International Journal of Academic Management Science Research (IJAMSR)
ISSN: 2643-900X
Vol. 8 Issue 5 May - 2024, Pages: 25-33

Modeling Gender Inequality Index (GII) in Nusa Tenggara Using A Multivariate Adaptive Approach Regression Splines

Nurul Izzah[1], Ardi Kurniawan[2*], Elly Ana[3]

[1]Department of Mathematics, Faculty of Science and Technology, Universitas Airlangga, Surabaya, Indonesia, nurul.izzah-2020@fst.unair.ac.id

[2]Department of Mathematics, Faculty of Science and Technology, Universitas Airlangga, Surabaya, Indonesia,

ardi-k@fst.unair.ac.id

[3]Department of Mathematics, Faculty of Science and Technology, Universitas Airlangga, Surabaya, Indonesia,

elly-a@fst.unair.ac.id

*Corresponding Author: ardi-k@fst.unair.ac.id

Abstract: *Gender inequality is one of the compelling issues that has been extensively discussed over the past few decades. In Indonesia, societal perception of women tends to relegate them to a secondary status. This limitation constrains women's opportunities and benefits in development. Efforts and development programs are evaluated through a gender lens using the Gender Inequality Index (GII). In the Sustainable Development Goals (SDGs), gender issues are integrated into Goal 5, aiming for gender equality and empowerment of all women and girls. This goal aligns with the fact that Indonesia has the highest GII among the Association of Southeast Asian Nations (ASEAN) due to gender disparities across its regions. According to data published by the Central Statistics Agency (BPS) (2021), the province with the highest GII is West Nusa Tenggara at 53.1%. The objective of this study is to describe and model the GII in the province of Nusa Tenggara, as well as interpret the results of the best model obtained. The method employed is a nonparametric regression approach, specifically Multivariate Adaptive Regression Splines (MARS). The research findings indicate that the best model obtained consists of a combination of 14 basis functions, a maximum of 2 interactions, and a minimum of 1 observation between knots. From this model, the predictor variables most influential on the response variable, ranked by importance, are Women's Participation in Parliament, Proportion of Women Ever Married aged 15-49 Years who Gave Birth to their First Live Child at Under 20 Years Old, Labor Force Participation Rate, Years of Schooling Expectancy, and Per Capita Expenditure. The best MARS model yields a Generalized Cross Validation value of 0.002; R^2 of 0.945, and Mean Square Error of 0.000613. It is hoped that these research findings will contribute to efforts to reduce gender inequality in the province of Nusa Tenggara and serve as a basis for the development of more effective policies in achieving gender equality..*

Keywords: Modeling, Gender Inequality Index, Spline, MARS

1. INTRODUCTION

Gender is a concept referring to the roles and relationships between men and women, determined by socio-cultural, political, and economic factors rather than biological differences. Gender equality implies equal status, conditions, and potentials for both genders to exercise their rights and contribute to national development (Hubeis, 2010). Gender issues are pivotal in development, with gender understood as socially constructed differences in roles and behaviors (Giyono & Maemunah, 2021). Supported by Utaimingsih (2017), gender is socially constructed, influenced by factors such as geography, economics, culture, politics, and ideology. Gender equality is crucial for effective governance and poverty reduction (Resroeningroem, 2018), yet societal perceptions often relegate women to secondary roles, limiting their participation in development. Gender-based development programs are evaluated using the Gender Inequality Index (IKG), which reflects human development failures due to gender inequality. Gender inequality, a significant issue globally (Carlsen, 2020), not only affects social and political aspects but also the economy. The World Bank (2018) predicts substantial losses due to gender inequality, with potential global wealth increases if women earned equally to men. Gender equality is a Sustainable Development Goal (SDG) aiming to eliminate discrimination, violence, and harmful behavior against women (UNDP, 2015). Indonesia, with high IKG rates, emphasizes the importance of regional gender, Despite progress in Java and Bali, research focusing on reducing IKG beyond these provinces, especially in Nusa Tenggara, is needed (BPS, 2021). Statistical analysis, particularly regression, is crucial for concrete insights into IKG (Massaid, 2019). Multivariate Adaptive Regression Splines (MARS) is a nonparametric method capable of capturing variable interactions, crucial for understanding factors influencing gender inequality (Otok, 2010). Previous studies on IKG factors emphasize the need for better government policies regarding women's empowerment in Indonesia (Yolanda, 2022) Therefore, this research aims to evaluate and develop local government policies, focusing on high IKG cases, particularly in West and East Nusa Tenggara Provinces. Key indicators influencing IKG include factors related to health, education, economy, gender orientation, and participation in the workforce (BPS, 2021). The selection of MARS for IKG modeling aims to provide the best model considering significant factors, hoping to contribute to reducing IKG in Nusa Tenggara Province.

2. RESEARCH METHODS

2.1 Types of Research

The type of research used in this study is quantitative research. Quantitative research methods are one type of research whose specifications are systematic, planned and clearly structured from the beginning to the making of the research design. quantitative research is a type of research that produces findings obtained using statistical procedures or other means of quantification (measurement). Quantitative research methods are used to examine certain populations or samples, sampling techniques are generally carried out randomly, data collection using research instruments, quantitative or statistical data analysis with the aim of testing predetermined hypotheses.

2.2 Data and Data Sources

This research employs a regression method with a non-parametric approach, namely the Multivariate Adaptive Regression Splines (MARS) method. The data used in this study are secondary data obtained from the official website of the Central Statistics Agency of the Republic of Indonesia. The data includes the Gender Inequality Index, the Proportion of Women Ever Married Aged 15-49 Years at First Birth Under Age 20, the Average Length of Schooling (AHH), Per Capita Expenditure (PPK), Women's Participation in Parliament (KPP), and Labor Force Participation Rate (TPAK). The observational units in this research are 32 Districts/Cities in East Nusa Tenggara and West Nusa Tenggara Provinces in the Year 2022.

2.3 Research Variables

Research variables are everything determined by the researcher in the form of attributes or objects that have certain variations to be studied and then conclusions drawn. Research variables are divided into predictor variables and response variables. Predictor variables are variables that are thought to cause influence on the response variable. Meanwhile, the response variable is a variable that appears as a result or impact of the influence caused by the predictor variable. The research variables used in this research are presented.

Table 1. Research Variables

Variable		Information	Scale
Response	Y	Gender Inequality Index (GII)	Ratio
Predictor	X_1	Proportion of women aged 15-49 who have ever been married and gave birth to their first child before the age of 20	Ratio
	X_2	Average Years of Schooling	Ratio
	X_3	Per-Capita Expenditure	Ratio
	X_4	Women's Participation in Parliament	Ratio
	X_5	Labor Force Participation Rate	Ratio

2.4 Steps of Analysis

The data analysis steps conducted in this study are as follows:

1. Describing the Gender Inequality Index (GII) along with its suspected influencing factors in all districts/cities in Nusa Tenggara by presenting each variable data in the form of diagrams using Microsoft Excel software.
2. Creating scatterplots between the response variable and each predictor variable with the help of SPSS software. This is used to facilitate reading the data distribution visually and as an initial detector for the use of nonparametric approaches.
3. Modeling the Gender Inequality Index (GII) data based on suspected influencing factors in Nusa Tenggara using the Multivariate Adaptive Regression Splines (MARS) method. The steps in modeling it are as follows:
   a) Inputting the data into the software consisting of one response variable and 5 predictor variables.
   b) Retrieving data from the software previously stored into the MARS software.
   c) Determining the value of Basis Function (BF), ranging from ten to twenty.

d) Determining the Maximum Interaction (MI), namely one, two, and three. This determination assumes that if the maximum interaction value exceeds three, it will result in a too complex model.

e) Determining the Minimum Observation (MO), namely zero, one, two, and three.

f) Combining the BF, MI, MO values until obtaining the best model with criteria of minimum Generalized Cross Validation (GCV) value, maximum R^2, and minimum Mean Square Error (MSE).

4. Simultaneously testing all basis function coefficient to determine whether the basis functions together influence the Gender Inequality Index (GII) variable using the $F-test$ statistic.
5. Partially testing each basis function coefficient to determine whether each basis function has a significant influence on the Gender Inequality Index (GII) variable using the t-test statistic.
6. Performing assumption testing for residuals, namely normality test, mean 0, and variance σ^2 using SPSS software.
7. Interpreting the model produced based on the factors that have been significantly tested against the Gender Inequality Index (GII) in Nusa Tenggara by summarizing the results obtained and explaining the meaning contained in the values of each basis function (BF) listed in the best model results.

3. RESULT AND DISCUSSION

3.1 Descriptive Statistics

Descriptive statistics serve as the initial stage in data exploration used to describe and illustrate the research object in general. Descriptive statistics can be presented in several forms, such as bar charts, line charts, pie charts, histograms, polygons, scatterplots, and so on. The use of descriptive statistics follows the researcher's needs by considering the usefulness of each form of data presentation.

The descriptive statistics used in this study are presented in the form of bar charts, scatterplots, measures of central tendency, and data dispersion. Bar charts aim to describe the Gender Inequality Index (GII) along with its suspected influencing factors in Nusa Tenggara. Scatterplots aim to determine the pattern of research data distribution as well as an initial detection for the use of nonparametric methods. If there is no pattern such as linear, quadratic, and cubic on the scatterplot between the response variable and each predictor variable, then nonparametric regression methods are considered and recommended for use in data analysis. The presentation of research data in the form of bar charts is assisted by Microsoft Excel software, and scatterplots with the help of an application. The bar charts and scatterplots between the response variable, namely the proportion of women ever married aged 15-49 who gave birth to their first child under the age of 20, the average years of schooling, per capita expenditure, women's involvement in parliament, and the labor force participation rate are presented as follows.

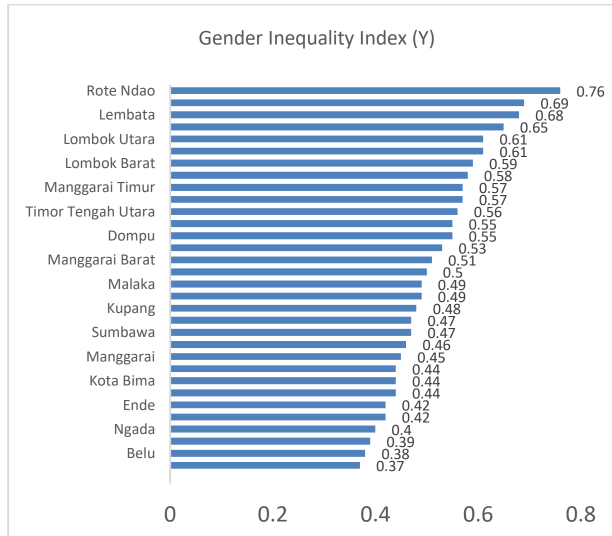

Figure 1. *Variable of bar chart Y*

Based on **Figure 1.**, it can be seen that the highest Gender Inequality Index in Nusa Tenggara is in Rote Ndao District at 0.76. Meanwhile, the lowest Gender Inequality Index in Nusa Tenggara is in Kupang City at 0.37. Additionally, the average is 0.516 and the variance is 0.009.

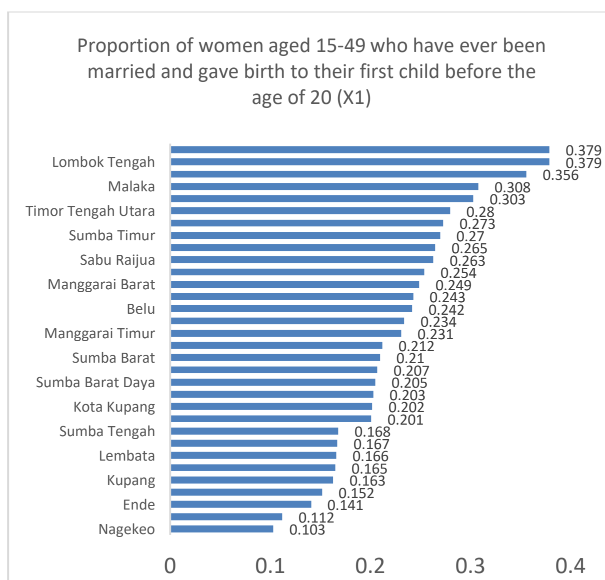

Figure 2. *Bar Chart of Variable X_1*

International Journal of Academic Management Science Research (IJAMSR)
ISSN: 2643-900X
Vol. 8 Issue 5 May - 2024, Pages: 25-33

Based on **Figure 2.**, it can be seen that the highest proportion of women aged 15-49 who were ever married and gave birth to their first child before the age of 20 in Nusa Tenggara is in West Lombok and Central Lombok Districts, namely 0.379. This is followed by Central Lombok District at 0.356. Meanwhile, the lowest proportion in Nusa Tenggara is in Nagekeo District, at 0.103.

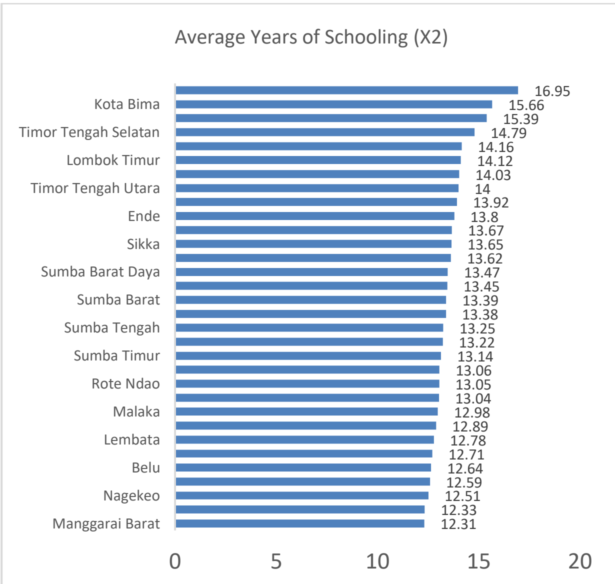

Figure 3. *Bar Chart of Variable* X_2

Based on **Figure 3.**. it can be seen that the highest Average Length of Schooling in Nusa Tenggara is in Kupang City, at 16.95. Then, there is Bima City at 15.66 and Mataram City at 15.39. Meanwhile, the lowest Average Length of Schooling is in Manggarai District, at 12.31.

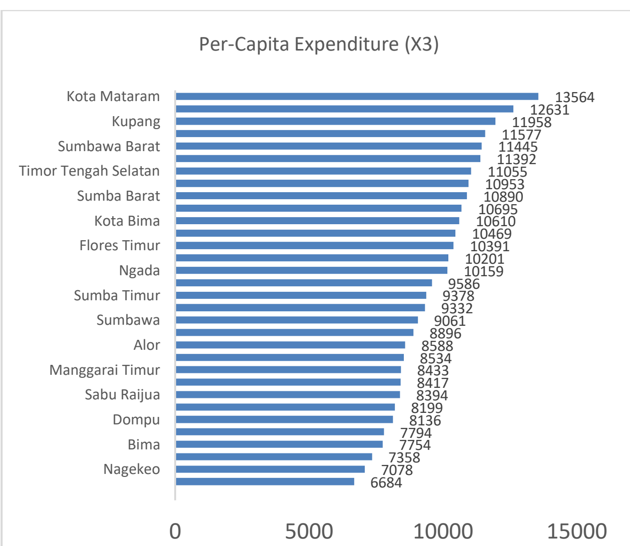

Figure 4. *Bar Chart of Variable* X_3

Based on **Figure 4.**, it can be seen that the highest Per Capita Expenditure in Nusa Tenggara is in Mataram City, at 13,564. Then, there is Ende District at 12,631 and Kupang District at 11,958. Meanwhile, the lowest Per Capita Expenditure in Nusa Tenggara is in North Lombok District, at 6,684.

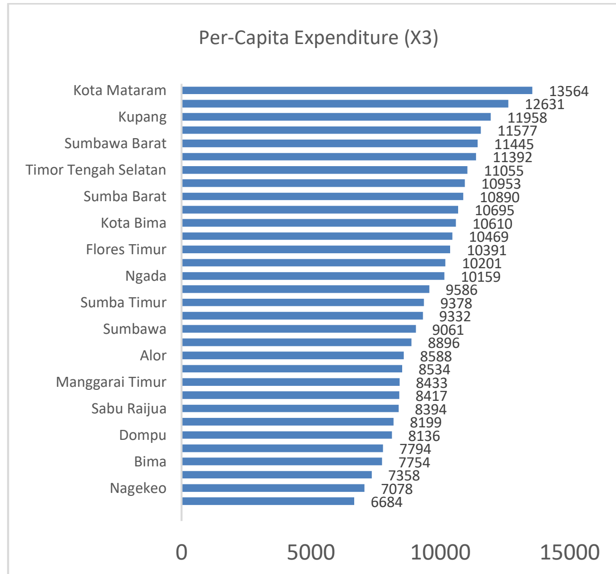

Figure 5. *Bar Chart of Variable* X_4

Based on **Figure 5.**, it can be observed that the percentage of Women's Participation in the Parliament of Nusa Tenggara Province is highest in Mataram City at 25%. Then, there is Belu District at 23.33% and Kupang City at 20%. Meanwhile, the lowest percentage of Women's Participation in Parliament in Nusa Tenggara is in Nagekeo and Lembata Districts, at 1%.

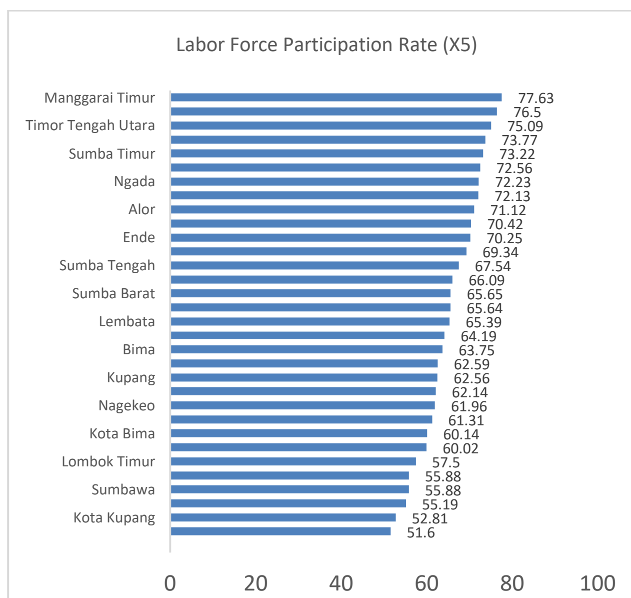

Figure 6. *Bar Chart of Variable* X_5

Based on **Figure 6.**, it can be observed that the percentage of Labor Force Participation Rate in Nusa Tenggara is highest in East Manggarai District at 77.63%. Then, there is South Central Timor District at 76.5% and North Central Timor District at 75.09%. Meanwhile, the lowest percentage of Labor Force Participation Rate in Nusa Tenggara is in Mataram City, at 51.6%.

3.2 Best Model Selection Analysis

In this study, the modeling of the Gender Inequality Index (GII) in 32 Districts/Cities in West Nusa Tenggara and East Nusa Tenggara Provinces using the MARS method based on 5 predictor variables is as follows.

3.2.1 Combination of Basis Functions

The calculations were performed using MARS software by combining Basis Function (BF), Maximum Interaction (MI), and Minimum Observations between knots (MO). The Basis Function (BF) values ranged from 10 to 20, Maximum Interaction (MI) was set at 1 and 2, and Minimum Observations between knots (MO) were 0, 1, 2, and 3. The criteria for the best model were the lowest or minimum Generalized Cross-Validation (GCV) value. If there were two or more models with the same minimum GCV value, the decision could be based on the maximum R^2 value and minimum Mean Square Error (MSE) value. The best model produced was at the fourteen basis function value. The results of the combinations at the fourteen basis functions are as follows:

Table 2. 14 Base Function Combinations

BF	MI	MO	GCV	R^2(Naive)	MSE
14	1	0	0,003	0,832	0,002
		1	0,003	0,945	0,000613
		2	0,003	0,830	0,002
		3	0,003	0,855	0,002
	2	0	0,003	0,841	0,002
		1	**0,002**	**0,945**	**0,000613**
		2	0,002	0,864	0,001
		3	0,003	0,857	0,001

Based on **Table 2**, It can be seen that the best model is obtained from the combination of fourteen basis functions, a maximum interaction of two, and a minimum observation between knots of one. The Generalized Cross-Validation (GCV) value is 0.002; R^2 is 0.945; and the Mean Square Error (MSE) is 0.000613 from this combination. An R^2 value of 0.945 means that 94.5% of the variation in the response variable, Gender Inequality Index (Y), can be explained by the predictor variables (X).

3.2.2 Basis Function Estimation

The estimation of the best model for modeling the Gender Inequality Index (IKG) in Nusa Tenggara is as follows.

Table 3. Best Significant Model Basis Function Estimation

Basis Function (BF)	Parameter Estimation
Intersept/Constant	0,583
$BF_1 = \max(0, X_4 - 3{,}330)$	-0,011
$BF_2 = \max(0, 3{,}330 - X_4)$	0,068
$BF_3 = \max(0, X_1 - 0{,}103)$	0,357
$BF_4 = max(0, X_5 - 51{,}600)$	-0,003
$BF_6 = max(0, 0{,}212 - X_1) \times BF_4$	-0,046

Based on **Table 3**, The following is the MARS model to estimate the Gender Inequality Index (GII) in Nusa Tenggara.

$$\hat{Y} = 0{,}583 - 0{,}011BF_1 + 0{,}068BF_2 + 0{,}357BF_3 - 0{,}003BF_4 - 0{,}046BF_6 \quad \textbf{(1)}$$

BF_1, BF_2, BF_3, BF_4 and BF_6 will be explained in point **3.4.1**

3.2.3 Residual Assumption Test

Testing the residual assumption is very necessary for further inference. The residuals are assumed to be normally distributed and the residuals from each observation are assumed to have a constant variance and have a value of $\sigma 2$. The hypothesis used in testing the residual assumption is as follows.

H_0: Residuals are normally distributed

H_1:The residuals are not normally distributed

Testing is carried out using software assistance. The results of the test statistical calculations using the Kolmogorov-Smirnov normality test are as follows.

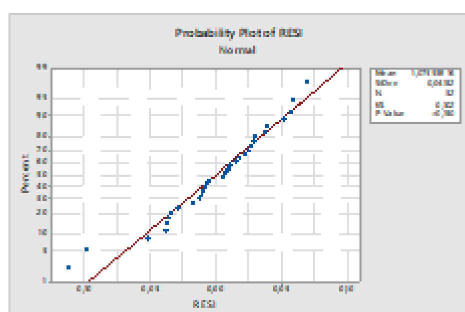

Figure 7. Residual Normality Assumption Test Plot

The critical area in this test is repulsion H_0 if the $p-value$ is less than the significance level (α). Based on **Figure 7.**, it can be seen that the $p - value$ is $0{,}150$, which is more than the significance level $(\alpha = 0{,}05)$. Therefore, a decision of failure to reject was obtained H_0, so the conclusion is that the residuals are normally distributed. This test is carried out to detect symptoms of heteroscedasticity in the residuals. Testing was carried out using the Glejser test. The hypothesis used in testing heteroscedasticity using the Glejser test is as follows.

H_0: $\sigma_1^2 = \sigma_2^2 = \cdots = \sigma_{32}^2 = \sigma^2$

H_1: There is at least one $\sigma_i^2 \neq \sigma^2; i = 1, 2, \ldots, 32$

The results of the heteroscedasticity detection test using the Glejser test with the help of Minitab software are as follows.

Table 4. Glejser Test

Glejser Test	*P-Value*	Decision
	0,056	Failed to Reject H_0

The critical area in this test is repulsion H_0 if the $p - value$ is less than the significance level (α). Based on **Table 4**, it can be seen that the $p - value$ is 0,56, which is more than the significance level (α = 0.05). Therefore, a decision of failure to reject was obtained H_0, so the conclusion is that there was no case of heteroscedasticity.

3.3 MARS Model Significance Test

3.3.1 Simultaneous Tet of MARS Model Base Function Coefficients

The hypothesis used in simultaneous testing of basis function coefficients is as follows.

H_0: $a_1 = a_2 = a_3 = a_4 = a_6 = 0$
H_1: There is at least one $a_m \neq 0; m = 1,2,3,4,6$

Table 5. Simultaneous Test of MARS Model Function Coefficients

Statistics Test	Value
F Test Statistics	90,041
P-value	$0{,}144329 \times 10^{-14}$

The critical area of simultaneous testing is rejecting H_0 if the F_{count} is more than (0,05,4,22) or the $p - value$ is less than the significance level (α = 0.05). Based on **Table 5**, the calculated $p - value$ is $60{,}352$, which is more than $F(0{,}05{,}4{,}22) = 2{,}82$. Apart from that, the resulting $p - value$ is $0{,}152881 \times 10^{-10}$, which is less than the significance level $(\alpha = 0.05)$. Therefore, the decision obtained is to reject H_0, so the conclusion is that there is at least one am that is not equal to zero, with $m = 1,2,3,4,6$. This can be interpreted that the model obtaine

3.3.2 Partial Test of MARS Model Base Function Coefficients

The hypothesis used in partial testing of the basis function coefficients is as follows.

H_0: $a_m = 0, m = 1,2,3,4,6$
H_1: $a_m \neq 0, m = 1,2,3,4,6$

Table 6. Partial Test of MARS Model Basis Founction Coefficients

Parameter	*Estimate*	S.E.	*T-Ratio*	*P-value*
Constant	0,583	0,028	20,769	0,000
BF_1	-0,011	0,001	−9,627	0,000
BF_2	0,068	0,008	9,086	0,000
BF_3	0,357	0,101	3,537	0,002
BF_4	-0,003	0,819	−4,095	0,000
BF_6	-0,046	0,014	-3,200	0,004

The critical area of the partial test is to reject H_0 if $|t_{count}|$ > (0,0025,23) or *p-value* less than the significance level (α = 0.05). Based on **Table 6**, the $|t_{count}|$ value is obtained of each basis function in the model is more than $t(0{,}025{,}23) = 2{,}069$. In addition, the $p - value$ of each basis function in the model is less than the significance level $(\alpha = 0.05)$. Therefore, the decision obtained is to reject H_0, so the conclusion is that there is at least one a_m that is not equal to zero, with $m = 1,2,3,4,6$. This can be interpreted that the model obtained shows a relationship between the basis function coefficients and the response variable.

3.3.3 Variable Importance Level

The level of importance of variables is one of the outputs generated by the MARS software. The level of importance of variables is used to rank predictor variables that influence the response variable. The level of importance of variables in modeling Gender Inequality Index (GII) data in Nusa Tenggara is as follows.

Table 7. Importance Level Variable

Variable	Variable Name	Level of Importance	GCV Reduction
X_4	Women's Participation in Parliament	100%	0,017

X_1	Proportion of women aged 15-49 who have ever been married and gave birth to their first child before the age of 20	40,441%	0,004
X_5	Labor Force Participation Rate	10,902%	0,002
X_2	Average Years of Schooling	0%	0,002
X_3	Per-Capita Expenditure	0%	0,002

Based on **Table 7**, it can be observed that the predictor variable most influencing the response variable is the Women's Participation in Parliament variable (X_4) with a level of importance of 100%. This is consistent with the actual conditions as stated in the research by Yolanda (2022), which indicates that the higher the proportion of women involved in parliament or, in other words, equal to the number of men, the Gender Inequality Index can be reduced. Additionally, the Women's Participation in Parliament variable can decrease the Generalized Cross Validation (GCV) value by 0,017 when included in the model.

Furthermore, the predictor variables affecting the response variable based on the order of importance are the Proportion of Women Ever Married Aged 15-49 Years Giving Birth to Their First Child Under 20 Years Old (X_1) with an importance level of 4,441% and the Labor Force Participation Rate (X_5) with an importance level of 10,902%. Meanwhile, the predictor variables with a 0% importance level are the Years of Schooling Expectation (X_2) and Per Capita Expenditure (X_3).

Considering the contribution of predictor variables in reducing the Generalized Cross Validation (GCV) value, the Women's Participation in Parliament variable (X_4) can reduce the GCV value by 0,017; the Proportion of Women Ever Married Aged 15-49 Years Giving Birth to Their First Child Under 20 Years Old (X_1) can reduce the GCV value by 0.004; the Labor Force Participation Rate (X_5); Years of Schooling Expectation (X_2); and Per-Capita Expenditure (X_3) can reduce the GCV value by 0.002 when included in the model. The decrease in GCV value applies when these variables are included in the model.

3.3.4 Mean Absolute Percentage Error (MAPE) Value

The Mean Absolute Percentage Error (MAPE) value is used to measure the accuracy of the best model estimation. Calculating the MAPE value can use the formula :

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\frac{|Y_t - \hat{Y}_t|}{Y_t} \times 100\% \quad \textbf{(2)}$$

The MAPE value obtained is 9,462%. This shows that the MAPE value is in the range of less than 10%, meaning that the estimation results show very accurate capabilities.

3.4 Best MARS Model Interpretation

The best MARS model has been obtained with a combination of fourteen basis functions, maximum interactions of two, and minimum observations between knots of one. After getting the best model and testing significant variables, as well as assumptions on the residuals, the response variable and the estimated results can be plotted to compare the two values. The plot is as follows.

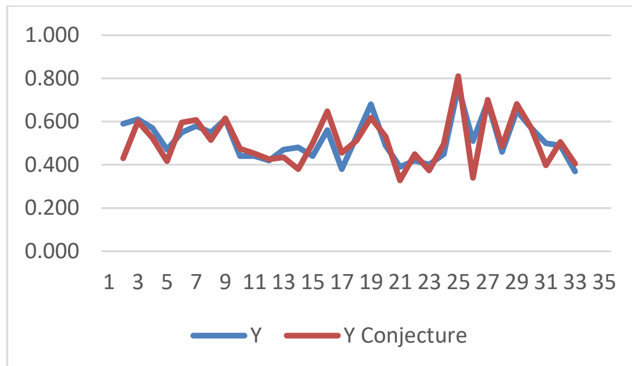

Figure 8. Plot Y with $\hat{Y}$

Based on **Figure 8.**, it can be seen that the estimated value ($\hat{Y}$) is close to the value from factual data, namely Gender Inequality Index (Y). Apart from that, based on the best model that has been obtained, the interpretation results are as follows.

3.4.1 Base One Functions (BF_1)

$$BF_1 = \begin{cases} (X_4 - 3{,}330) ; & \text{for } X_4 > 3{,}330 \\ 0 ; & \text{for } X_4 \text{ the other} \end{cases}$$

The interpretation of the value of basis function one (BF_1) with a coefficient of -0.011 in the best model signifies that each increase of one unit in BF_1 will decrease the Gender Inequality Index by 0.011, with basis functions (BF_2), (BF_3), (BF_4), and (BF_6) considered constant. Additionally, another implication is that the Percentage of Women's Participation in Parliament will provide a significant contribution if it exceeds 3.330, resulting in a decrease of 0.011 in the Gender Inequality Index in the Nusa Tenggara Province.

3.4.2 Base Two Functions (BF_2)

$$BF_2 = \begin{cases} (3{,}30 - X_4) ; & \text{for } X_4 < 3{,}30 \\ 0 ; & \text{for } X_4 \text{ the other} \end{cases}$$

The interpretation of the value of basis function two (BF_2) with a coefficient of 0.068 in the best model, implies that each increase of one unit in (BF_2) will increase the Gender Inequality Index by 0.068, with basis functions (BF_1), (BF_3), (BF_4), and (BF_6) considered constant. Additionally, another implication is that the Percentage of Women's Participation in Parliament will provide a significant contribution if it is less than 3.30, resulting in an increase of 0.068 in the Gender Inequality Index in the Nusa Tenggara Province.

3.4.3 Base Three Functions (BF_3)

$$BF_3 = \begin{cases} (X_1 - 0{,}103) \; ; & \text{for } X_1 > 0{,}103 \\ 0 \; ; & \text{for } X_1 \text{ the other} \end{cases}$$

The interpretation of the value of basis function three (BF_3) with a coefficient of 0.357 in the best model, implies that each increase of one unit in (BF_3) will increase the Gender Inequality Index by 0.357, with basis functions (BF_1), (BF_2), (BF_4), and (BF_6) considered constant. Additionally, another implication is that the Proportion of Women Ever Married aged 15-49 who gave Birth to their First Child Before the Age of 20 will provide a significant contribution if it is greater than 0.103, resulting in an increase in the Gender Inequality Index in the Nusa Tenggara Province.

3.4.4 Base Four Functions (BF_4)

$$BF_4 = \begin{cases} (X_5 - 51{,}600) \; ; & \text{for } X_5 > 51{,}600 \\ 0 \; ; & \text{for } X_5 \text{ the other} \end{cases}$$

The interpretation of the value of basis function four (BF_4) with a coefficient of -0.003 in the best model implies that each increase of one unit in (BF_4) will decrease the Gender Inequality Index by 0.003, with basis functions (BF_1), (BF_2), (BF_3), and (BF_6) considered constant. Additionally, another implication is that the Labor Force Participation Rate (Women) will provide a significant contribution if it is greater than 51.600, resulting in a decrease in the Gender Inequality Index in the Nusa Tenggara Province.

3.4.5 Base Five Functions (BF_5)

$$BF_6 = \begin{cases} (0{,}212 - X_1)((X_5 - 51{,}600) \; ; & \text{for } X_1 < 0{,}212 \; and \; X_5 > 51{,}600 \\ 0 \; ; & \text{for } X_1 \text{dan } X_5 \text{ the other} \end{cases}$$

The interpretation of the value of basis function six (BF_6) with a coefficient of -0.046 in the best model, implies that each increase of one unit in (BF_6) will decrease the Gender Inequality Index by 0.046, with basis functions (BF_1), (BF_2), (BF_3), and (BF_4) considered constant. Additionally, another implication is that if the variable Proportion of Women Ever Married aged 15-49 who Gave Birth to the First Child Under the Age of 20 is less than 0.212 and the variable Labor Force Participation Rate is greater than 51.600, it will provide a significant contribution in the form of a decrease in the Gender Inequality Index in the Nusa Tenggara Province.

4. CONCLUTIONS

Based on the results of the analysis and discussion that has been carried out, the conclusions obtained from the research as follows:

1. Descriptive statistics for each variable in this study indicate that the highest Gender Inequality Index (GII) in the Nusa Tenggara Province is recorded in Rote Ndao Regency at 0.76, while the lowest GII is observed in Kupang City at 0.37. Additionally, the mean GII is 0.516 with a variance of 0.009. On the other hand, scatterplots of each predictor variable against the response variable show no specific data distribution pattern (trend), indicating suitability for non-parametric regression approach.
2. The best model obtained using the Multivariate Adaptive Regression Spline (MARS) method is a combination of fourteen basis functions, maximum interaction of two, and minimum observation between knots of one. The Generalized Cross Validation (GCV) value is 0.002; R^2 is 0.945; and Mean Square Error (MSE) is 0.000613 from this combination. An R^2 value of 0.945 indicates that 94.5% of the variation in the Gender Inequality Index (Y) can be explained by the predictor variables (X). Below is the best model obtained using the MARS method.

$$\hat{Y} = 0{,}583 - 0{,}011BF_1 + 0{,}068BF_2 + 0{,}357BF_3 - 0{,}003BF_4 - 0{,}046BF_6$$

3. The interpretation of the best model in this study is as follows:
   a) Based on the obtained best model, the interpretation of the value of basis function one (BF_1) indicates that the Percentage of Women's Participation in Parliament will significantly contribute to a decrease in GII in the districts/cities with values above 3.330 in the Nusa Tenggara Province by 0.011.
   b) Based on the obtained best model, the interpretation of the value of basis function two (BF_2) suggests that the Percentage of Women's Participation in Parliament will significantly contribute to an increase in GII in the districts/cities with values less than 3.30 in the Nusa Tenggara Province by 0.068.
   c) Based on the obtained best model, the interpretation of the value of basis function three (BF_3) implies that the Proportion of Women Ever Married Aged 15-49 Who Gave Birth to Their First Child Before Age 20 will significantly contribute to an increase in GII in

the districts/cities with values above 0.103 in the Nusa Tenggara Province by 0.357.

d) Based on the obtained best model, the interpretation of the value of basis function four (BF_4) indicates that the Labor Force Participation Rate (Female) will significantly contribute to a decrease in GII in the districts/cities with values above 51.600 in the Nusa Tenggara Province by 0.003.

e) Based on the obtained best model, the interpretation of the value of basis function six (BF_6) suggests that the variable Proportion of Women Ever Married Aged 15-49 Who Gave Birth to Their First Child Before Age 20, with values less than 0.212, and the variable Labor Force Participation Rate with values above 51.600 will significantly contribute to a decrease in GII in the districts/cities of the Nusa Tenggara Province by 0.046.

5. Acknowledgment

This research could not have been carried out without help and support from various parties. For this reason, the author would like to thank the Statistics Study Program, Faculty of Science and Technology, Airlangga University. Thanks to everyone involved in this endeavor. Your support and encouragement have been invaluable, and the author were truly grateful for the opportunity to undertake this research..

6. References

[1] BPS. (2021). Kajian Penghtungan Indeks Ketimpangan Gender. Jakarta: Badan Pusat Statstik Republik Indonesia.

[2] Carlsen, L. (2020). Gender inequality and development. Sustainability Science,, 15(3) 759-780..

[3] Hubeis, A. V. (2010). Pendekatan Gender dan Pembangunan dalam Pemberdayaan Perempuan dari Masa ke Masa. Bogor: IPB Press.

[4] Maemunah., G. U. (2021). Perempuan Dalam Indeks Pembangunan Manusia (Kajian IPM-IPG dan IDG Sumber Kemenpp Tahun 2019). Jurnal Jendela Hukum, 8(2), 75-86.

[5] Massaid, A. H. (2019). Modelling of Poverty Percentage of Non-Food Per Capita Expenditures in Indonesia Using Least Square Spline Estimator. IOP Conference Series: Materials Science and , 564(5). 050244.

[6] Otok, B. (2010). Pendekatan Multivariate Adaptive Regression Spline (MARS) pada Pengelompokkan Zona Musim Suatu Wilayah. Jurnal Statistika , Vol.10.

[7] Resroeningroem, R. (2018). Dinamika Kedepan Pengarusutamaan Gender.Seminar Nasional dan Diskusi Panel Multidisiplin Hasil Penelitian dan Pengabdian Kepada Masyarakat 2018.

[8] UNDP. (2015). Overview Human Development Report 2015. New York: Work for Human Development.

[9] Utaimingsih, A. (2017). Gender dan wanita karir. Unversitas Brawijaya Press.

[10] Yolanda, A. M. (2022). Pemodelan Klasifikasi pada Indeks ketimpangan Gender (IKG) Tahun 2020 dengan Metode Naive Bayes. Jurnal Keluarga Berencana, 7(1), 231-244.