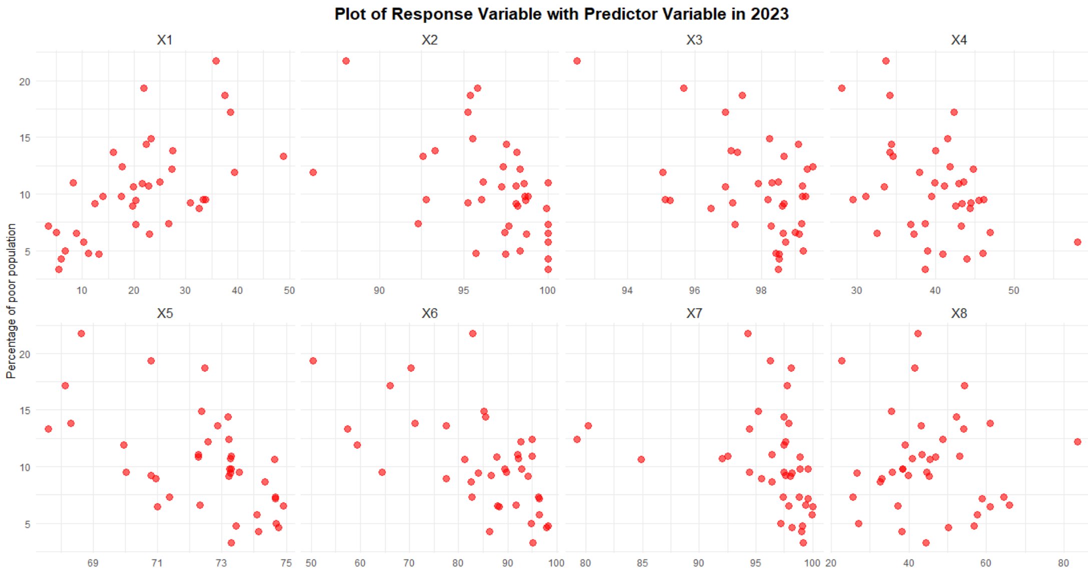

Fig 1. Response and Predictor Variable Plots

Based on Figure 1, The plot results of the percentage of poor people and indicator variables that affect poverty in East Java in 2023 show that each independent variable has a different pattern and does not show a tendency to form a particular pattern. Because of the limited indication of the shape of the pattern and the unclear nature of several patterns of relationships between response variables and predictor variables, the considerations for choosing a nonparametric regression approach in modeling the data. The nonparametric regression approach used in this study is Multivariate Adaptive Regression Spline because the data used has a high dimension.

4.2. Multivariate Adaptive Regression Spline Modeling

Next is to determine the MARS model using the stepwise method, namely the forward method and the backward method. The forward method is used to obtain α and the maximum number of basis functions. While the Backward method is used to determine the best model based on the lowest Generalize Cross Validation (GCV) value in the existing model.

The process of modeling poverty indicators in East Java is carried out by trial and error by combining the number of basis functions (BF), maximum interaction (MI) and minimum observation (MO), namely BF = 16,24,36 MI = 1,2,3 MO = 0,1,2,3 so that the number of models produced is 36 models. From each of these experiments, the GCV value will be obtained and the predictor variables included in the model will be obtained.

[DOI: 10.14710/JSUNIMUS.12.2.2024.19-28]

Table 2. Best Combination Results

No	BF	MI	MO	GCV	R^2
1	16	1	0	10.025618	0.6330367
2	16	1	1	11.086684	0.5941990
3	16	1	2	10.812208	0.8183895
4	16	1	3	11.276490	0.5872516
5	16	2	0	10.318677	0.8097794
6	16	2	1	9.913085	0.6714202
7	16	2	2	12.826010	0.6438712
8	16	2	3	12.739949	0.6462607
9	16	3	0	10.953373	0.6958671
10	16	3	1	9.913085	0.6714202
11	16	3	2	12.826010	0.6438712
12	16	3	3	12.739949	0.6462607
13*	**24**	**1**	**0**	**9.231184**	**0.8449456**
14	24	1	1	10.350436	0.8261458
15	24	1	2	12.777567	0.4038270
16	24	1	3	12.695685	0.5933318
17	24	2	0	10.318677	0.8097794
18	24	2	1	9.913085	0.6714202
19	24	2	2	12.826010	0.6438712
20	24	2	3	12.739949	0.6462607
21	24	3	0	10.953373	0.6958671
22	24	3	1	9.913085	0.6714202
23	24	3	2	12.826010	0.6438712
24	24	3	3	12.739949	0.6462607
25*	32	1	0	9.231184	0.8449456
26	32	1	1	10.350436	0.8261458
27	32	1	2	12.777567	0.4038270
28	32	1	3	12.695685	0.5933318
29	32	2	0	10.318677	0.8097794
30	32	2	1	9.913085	0.6714202
31	32	2	2	12.826010	0.6438712
32	32	2	3	12.739949	0.6462607

*= Nilai GCV terrendah

In table 2, two models are obtained from 36 models that have the same lowest GCV. The criteria for choosing the best model is to find the lowest GCV value. If models are obtained that have the same GCV value, then the next consideration is the highest R2 value. However, if the models also have the same rsquare value, then the next decision is to look at the model with the smallest combination. So the best model obtained based on the smallest combination is BF = 24, MI = 1, MO = 0 with a model like:

$$\hat{Y} = -1{,}520 + 0{,}304BF1 - 0{,}421BF2 + 7{,}656BF3 + 2{,}823B \\ -3{,}758BF5 + 3{,}344BF6 - 5{,}318BF7 + 0{,}433BF8 \quad (4)$$

[DOI: 10.14710/JSUNIMUS.12.2.2024.19-28]

with,
BF1 = max (0, X_1 – 5.05) , BF2 = max (0, X_1 – 22.3), BF3 = max (0, 95.05 – X_3)
BF4 = max (0, X_3 – 95.05), BF5 = max (0, X_3 – 96.91), BF6 = max (0, X_5 – 70.79)
BF7 = max (0, X_5 – 72.57), BF8 = max (0, 82.63 – X_6)

4.3. Bootstrap Aggregating Multivariate Adaptive Regression Spline (Bagging MARS) Modeling

Next, form a Bagging MARS model with bootstrap replications of 35, 45, 50, 75 and 100, and use the best parameters of the MARS model, BF = 24, MI = 1, MO = 0. Each replication will produce the same number of models as the number of replications. Then the GCV value is averaged, resulting in the GCV Bagging MARS value for each replication.

Table 3. MARS Bagging Results

Replication	GCV	R^2
35	4.16604	0.941232
45*	**3.84492**	**0.94611**
50	4.43979	0.912522
75	4.40268	0.93211
100	4.16206	0.93041

Based on table 3, Of the 35, 45, 50, 75 and 100 replications, the lowest GCV value is in the 45th replication, which is 3.84492. The best Bagging MARS model with 45 replications is in the 45th model with a GCV value of 0.28292 and the model is as follows:

$$\hat{Y} = 12{,}007 - 0{,}455BF1 - 0{,}196BF2 + 0{,}543BF3 + 0{,}703\mathrm{BF4} + 0{,}674BF5 + 3{,}975BF6 + 0{,}822BF7 - 0{,}571BF8 + 0{,}172BF9 - 0{,}576BF10 - 0{,}827BF11 + 0{,}431BF12 - 0{,}128BF13 + 0{,}224BF14 + 4{,}244BF15 - 0{,}049BF16 - 0{,}018BF17 \quad (5)$$

with,
BF1 = max (0, 21.56 – X_1), BF2 = max (0, X_1 – 21.56), BF3 = max (0, 98.15 – X_2),
BF4 = max (0, X_2 – 98.15), BF5 = max (0, 98.2 – X_3), BF6 = max (0, X_3 – 98.2)
BF7 = max (0, X_4 – 37.99), BF8 = max (0, X_4 – 40), BF9 = max (0, 44.44 – X_4),
BF10 = max (0, X_4 – 44.44), BF11 = max (0, X_6 – 82.5), BF12 = max (0, X_6 – 86.36)
BF13 = max (0, 91.73 – X_6), BF14 = max (0, 98.86 – X_7), BF15 = max (0, X_7 – 98.86),
BF16 = max (0, 44.35 – X_8), BF17 = max (0, X_8 – 44.35)

With an example of interpretation at BF1 = max (0, 21.56 – X_1) with a coefficient of -0.455 in the model, which means that when the percentage of the population over 15 years old who have not completed elementary school (X_1) is less than 21.56%, every 1% decrease will reduce the percentage of poor people by 0.455%.

5. CONCLUSION

The characteristics of the poverty rate in East Java in 2023 are uneven, where the Regency area has a higher percentage than the City area. The best model for the Regency/City poverty indicator in East Java 2023 with the Bagging MARS method using 45 replications obtained a GCV value of 3.84492. Indicators that have a significant impact on poverty by Regency/City in East Java 2023 using the Bagging MARS model from the most influential, namely

[DOI: 10.14710/JSUNIMUS.12.2.2024.19-28]

Percentage of population over 15 years old who have not completed elementary school (X_1), Percentage of poor households using clean water (X_6), Percentage of poor households that receive Jamkesmas services (X_8), School Participation Rate at age 13-15 years (X_3), Percentage of poor households that have their own or shared defecation (X_7), Percentage of population over 15 years old who are not working (X_4) and Literacy Rate of poor population aged 15-55 years (X_2).

REFERENCES

[1] E. W. Azizah, Sudarti, and H. Kusuma, "Pengaruh Pendidikan, Pendapatan Perkapita Dan Jumlah Penduduk Terhadap Kemiskinan Di Provinsi Jawa Timur," *Jurnal Ilmu Ekonomi*, vol. Vol.2, no. Jilid 1, pp. 167–180, Aug. 2018.

[2] L. Leonita and R. S. Kurnia, "Pengaruh Pdrb, Pengangguran Dan Pembangunan Manusia Terhadap Kemiskinan Di Indonesia," *Jurnal Ekonomi, Manajemen dan Akuntansi*, vol. 3, no. 2, pp. 1–8, Sep. 2019, doi: 10.24269/iso.v3i2.252.

[3] R. F. Saragih, P. R. Silalahi, and K. Tambunan, "Pengaruh Indeks Pembangunan Manusia, Tingkat Pengangguran Terbuka Terhadap Tingkat Kemiskinan di Indonesia Tahun 2007–2021," *Jurnal Pendidikan, Sosial dan Humaniora*, vol. Vol.1, no. No.2, Feb. 2022.

[4] C. Mustika, "Pengaruh PDB dan Jumlah Penduduk Terhadap Kemiskinan di Indonesia Periode 1990-2008," *Jurnal Paradigma Ekonomika*, vol. 1, no. 4, 2011.

[5] R. B. Budhijana, "Analisis Pengaruh Pertumbuhan Ekonomi, Index Pembangunan Manusia (IPM) dan Pengangguran Terhadap Tingkat Kemiskinan Di Indonesia Tahun 2000-2017," *Jurnal Ekonomi Manajemen dan Perbankan*, vol. Vol.5, no. No.1, Mar. 2020, doi: 10.35384/jemp.v5i1.170.

[6] J. H. Friedman, "Multivariate Adaptive Regression Splines," *The Annals of Statistics*, vol. 19, no. 1, pp. 1–67, Mar. 1991, Accessed: Feb. 26, 2024. [Online]. Available: http://links.jstor.org/sici?sici=0090-5364%28199103%2919%3A1%3C1%3AMARS%3E2.0.CO%3B2-D

[7] W. Pintowati and B. Widjanarko O, "Pemodelan Kemiskinan di Propinsi Jawa Timur dengan Pendekatan Multivariate Adaptive," *JURNAL SAINS DAN SENI ITS*, vol. Vol 1, no. No. 1, Sep. 2012, doi: http://dx.doi.org/10.12962/j23373520.v1i1.2072.

[8] L. Nahriyah, "Bootstrap Aggregating Multivariate Adaptive Regression Splines (Bagging Mars) Untuk Klasifikasi Pasien HivAids Di Kabupaten Pasuruan," Institut Teknologi Sepuluh Nopember, Surabaya, 2017.

[9] M. Jamilah An H, A. Rusgiyono, and D. Safitri, "Pemodelan Produk Domestik Regional Bruto (PDRB) di Provinsi Jawa Tengah Menggunakan Bootstrap Agregasi Multivariate Adaptive Regression Splines (Bagging MARS)," *Jurnal Gaussian*, vol. Vol.8, no. No.1, pp. 139–148, 2019, doi: https://doi.org/10.14710/j.gauss.8.1.139-148.

[10] T. Mijayanti and Helma, "Bootstrap Aggregating Multivariate Adaptive Regression Splines (Bagging MARS) dan Penerapannya pada Pemodelan Produk Domestik

Regional Bruto (PDRB) di Provinsi Sumatera Barat," *Jurnal Of Mathematics UNP*, vol. Vol 6, no. No. 4, Dec. 2021, doi: http://dx.doi.org/10.24036/unpjomath.v6i4.12233.

[11] M. Hasyim *et al.*, "Bootstrap Aggregating Multivariate Adaptive Regression Splines (Bagging MARS) to Analyse the Lecturer Research Performance in Private University," *J Phys Conf Ser*, vol. 1114, no. 1, Dec. 2018, doi: 10.1088/1742-6596/1114/1/012117.

[12] I. N. Budiantara, "Penelitian Bidang Regresi Spline Menuju Terwujudnya Penelitian Statistika Yang Mandiri dan Berkarakter," *Seminar Nasional FMIPA Undiksha*, pp. 9–28, 2011.

[13] M. N. Rahmaniah, Y. Novia N, and I. Purnamasari, "Bootstrap Aggregating Multivariate Adaptive Regression Splines (Studi Kasus : Identifikasi Komponen Penciri Akreditasi Sekolah/Madrasah Pada Tingkat SD/MI di Provinsi Kalimantan Timur Tahun 2015)," *Jurnal EKSPONENSIAL*, vol. 7, no. 2, Nov. 2016.

[14] R. L. Eubank, *Non Parametric Regression and Spline Smoothing*, Second edition., vol. 157. 1991.

[15] Bradley. Efron and Robert. Tibshirani, *An introduction to the bootstrap*. Chapman & Hall, 1994.

[16] L. Breiman, "Bagging Predictors," *Kluwer Academic Publishers*, vol. 24, pp. 123–140, 1996, doi: https://doi.org/10.1007/BF00058655.