**JTAM (Jurnal Teori dan Aplikasi Matematika)**
http://journal.ummat.ac.id/index.php/jtam
p-ISSN 2597-7512 | e-ISSN 2614-1175
Vol. 9, No. 2, April 2025, pp. 420-430

# Modeling Prevalence of Hypertension in Indonesia with Multivariate Adaptive Regression Splines Method

**Suliyanto[1], Toha Saifudin[1], Sheila Sevira Asteriska Naura[1], Sanda Insania Dewanty[1], Indana Zulfa Wulandari[1], Nabila Shafa Aflaha[1], Niswa Faizah Aulia[1]**
[1]Department of Mathematics, Airlangga University, Indonesia
suliyanto@fst.unair.ac.id

**Article History:**
Received : 08-12-2024
Revised : 10-03-2025
Accepted : 11-03-2025
Online : 24-04-2025

**Keywords:**
Hypertension;
MARS;
Risk Factors;
Public Health;
SDGs.

**ABSTRACT**

Hypertension is one of the important public health problems in Indonesia, which contributes to the high prevalence of non-communicable diseases. This study aims to model the prevalence of hypertension in Indonesia using the Multivariate Adaptive Regression Splines (MARS) method to identify significant predictors and their interactions. The data used was secondary data from the 2023 Indonesian Health Survey, including variables such as smoking prevalence, physical inactivity, dietary habits (consumption of fatty and sweet foods), lack of fruit and vegetable consumption, and obesity prevalence. The MARS method was used to analyse the nonlinear relationships and interactions between these predictors. After a trial-and-error process to determine the optimal number of basis functions (BF), maximum interactions (MI), and minimum observations (MO), the best model was achieved with BF = 18, MI = 3, and MO = 1. This model produced a Generalised Cross Validation (GCV) value of 13.428 and R-Square of 0.278. This fairly low R-Square value indicates that the factors analysed have contributed to the variation in hypertension prevalence, but there are still other aspects that can be taken into account to improve the predictive power of the model. The significant predictor variables were consumption of fatty foods (X3), lack of physical activity (X2), and consumption of sweets (X4), with the highest importance on X3 (100%). The findings reveal that interactions between variables, such as dietary habits and physical inactivity, significantly influence the prevalence of hypertension. For example, higher consumption of fatty and sweet foods combined with low physical activity increases the risk of hypertension. These results demonstrate the effectiveness of the MARS method in capturing complex and nonlinear relationships and serve as findings that highlight the need for health policies that focus on healthy diets and increased physical activity, in line with Goal 3 of the SDGs, "Good Health and Well-Being," which aims to reduce premature mortality from noncommunicable diseases. Recommended interventions include nutrition education campaigns and community-based exercise programs to reduce the prevalence of hypertension in Indonesia.

https://doi.org/10.31764/jtam.v9i2.28392


——————————— ◆ ————————————

## A. INTRODUCTION

Hypertension, also known as high or raised blood pressure, is a condition in which the blood vessels have persistently raised pressure. According to the World Health Organization (WHO), hypertension can occur if the systolic pressure, which is the blood pressure when the heart contracts, exceeds 140 mmHg, or the diastolic pressure, which is the blood pressure when the heart rests between heartbeats, exceeds 90 mmHg. Hypertension is one of the factors contributing to the highest mortality rate by non-communicable diseases whose number of sufferers continues to increase every year (Zhou et al., 2021). Hypertension can attack various ages ranging from children, adolescents, adults, to the elderly and even pregnant women (Vintilă & Dorobanțu, 2023). Several studies explain that there are several factors that are at risk for hypertension in Indonesia, such as age, gender, family history, obesity, smoking habits, lack of physical activity, excessive salt consumption, dyslipidaemia, alcohol consumption, and psychosocial and stress (Mutebi et al., 2023).

According to data from the World Health Organization (WHO) in 2019, 33% of the world's population or around 1.3 billion people suffer from hypertension. The greatest number of people with hypertension live in the most populous WHO regions, Western Pacific Region and South-East Asia Region (WHO, 2023). Based on the results of the Indonesian Health Survey (HIS) in 2023, according to blood pressure measurement data, hypertension is one of the non-communicable diseases suffered by many Indonesians aged 18 years and over, which is 30.8%. This makes Indonesia the 5th country with the most cases of hypertension in the world. This condition makes hypertension one of the major challenges in achieving the Sustainable Development Goals (SDGs), especially in goal 3, namely "Healthy and Prosperous Life" which targets a one-third reduction in premature deaths from non-communicable diseases, including hypertension by 25%.

The method that can be done in modelling the risk factors for hypertension without involving spatial effects is Multivariate Adaptive Regression Splines (MARS). MARS is a nonparametric regression method that does not require certain assumptions about the relationship between response variables and predictor variables (Bekar Adiguzel & Cengiz, 2023); (López & Kholodilin, 2023). MARS is more flexible in handling the complexity of health data that is often non-linear (Chang et al., 2023). MARS has the advantage of handling interactions between variables and simplifying the model without reducing the level of accuracy (Özmen et al., 2022). MARS is also effective in overcoming multicollinearity problems by selecting the most relevant basis functions, thereby improving its predictive performance (Sahraei et al., 2021).

Research conducted by (Van Oort et al., 2020) revealed factors that have a significant effect on the incidence of hypertension, three of which are smoking habits, physical activity, and obesity. Other studies were also conducted by (Maniero et al., 2023), (Yuan et al., 2020), and (Teo & Rafiq, 2021) which revealed that several factors can contribute to hypertension, especially related to lifestyle and non-pharmacological interventions, including smoking, excessive salt intake, low potassium intake, fat intake, unhealthy diet, overweight or obesity, lack of physical activity, excessive alcohol consumption, and psychological stress. In addition, (Ningsih, 2024) also conducted research using the Multivariate Adaptive Regression Splines (MARS) method to model the factors that influence the prevalence of hypertension. The results concluded that the proportion of the population with the habit of not consuming fruits and vegetables per day in a week, the proportion of the population with the habit of consuming fatty foods more than once a day, the prevalence of obesity nutritional status, and the proportion of the population with the habit of consuming salty foods more than once a day had a significant effect on the prevalence of hypertension. However, the data used in the study was limited to East Java Province and used data from the East Java Riskesdas Report 2018, so it can be considered less relevant as a reference for decision making at this time.

Recent studies have identified additional risk factors for hypertension in Indonesia. A study by Andini & Siregar (2024) found that individuals working more than 40 hours per week had a higher prevalence of hypertension, suggesting that long working hours may contribute to increased blood pressure levels. This finding aligns with global research indicating that chronic work stress and extended working hours can lead to lifestyle changes, such as reduced physical activity and poor dietary habits, which increase hypertension risk (Virtanen & Kivimaki, 2018). Furthermore, environmental factors, including air pollution and noise exposure, have also been linked to elevated blood pressure (Yang et al., 2020). These findings suggest that beyond individual lifestyle choices, broader social and environmental determinants play a crucial role in hypertension incidence.

Although previous studies have provided valuable insights into the factors influencing hypertension, there is still a lack of comprehensive analysis using more recent and broader data across Indonesia. Furthermore, while Ningsih (2024) utilized the MARS method, the geographical scope and data timeliness were limited to 2018, reducing the generalizability of the findings. This study aims to bridge this gap by applying the MARS method to model hypertension risk factors using more current and nationally representative data. By doing so, it is expected to provide a more accurate understanding of the factors influencing hypertension in Indonesia and highlight the advantages of MARS in handling complex, nonlinear relationships and variable interactions.

Based on the above description, the author is interested in analysing the coding of hypertension in Indonesia using statistical methods, namely and Multivariate Adaptive Regression Splines (MARS) because it is superior in handling nonlinear relationships and interactions between complicated predictor variables. The results of this study are expected to provide an overview of the factors that influence the prevalence of hypertension, so that it can make an important contribution to efforts to prevent and control hypertension in Indonesia.

## B. METHODS

The type of research used in this study is quantitative research with the data used were data on the prevalence of hypertension in the population aged $\geq 18$ years in Indonesia, the prevalence of smoking in the population aged $\geq 10$ years, the proportion of physical inactivity, the proportion of fatty food consumption habits, the proportion of sweet food consumption habits, the proportion of not consuming fresh fruits and vegetables, and the prevalence of nutritional status of the population aged $\geq 18$ years who were obese obtained from the publication of Health Survey Report Indonesia 2023 (Badan Kebijakan Pembangunan Kesehatan, 2023). The selection of predictor variables in this research was based on previous research which revealed that these six predictor variables have a significant influence on the prevalence of hypertension (Zhao et al., 2024), (Van Oort et al., 2020), (Maniero et al., 2023), (Yuan et al., 2020), (Teo & Rafiq, 2021), and (Ningsih, 2024). The variables used in this study are presented in Table 1 below.

**Table 1**. Research Variables

| Variables | Description | Scale | Data Source |
|---|---|---|---|
| $Y$ | Hypertension prevalence in Indonesia | Ratio | Laporan Survei Kesehatan Indonesia 2023 |
| $X_1$ | Prevalence of population smoking | Ratio | Laporan Survei Kesehatan Indonesia 2023 |
| $X_2$ | Proportion of physical inactivity | Ratio | Laporan Survei Kesehatan Indonesia 2023 |
| $X_3$ | Proportion of fatty food consumption habits | Ratio | Laporan Survei Kesehatan Indonesia 2023 |
| $X_4$ | Proportion of sweet food consumption habits | Ratio | Laporan Survei Kesehatan Indonesia 2023 |
| $X_5$ | Proportion not consuming fresh fruits and vegetables | Ratio | Laporan Survei Kesehatan Indonesia 2023 |
| $X_6$ | Prevalence of nutritional status of obese population | Ratio | Laporan Survei Kesehatan Indonesia 2023 |

The data obtained were processed using MARS software with the stages of data analysis shown in Figure 1.

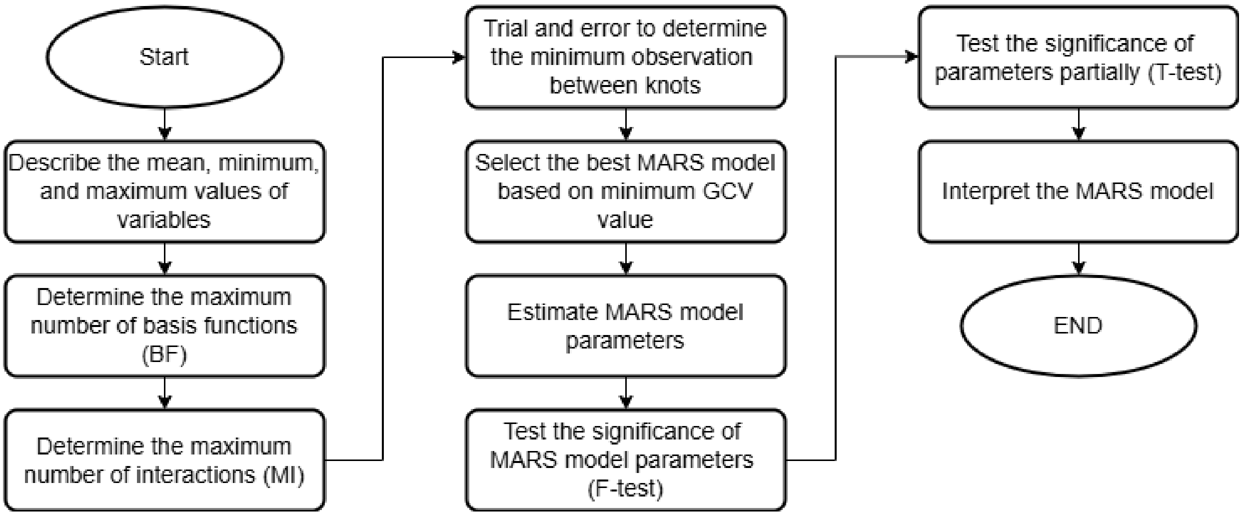


**Figure 1.** MARS Analysis Flowchart

The following is an explanation of the analysis stages of the flowchart in Figure 1.

a. Describe the mean, minimum, and maximum values of hypertension incidence in Indonesia and the factors that are thought to influence it.
b. Modelling the incidence of hypertension in Indonesia using the Multivariate Adaptive Regression Splines (MARS) method with the following steps.
   1) Determine the maximum number of basis functions (BF), with the recommended basis functions to be used being between two and four times the number of predictor variables (Friedman, 1991), so the number of BFs used are 12, 18, and 24.
   2) Determine the maximum number of interactions (MI), with the recommended number of MI being 1, 2, and 3. If the number of MI is more than 3, the resulting model will be very complex (Friedman & Roosen, 1995).
   3) Perform trial and error to determine the minimum observation between knots. Knot is the end of one segment and the beginning of another segment (Maleki & Pak, 2022).

4) The selection of the best MARS model is done by determining the optimal number of knots, because the optimal number of knots will result in a good MARS model. The process of selecting a model is using the stepwise method. Forward stepwise is used to find the maximum number of base functions by minimizing the Average Sum of Square Residual (ASR). After that, backward stepwise is done to simplify the model by removing the function base that contributes little to the response by reviewing the Generalized Cross Validation value. The best model is determined by the smallest GCV value. The Generalized Cross validation or GCV method is defined as follows (Putra et al., 2023).

$$GCV(\lambda) = \frac{n^{-1}\sum_{i=1}^{n}[y_i-\hat{y}_i]^2}{\left(1-\frac{trace[\boldsymbol{H}(\lambda)]}{n}\right)^2} \tag{1}$$

with: $y_i$ = response variable; $\hat{y}_i$ = prediction of $y_i$ ; $n$ = data count; $\boldsymbol{H}(\lambda)$ = $\mathbf{B}(\mathbf{B}^{\mathbf{T}}\mathbf{B})^{-1}\mathbf{B}^{\mathbf{T}}$where $\mathbf{B}$ is the function-base matrix
5) Estimating MARS model parameters
6) Testing the significance of MARS model parameters simultaneously (F test)
7) Testing the significance of parameters partially (T test)
8) Interpret the MARS model obtained.

## C. RESULT AND DISCUSSION

### 1. Descriptive Statistic

The characteristics of the incidence of hypertension in Indonesia and the factors that are thought to influence it can be presented in the form of descriptive statistics as presented in Table 2 below.

**Table 2.** Characteristics of Variables

| Variables | N | Mean | Minimum | Maximum |
|---|---|---|---|---|
| Y | 38 | 28,134 | 19,9 | 40,7 |
| X1 | 38 | 19,918 | 10,0 | 27,7 |
| X2 | 38 | 43,503 | 27,8 | 62,0 |
| X3 | 38 | 27,553 | 12,2 | 54,2 |
| X4 | 38 | 23,592 | 11,0 | 46,5 |
| X5 | 38 | 14,495 | 5,0 | 30,7 |
| X6 | 38 | 22,968 | 13,3 | 31,8 |

Based on the data presented in Table 2, it can be concluded that the prevalence of hypertension in Indonesia has an average of 28.134%, with the lowest value recorded at 19.9% in Papua Mountains Province and the highest value reaching 40.7% in Central Kalimantan Province. This prevalence indicates a significant variation between regions in Indonesia. In addition, the prevalence of smoking in the ≥10 years age group had an average of 19.918%, with a minimum value of 10.0% found in Papua Mountains Province, while the maximum value reached 27.7% in West Nusa Tenggara Province. This shows that smoking is still a public health challenge in various regions.

The proportion of the population that lacks physical activity has a fairly high average of 43.503%, with a minimum value of 27.8% in East Java Province and a maximum value of 62.0%

in Central Papua Province. This figure shows that sedentary lifestyle is still a health problem that needs serious attention. The habit of consuming fatty foods also shows an average of 27.553%, with the lowest proportion of 12.2% in East Nusa Tenggara Province and the highest reaching 54.2% in Central Java Province. Meanwhile, the habit of consuming sweet foods has an average of 23.592%, with a minimum value of 11.0% in Bali Province and a maximum of 46.5% in West Java Province.

The low consumption of fresh fruits and vegetables is also a concern, with an average of 14.495%. The minimum value of 5.0% was found in Lampung Province, while the maximum value reached 30.7% in Gorontalo Province. This consumption pattern is one of the important indicators in evaluating the diet of the Indonesian people. Finally, the prevalence of obese nutritional status in Indonesia has an average of 22.968%, with a minimum value of 13.3% found in East Nusa Tenggara Province and a maximum value of 31.8% in DKI Jakarta. The high prevalence of obesity, especially in urban areas, indicates changes in diet and lifestyle that need to be anticipated to prevent the risk of non-communicable diseases. Taken together, these data illustrate the various aspects of lifestyle that contribute to public health in Indonesia, such as diet, physical activity and smoking habits, all of which can influence the prevalence of hypertension and other health conditions. Targeted, data-driven interventions are needed to address these issues and improve the quality of life for people across Indonesia.

## 2. MARS Analysis

Furthermore, the analysis was carried out using the MARS method. To obtain the MARS model, first determine the maximum number of Basis Functions (BF), Maximum Interaction (MI), and Minimum Observation (MO). To get the best model, trial and error was conducted on the combination of BF, MI, and MO so as to obtain a MARS model with minimum GCV value and maximum R-Square value. Based on results of trial and error MARS model and by considering the principle of parsimony, the best MARS model is obtained with a combination of BF = 18, MI = 3, and MO = 1. The GCV value for the best MARS model is 13.428 and R-Square is 0.278 and the model form is as shown in the following equation.

$$\hat{y} = 24{,}883 - 0{,}015^{*}BF_2 + 0{,}011^{*}BF_8 + 0{,}243^{*}BF_{12} \qquad (2)$$

with:

$BF_1 = \max(0, X_3 - 12{,}20)$

$BF_2 = \max(0, X_4 - 411{,}00) * BF_1$

$BF_6 = \max(0, X_4 - 366{,}00) * BF_1$

$BF_8 = \max(0, X_2 - 52{,}00) * BF_6$

$BF_{12} = \max(0, X_3 - 14{,}90)$

The MARS model generated in this study shows that the interaction between sweet food consumption ($X_4$), fatty food consumption ($X_3$), and physical activity ($X_2$) has a significant influence on the prevalence of hypertension in Indonesia. In the model equation, the basis functions involved have their respective roles.

$BF_2$ Contains the interaction between the variables of sweet food consumption habits $(X_4)$ and fatty food consumption habits $(X_3)$ on the prevalence of hypertension. The basic function of $BF_2$ in the MARS model is defined as $BF_2 = \max(0,\ X_4 - 411.00) * BF_1$, with a coefficient $-0.015$. The function $BF_1$ which is the main component in $BF_2$ is defined as $BF_1 = \max(0, X_3 - 12.20)$ thus $BF_2$ can be represented as $BF_2 = \max(0, X_4 - 411.00) * BF_1 = \max(0, X_3 - 12.20)$. This means that for every one unit increase in $BF_2$. the prevalence of hypertension in each province in Indonesia will decrease by 0.015 percent. $BF_2$ will be meaningful or have an impact on the prevalence of hypertension in Indonesia $(Y)$ when the value of the proportion of fatty food consumption habits $(X_3)$ is more than 12.2% and the value of the proportion of sweet food consumption habits $(X_4)$ is more than 411%. The interaction between sweet food consumption habit $(X_4)$ and fatty food consumption habit $(X_3)$ has a negative influence on the prevalence of hypertension. The higher the consumption of sweets and fatty foods. the lower the probability of a person having hypertension. These results may indicate the presence of certain dietary patterns that are not fully captured in the data, or the possibility of metabolic adaptation in certain population groups. Research by (Reynolds et al., 2019), suggests that while consumption of fatty foods can be detrimental, its effects depend on a combination of the overall diet, including fibre and protein consumption that support other cardiovascular health. According to (Casas et al., 2018) also mentioned that unsaturated fats can have a protective impact on cardiovascular health if consumed in the right amount.

$BF_8$ Contains the interaction between the variables of physical inactivity $(X_2)$. sweet food consumption habits $(X_4)$. and fatty food consumption habits $(X_3)$ on the prevalence of hypertension. The basic function of $BF_8$ in the MARS model is defined $BF_8 = \max(0, X_2 - 52.00) * BF_6$, with a coefficient 0.011. The function $BF_6$ and $BF_1$ which is the main component in $BF_8$, is defined as $BF_6 = \max(0, X_4 - 366.00) * BF_1$ and $BF_1 = \max(0, X_3 - 12.20)$ thus $BF_8$ can be represented as $BF_8 = \max(0, X_2 - 52.00) * \max(0, X_4 - 366.00) * \max(0, X_3 - 12.20)$. This means that every one unit increase in $BF_8$ will increase the prevalence of hypertension by 0.011 percent in provinces in Indonesia. $BF_8$ will have a significant impact on the prevalence of hypertension in Indonesia $(Y)$ when the proportion of fatty food consumption $(X_3)$ is more than 12.2%. the proportion of sweets consumption is more than 366%. and the proportion of physical inactivity $(X_2)$ is more than 52%. The interaction between lack of physical activity $(X_2)$. proportion of sweet food consumption habits $(X_4)$. and proportion of fatty food consumption habits $(X_3)$ has a positive influence on the prevalence of hypertension. This suggests that the combination of lack of physical activity. consuming a lot of sweets and fatty foods will increase the risk of hypertension. This confirms that the combination of a sedentary lifestyle and unhealthy diet collectively increases the risk of hypertension. This supports the study by (Forouhi et al., 2018), which showed that a sedentary lifestyle combined with a poor diet is a major factor in increasing the risk of cardiovascular disease.

While $BF_{12}$ only interacts with fatty food consumption habits on the prevalence of hypertension. The basic function of $BF_8$ in the MARS model is defined $BF_{12} = \max(0, X_3 - 14.90)$. It can be interpreted that every one unit increase in $BF_{12}$ will increase the prevalence of hypertension by 0.243 percent in provinces in Indonesia. $BF_{12}$ will be meaningful or have an impact on the prevalence of hypertension in Indonesia $(Y)$ when the value of the proportion of fatty food consumption habits $(X_3)$ is more than 14.9%. This indicates that an increase in fatty food consumption alone (without considering other variables in $(BF_{12})$ will increase the risk of hypertension. Furthermore, from the MARS model in equation (1), it can be seen that there are three predictor variables that enter the model, namely variables (X2), (X3), and (X4). To see the extent to which these variables affect the formation of the MARS model, it can be seen in the following variables importance Table 3.

**Table 3.** Predictor Variables Importance Level

| Variables | Importance Level | -GCV |
|---|---|---|
| Proportion of fatty food consumption habits (X3) | 100,000 | 18,587 |
| Proportion of physical inactivity (X2) | 76,621 | 16,457 |
| Proportion of sweet food consumption habits (X4) | 76,233 | 16,426 |
| Prevalence of population smoking (X1) | 0,000 | 13,428 |
| Proportion not consuming fresh fruits and vegetables (X5) | 0,000 | 13,428 |
| Prevalence of nutritional status of obese population (X6) | 0,000 | 13,428 |

Based on Table 3. it can be seen that the variable proportion of fatty food consumption habits (X3) is the most important variable in the MARS model with an importance level of 100%, then followed by the variable proportion of physical inactivity (X2) with an importance level of 76.621%. In third place is the variable proportion of sweet food consumption habits (X4) with an importance level of 76.233%. While in the fourth, fifth, and sixth order there are variables prevalence of population smoking (X1), proportion not consuming fresh fruits and vegetables (X5), and prevalence of nutritional status of obese population (X6) which have no level of importance (0%) because they have been represented by the previous three variables. The -GCV value shows how much the GCV value is reduced if the variable is excluded from the MARS model.

From the best MARS model, parameter significance testing can be carried out. There are two stages of testing the significance of parameters in the MARS model, namely simultaneous testing and partial testing (Addini et al., 2023).

a. Simultaneous regression testing

Simultaneous test is conducted to determine the effect of independent variables simultaneously on the dependent variable based on the base function coefficient $(\alpha)$ using the F test (Rodriguez-Clare & Dingel, 2021). The hypothesis for the simultaneous significance test is as follows (Junaedi et al., 2023): $H_0$ is $\alpha_2 = \alpha_8 = \alpha_{12} = 0$ (model is not significant); $H_1$ is at least one $\alpha_m \neq 0$; $m = 2, 8, 12$ (model is significant); and Significance level is 5%. With test statistics as shown in Table 4.

**Table 4.** Simultaneous Testing Results

| F-count | P-value |
|---|---|
| 15.533 | 0.00000156 |

Based on Table 4. with the critical region, namely reject $H_0$ if the $F_{count} > F_{0.05(3,30)}$ or p-value < 0.05. So, the decision to reject $H_0$ is obtained because the value of $F_{count}(15.533) > F_{table}(2.922)$, which means that the model is significant and can be used to model hypertension cases.

b. Partial regression testing

After conducting a simultaneous test which shows that together the independent variables affect the dependent variable, the next step is to conduct a partial test. This test aims to find out what independent variables affect the dependent variable. The test statistic that will be used in this test is the t test (Rodriguez-Clare & Dingel, 2021). The hypothesis used for the partial significance test is as follows (Junaedi et al., 2023): $H_0$ is $\alpha_m = 0$ (coefficient $\alpha_m$ has no effect on the model); $H_1$is $\alpha_m \neq 0$ for each $m$, with $m = 2,3,4$; Significance level: 5%. With test statistics as shown in Table 5.

**Table 5.** Partial Testing Results

| Parameters | Estimation | S.E | $t_{count}$ | p-value |
|---|---|---|---|---|
| Constant | 24.883 | 0.774 | 32.137 | $0.999 \times 10^{-15}$ |
| BF 2 | -0.015 | 0.003 | -5.102 | $0.127 \times 10^{-4}$ |
| BF 8 | 0.011 | 0.003 | 4.085 | $0.254 \times 10^{-3}$ |
| BF 12 | 0.243 | 0.048 | 5.010 | $0.167 \times 10^{-4}$ |

The partial test results in Table 5. show that the coefficients $\alpha_2$, $\alpha_3$, and $\alpha_4$ have a significant effect on hypertension because the value of $|t_{count}| > t_{table}$ (2.355) or $p - value < \alpha(5\%)$. So, they can be used to model hypertension cases. The results of this study support the statement from the research conducted by (Van Oort et al., 2020) and (Zhao et al., 2024) which concluded that the consumption of fatty foods, the consumption of sweet foods, and physical inactivity can significantly affect the prevalence of hypertension.

## D. CONCLUSION AND SUGGESTIONS

Modelling the prevalence of hypertension in Indonesia using the Multivariate Adaptive Regression Splines (MARS) method is as follows $\hat{y} = 24{,}883 - 0{,}015^{*}BF_2 + 0{,}011^{*}BF_8 + 0{,}243^{*}BF_{12}$. The consumption of fatty foods (X3) has the greatest influence with 100% importance, followed by lack of physical activity (X2) with 76.621% importance, and consumption of sweets (X4) with 76.233% importance. BF2 represents an interaction between the consumption of fatty and sweet foods, which in some circumstances tends to reduce the incidence of hypertension. However, BF8, which includes interactions between physical inactivity, fatty food consumption, and sugary food consumption, shows a significant increase in the prevalence of hypertension, highlighting the cumulative effect of an unhealthy lifestyle. BF12 study shows that fatty food consumption is in itself a strong predictor of the onset of hypertension, reinforcing the importance of dietary habits in preventing hypertension. The analysis also shows that the model has a generalized cross-validation (GCV) value of 13.428, indicating good predictive ability. Simultaneous and partial significance tests showed that the model and parameters were significant at the 95% confidence level.

Overall, the MARS method showed a good ability to capture complex and nonlinear patterns of relationships between predictor variables. Based on these results, concrete recommendations to prevent hypertension include public health campaigns that promote a balanced diet by reducing the consumption of fatty and sugary foods, as well as initiatives that encourage regular physical activity. To reduce the risk of high blood pressure, it is important to implement community-based physical activity programs and improve access to healthier food options. Future research should consider additional variables such as sodium intake, genetic predisposition, and socioeconomic factors. In addition, incorporating spatial aspects, such as regional dietary habits and access to health care, can improve the accuracy of the model and provide a more comprehensive analysis of the prevalence of hypertension in Indonesia.

## ACKNOWLEDGEMENT

We would like to express our deepest gratitude to the Department of Mathematics, Airlangga University for providing the resources and support necessary for conducting this research. Additionally, we extend our appreciation to the authors and researchers whose works have significantly contributed to the theoretical and methodological framework of this study. Lastly, we thank all parties who provided data and other assistance to make this study possible.

## REFERENCES

Addini, P. F., Hadi, W., Muslim, P., & Harahap, R. (2023). Application Of The Multivariate Adaptive Regression Spline (Mars) Method In Analyzing Misclassification Of Elementary School Accreditation Data In The City Of Tebing Tinggi. *Jurnal Scientia*, *12*(1), 617–620. http://infor.seaninstitute.org/index.php

Andini, F. A. D., & Siregar, A. Y. M. (2024). Work hours and the risk of hypertension: the case of Indonesia. *BMC Public Health*, *24*(1), 2480. https://doi.org/10.1186/s12889-024-20003-z

Badan Kebijakan Pembangunan Kesehatan. (2023). *Survei Kesehatan Indonesia Dalam Angka*.

Bekar Adiguzel, M., & Cengiz, M. A. (2023). Model selection in multivariate adaptive regressions splines (MARS) using alternative information criteria. *Heliyon*, *9*(9), e19964. https://doi.org/10.1016/j.heliyon.2023.e19964

Casas, R., Castro-Barquero, S., Estruch, R., & Sacanella, E. (2018). Nutrition and cardiovascular health. In *International Journal of Molecular Sciences* (Vol. 19, Issue 12). MDPI AG. https://doi.org/10.3390/ijms19123988

Chang, C. C., Yeh, J. H., Chiu, H. C., Liu, T. C., Chen, Y. M., Jhou, M. J., & Lu, C. J. (2023). Assessing the length of hospital stay for patients with myasthenia gravis based on the data mining MARS approach. *Frontiers in Neurology*, *14*, 1283214. https://doi.org/https://doi.org/10.3389/fneur.2023.1283214

Forouhi, N. G., Krauss, R. M., Taubes, G., & Willett, W. (2018). Dietary fat and cardiometabolic health: Evidence, controversies, and consensus for guidance. *BMJ (Online)*, *361*, 1–8. https://doi.org/10.1136/bmj.k2139

Friedman, J. H. (1991). Multivariate Adaptive Regression Splines. *The Annals of Statistics*, *19*(1), 1 – 67. https://doi.org/10.1214/aos/1176347963

Friedman, J. H., & Roosen, C. B. (1995). An introduction to multivariate adaptive regression splines. *Statistical Methods in Medical Research*, *4*(3), 197–217. https://doi.org/10.1177/096228029500400303

Junaedi, M., Nasikhin, Hasanah, S., & Hassan, Z. (2023). Learning Patterns in Influencing Attitudes of Religious Tolerance in Indonesian Universities. *Education Sciences*, *13*(3), 285. https://doi.org/10.3390/educsci13030285

López, F., & Kholodilin, K. (2023). Putting MARS into space. Non-linearities and spatial effects in hedonic models. *Papers in Regional Science*, *102*(4), 871–896. https://doi.org/10.1111/pirs.12738

Maleki, J., & Pak, A. (2022). A Rapid Design Procedure for Tied-Back Soil Walls Using Multivariate Adaptive Regression Splines (MARS) Method. *Geotechnical and Geological Engineering*, *41*(2), 1521–1535. https://doi.org/10.1007/s10706-022-02351-y

Maniero, C., Lopuszko, A., Papalois, K. B., Gupta, A., Kapil, V., & Khanji, M. Y. (2023). Non-pharmacological factors for hypertension management: a systematic review of international guidelines. In *European Journal of Preventive Cardiology* (Vol. 30, Issue 1, pp. 17–33). Oxford University Press. https://doi.org/10.1093/eurjpc/zwac163

Mutebi, R. K., Semulimi, A. W., Mukisa, J., Namusobya, M., Namirembe, J. C., Nalugga, E. A., Batte, C., Mukunya, D., Kirenga, B., Kalyesubula, R., & Byakika-Kibwika, P. (2023). Prevalence of and Factors Associated with Hypertension Among Adults on Dolutegravir-Based Antiretroviral Therapy in Uganda: A Cross Sectional Study. *Integrated Blood Pressure Control*, *16*, 11–21. https://doi.org/10.2147/IBPC.S403023

Ningsih, P. S. (2024). *Pemodelan Multivariate Adaptive Regression Splines pada Prevalensi Hipertensi di Provinsi Jawa Timur*. Institut Teknologi Sepuluh Nopember.

Özmen, A., Zinchenko, Y., & Weber, G.-W. (2022). Robust multivariate adaptive regression splines under cross-polytope uncertainty: an application in a natural gas market. *Annals of Operations Research*, *324*, 1–31. https://doi.org/10.1007/s10479-022-04993-w

Putra, R., Fadhlurrahman, M. G., & Gunardi. (2023). Determination of the best knot and bandwidth in geographically weighted truncated spline nonparametric regression using generalized cross validation. *MethodsX*, *10*, 101994. https://doi.org/10.1016/j.mex.2022.101994

Reynolds, A., Mann, J., Cummings, J., Winter, N., Mete, E., & Te Morenga, L. (2019). Carbohydrate quality and human health: a series of systematic reviews and meta-analyses. *The Lancet*, *393*(10170), 434–445. https://doi.org/10.1016/S0140-6736(18)31809-9

Rodriguez-Clare, A., & Dingel, J. (2021). *The Effect Of Compensation, Leadership Style And Work Discipline On The Performance Of Hospital Employee In United States*. *12*(1), 33–47. https://medalionjournal.com/

Sahraei, M. A., Duman, H., Çodur, M. Y., & Eyduran, E. (2021). Prediction of transportation energy demand: Multivariate Adaptive Regression Splines. *Energy*, *224*(12), 120090. https://doi.org/https://doi.org/10.1016/j.energy.2021.120090

Teo, K. K., & Rafiq, T. (2021). Cardiovascular Risk Factors and Prevention: A Perspective From Developing Countries. In *Canadian Journal of Cardiology* (Vol. 37, Issue 5, pp. 733–743). Elsevier Inc. https://doi.org/10.1016/j.cjca.2021.02.009

Van Oort, S., Beulens, J. W. J., Van Ballegooijen, A. J., Grobbee, D. E., & Larsson, S. C. (2020). Association of Cardiovascular Risk Factors and Lifestyle Behaviors with Hypertension: A Mendelian Randomization Study. *Hypertension*, *76*(6), 1971–1979. https://doi.org/10.1161/HYPERTENSIONAHA.120.15761

Vintilă, A.-M., & Dorobanțu, M. (2023). Hypertension in Young Adults. In M. Dorobantu, V. Voicu, G. Grassi, E. Agabiti-Rosei, & G. Mancia (Eds.), *Hypertension and Heart Failure: Epidemiology, Mechanisms and Treatment* (pp. 459–476). Springer International Publishing. https://doi.org/10.1007/978-3-031-39315-0_29

Virtanen, M., & Kivimaki, M. (2018). Long Working Hours and Risk of Cardiovascular Disease. *Current Cardiology Reports*, *20*(11), 123. https://doi.org/10.1007/s11886-018-1049-9

WHO. (2023). *Global report on hypertension The race against a silent killer*. https://www.who.int/publications/i/item/9789240081062

Yang, B.-Y., Fan, S., Thiering, E., Seissler, J., Nowak, D., Dong, G.-H., & Heinrich, J. (2020). Ambient air pollution and diabetes: A systematic review and meta-analysis. *Environmental Research*, *180*, 108817. https://doi.org/10.1016/j.envres.2019.108817

Yuan, S., Yu, H. jie, Liu, M. wei, Tang, B. wen, Zhang, J., Gasevic, D., Larsson, S. C., & He, Q. qiang. (2020). Fat Intake and Hypertension Among Adults in China: The Modifying Effects of Fruit and Vegetable Intake. *American Journal of Preventive Medicine*, *58*(2), 294–301. https://doi.org/10.1016/j.amepre.2019.09.004

Zhao, Q., Wu, Q., Zhong, H., Yan, B., Wu, J., & Guo, W. (2024). Association of dietary habits with body mass index and waist circumference, and their interaction effect on hypertension. *Medicine (United States)*, *103*(20), E38178. https://doi.org/10.1097/MD.0000000000038178

Zhou, B., Perel, P., Mensah, G., & Ezzati, M. (2021). Global epidemiology, health burden and effective interventions for elevated blood pressure and hypertension. *Nature Reviews Cardiology*, *18*(11), 785-802. https://doi.org/10.1038/s41569-021-00559-8