BAREKENG: Journal of Mathematics and Its Applications
March 2026 Volume 20 Issue 1 Page 0255-0270
P-ISSN: 1978-7227 E-ISSN: 2615-3017
https://doi.org/10.30598/barekengvol20iss1pp0255-0270

# COMPARISON OF BINARY PROBIT REGRESSION AND FOURIER SERIES NONPARAMETRIC LOGISTIC REGRESSION IN MODELING DIABETES STATUS AT HAJJ GENERAL HOSPITAL SURABAYA

**Bambang Widjanarko Otok [1], Muhammad Zulfadhli [2*], Riwi Dyah Pangesti [3], Muhammad Idham Kurniawan [4], Albertus Eka Putra Haryanto [5], Darwis [6], Iwan Kurniawan [7]**

*[1,2,4] Department of Statistics, Faculty of Science and Data Analytics, Institut Teknologi Sepuluh Nopember*
*Jln. Teknik Kimia, Keputih, Sukolilo, Surabaya, Jawa Timur, 60111, Indonesia*
*[3]Department of Statistics, Faculty of Mathematics and Natural Sciences, Bengkulu University*
*Jln. WR. Supratman, Kandang Limun, Muara Bangka Hulu, Sumatera, Bengkulu 38371, Indonesia*
*[5]Regional Economic Development Institute (REDI)*
*Jln. Arief Rahman Hakim No.152, Keputih, Sukolilo, Surabaya, Jawa Timur, 60111, Indonesia*
*[6]Islamic Religious Education, Faculty of Education and Teaching, STAIN Majene*
*Jln. Balai Latihan Kerja No.7, Totoli, Banggae, Majene, Sulawesi Barat, 91415, Indonesia*
*[7]Public Sector Business Administration, Faculty of Education, Politeknik STIA LAN Bandung*
*Jln. Hayam Wuruk No.34-38, Citarum, Bandung Wetan, Bandung, Jawa Barat, 40115, Indonesia*
*Corresponding author's e-mail: * **muhammadzulfadhli23@gmail.com***

***Article Info***

***Article History:***
*Received: 3rd March 2025*
*Revised: 28th April 2025*
*Accepted: 25th July 2025*
*Available online: 24th November 2025*

***Keywords:***
*Categorical data;*
*Bernoulli distribution;*
*Binary Probit Regression (BPR;*
*Binary Logistic Regression (BLR);*
*Fourier Series Nonparametric Binary Logistic Regression (FSNBLR).*

***ABSTRACT***

*Diabetes mellitus is a chronic disease with a rising global prevalence, including in Indonesia. Early detection and accurate modeling are crucial for effective prevention and management. Binary Logistic Regression (BLR) is commonly used for binary outcome modeling; however, in practice, the relationship between binary outcomes and continuous predictors is often nonlinear, making BLR less suitable. To address these limitations, alternative methods such as Binary Probit Regression (BPR) and Flexible Semiparametric Nonlinear Binary Logistic Regression (FSNBLR) have been developed. This study aims to compare the performance of BLR, BPR, and FSNBLR models in classifying diabetes mellitus cases at Hajj General Hospital Surabaya. All three models were estimated using the Maximum Likelihood Estimation (MLE) method. Since the resulting estimators do not have closed-form solutions, numerical iteration using the Newton-Raphson method was applied. Model performance was assessed using Area Under the Curve (AUC), accuracy, sensitivity, and specificity. The FSNBLR model outperformed both the BLR and BPR models. It achieved the highest AUC value of 81.86%, while BLR (66.30%) and BPR (66.30%). That is indicated FSNBLR superior discriminative ability. In addition, the FSNBLR model recorded higher accuracy, sensitivity, and specificity compared to the other two models. The FSNBLR model demonstrated better predictive performance in identifying diabetes mellitus cases, especially in scenarios involving nonlinear relationships between predictors and the outcome variable. These findings suggest that flexible semiparametric approaches offer greater effectiveness in medical classification tasks, particularly for chronic conditions like diabetes mellitus.*

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution-ShareAlike 4.0 International License (https://creativecommons.org/licenses/by-sa/4.0/).

***How to cite this article:***
B. W. Otok, M. Zulfadhli, R. D. Pangesti, M. I. Kurniawan, A. E. P. Haryanto, Darwis and I. Kurniawan., "COMPARISON OF BINARY PROBIT REGRESSION AND FOURIER SERIES NONPARAMETRIC LOGISTIC REGRESSION IN MODELING DIABETES STATUS AT HAJJ GENERAL HOSPITAL SURABAYA," *BAREKENG: J. Math. & App.*, vol. 20, iss. 1, pp. 0255-0270, Mar, 2026.

*Copyright © 2026 Author(s)*
Journal homepage: *https://ojs3.unpatti.ac.id/index.php/barekeng/*
Journal e-mail: *barekeng.math@yahoo.com; barekeng.journal@mail.unpatti.ac.id*

**Research Article · Open Access**

## 1. INTRODUCTION

Binary Logistic Regression (BLR) is a statistical method used to model the relationship between a binary dependent variable and one or more independent variables. The model assumes a linear relationship between the logit (log-odds) of the probability of the event and the predictors, but is often insufficient to handle data that have probit links because the underlying distribution and link function differ, leading to potential bias and poor fit when the logit link is incorrectly imposed. To overcome these limitations, Binary Probit Regression (BPR) offers an approach that accounts. However, in some cases, the relationship between the dependent and independent variables may be more complex and require a nonparametric approach. One of the evolving methods using logit function is Nonparametric Binary Logistic Regression (NBLR) [1].

NBLR can be used to determine the relationship between the response and predictor variables when the function of the regression curve is unknown. The NBLR curve is assumed to be smooth in the sense that it is contained in a certain function space. The data were expected to find their own form of estimation, without being influenced by the subjective factors of the researcher. Thus, the NBLR approach is highly flexible and it can be implemented based on observed data using smoothing techniques. There are many smoothing techniques, including the Spline estimator [2], Fourier Series estimator [3], Wavelet estimator [4], Kernel estimator [5], Local Polynomial estimator [6], and Multivariate Adaptive Regression Splines (MARS) estimator [7], [8].

Spline estimators are used for data with changing patterns that depend on knot points [2] and are especially suitable when the underlying relationship is smooth but may have varying curvature across different regions. A local polynomial estimator that has been used to reduce the bias properties and asymptotic variance of the local polynomial estimator in nonparametric regression with more than one response variable [9], making it appropriate when modeling local trends in complex, multidimensional data. Wavelet estimator that has been used to to model observations of signals contaminated with Gaussian additive noise, particularly when the data contain localized features or abrupt changes that need to be captured at multiple scales [10]. Kernel estimators are preferred when smoothing noisy data where a simple, nonparametric local averaging is sufficient [5]. The Fourier Series estimator is used for patterned data that tend to repeat [3]. Among these estimators, the Fourier Series method was used to. This method is very specialized and well used in data cases in which the response and predictor variables exhibit a repeating pattern following a certain trend [11]. The Fourier Series estimator best optimizes the accuracy and computational cost of additive nonparametric regression models [12]. Not only predictors with one predictor variable (univariable) but also with many predictor variables (multivariable) [3], [10], [13], [14].

Fourier Series was first introduced by [3], and then [10] studied the Fourier series estimator in nonparametric regression. Furthermore, [15] applied the Fourier Series in semiparametric regression. [16] developed a birresponse semiparametric regression using Fourier Series, until it became a Fourier Series nonparametric regression mixture estimator by [17], [18], [19], and a Fourier Series semiparametric mixture estimator by [20]. However, previous studies that developed using this method only used quantitative data, such as [21], [22], [23], [24]. However, in reality, there is often a relationship between response and predictor, where the response is categorical data.

Some researchers have developed nonparametric regression estimators for categorical data, such as using Local Likelihood Logit Estimation [25], using the Decision Tree approach [26], and using the B-Spline function [27] and recently, researchers have developed estimators for nonparametric regression using categorical data, such as Fourier Series Nonparametric Logistic Regression (FSNBLR) [28]. In addition, Fourier Series has been explored as a smoothing technique in nonparametric regression due to its ability to represent complex, periodic, or oscillating relationships using a combination of sine and cosine functions. Although originally developed for continuous signals, the Fourier basis can be adapted to model nonlinear patterns in categorical response data by transforming predictor variables and capturing latent cyclic or wave-like structures in the data. This approach is particularly useful when the relationship between predictors and the categorical response is non-linear and non-monotonic, which is often the case in medical and behavioral data.

Previous studies have only compared conventional methods such as Binary Logistic Regression (BLR) and Binary Probit Regression (BPR), without including more recent methods like FSNBLR that have been developed to address nonlinear patterns in categorical data. Conventional methods like BLR and BPR assume a linear relationship between predictors and the transformed response variable, which makes them less capable of capturing complex or repeating patterns that often occur in real-world data, such as in medical conditions like diabetes mellitus. However, no previous study has compared BLR, BPR, and FSNBLR estimators. Therefore, this study aims to compare BLR, BPR, and FSNBLR methods in the case of diabetes mellitus at Hajj General Hospital Surabaya.

However, no previous study has compared BLR, BPR and FSNBLR estimator. Therefore, this study aims to compare BLR, BPR, and FSNBLR methods in the case of the diabetes mellitus in hajj general hospital Surabaya, to identify which method performs best in handling categorical response data exhibiting nonparametric patterns.

## 2. RESEARCH METHODS

In obtaining a BPR, BLR, and FSNBLR estimators for categorical data, several steps are required: building a BPR, BLR and FSNBLR model, then creating a Log Likelihood function and deriving it for each model parameter. Finally, numerical iterations were performed using the Newton–Raphson iteration.

### 2.1 Probability Distribution

Given $x_1, x_2, \ldots, x_p$ are as many as $p$ predictor variables. Furthermore, the variable $Y$ is a random Bernoulli distribution variable [1] with a probability distribution of

$$Y \sim B\big(1, \pi(x)\big), \pi(x) = \pi\big(x_1, x_2, \ldots, x_p\big)$$

where the success probability

$$P(Y_i = 1) = \pi(x_i)$$

and the unsuccessful probability

$$P(Y_i = 0) = 1 - \pi(x_i)$$

$\pi(x_i)$ is defined in the probability distribution function $P(Y_i = y_i)$, where $i$ is the number of observations $(i = 1,2,\ldots,n)$ as follows.

$$P(Y_i = y_i) = \pi(x_i)^{y_i}\big(1 - \pi(x_i)\big)^{1-y_i} = \left(\frac{\pi(x_i)}{1 - \pi(x_i)}\right)^{y_i}\big(1 - \pi(x_i)\big) \tag{1}$$

### 2.2 BLR Estimator

#### 2.2.1 Logit Function (Link Function)

Eq. (1) can be expressed as a natural logarithmic function

$$\ln P(Y_i = y_i) = y_i \ln\left(\frac{\pi(x_i)}{1 - \pi(x_i)}\right) + \ln\big(1 - \pi(x_i)\big). \tag{2}$$

When made in exponential form, Eq. (2) forms an exponential family distribution function as follows.

$$\exp(\ln P(Y_i = y_i)) = \exp\left(y_i \ln\left(\frac{\pi(x_i)}{1 - \pi(x_i)}\right) + \ln\big(1 - \pi(x_i)\big)\right), \tag{3}$$

where Eq. (3) the exponential family distribution function is defined as follows.

$$f(y_i, \theta) = \exp\left(\frac{y_i\, \theta - b(\theta)}{a(\emptyset)} + c(\theta, \emptyset)\right). \tag{4}$$

Therefore, its probability distribution function belongs to the exponential family of distribution functions.

$$P(Y_i = y_i) = \exp\left(\frac{y_i \ln\left(\frac{\pi(x_i)}{1 - \pi(x_i)}\right)}{1} + \ln\big(1 - \pi(x_i)\big)\right), \tag{5}$$

where

$$\theta \quad = \ln\left(\frac{\pi(x_i)}{1-\pi(x_i)}\right) \qquad a(\emptyset) \quad = 1$$

$$b(\theta) \ = \ln(1-\pi(\boldsymbol{x_i})) \qquad c(\theta,\emptyset) = 0.$$

The variable $\theta$ in Eq. (5) is a logit function, the logit function for the regression obtained is

$$\theta = \ln\left(\frac{\pi(\boldsymbol{x_i})}{1-\pi(\boldsymbol{x_i})}\right). \tag{6}$$

The logit function (link function) simplifies a long regression model and facilitates parameter estimation. To achieve this goal, logit transformation is performed.

### 2.2.2 Logit Transformation Model

The logit transformation model is defined as follows.

$$\ln\left(\frac{\pi(\boldsymbol{x_i})}{1-\pi(\boldsymbol{x_i})}\right) = f(x_{1i}, \ \dots \ , x_{pi}), \tag{7}$$

where $f(x_{1i}, \ \dots \ , x_{pi})$ in Eq. (7) as follows.

$$\ln\left(\frac{\pi(\boldsymbol{x_i})}{1-\pi(\boldsymbol{x_i})}\right) = \sum_{j=1}^{p}(\beta_0+\beta_j x_{ji})\,; \ i = 1,2,\dots n. \tag{8}$$

By using Eq. (8), BLR model is obtained as follows.

$$\pi(\boldsymbol{x_i}) = \frac{\exp\sum_{j=1}^{p}(\beta_0+\beta_j x_{ji})}{1+\exp\sum_{j=1}^{p}(\beta_0+\beta_j x_{ji})}\ ; \ i = 1,2,\dots,n, \tag{9}$$

where $\beta_0$ and $\beta_j$ , $j = 1,2,\dots,p$ are the model parameters of the logit function.

### 2.2.3 Likelihood Function $l(\boldsymbol{\beta})$

The form of the likelihood function $l(\beta)$

where

$$\beta = (\beta_0 \quad \beta_1 \quad \cdots \quad \beta_p),$$

is obtained using the Maximum Likelihood Estimation (MLE) method.

$$l(\beta) = \prod_{i=1}^{n} P(Y_i = \ y_i) = \pi(x_i)^{\sum_{i=1}^{n} y_i}\left(1-\pi(x_i)\right)^{n-\sum_{i=1}^{n} y_i}. \tag{10}$$

Parameter estimation in logistic regression can be performed using the MLE method by maximizing the first derivative of the log likelihood function.

### 2.2.4 Log-Likelihood Function $L(\boldsymbol{\beta})$

The likelihood Eq. (10) can be easily maximized as follows $\ln l(\beta)$.

$$\begin{aligned} L(\beta) = \ln[l(\beta)] &= \sum\nolimits_{i=1}^{n} y_i \ln[\pi(x_i)] + \sum\nolimits_{i=1}^{n} (1-y_i)\ln[1-\pi(x_i)] \\ &= \sum\nolimits_{i=1}^{n} \left\{y_i\left(f(x_{1i}, \ \dots \ , x_{pi})\right) - \ln\left[1+\exp(f(x_{1i}, \ \dots \ , x_{pi}))\right]\right\}. \end{aligned} \tag{11}$$

### 2.2.5 Newton-Raphson Iteration

$$\beta^{(t+1)} = \beta^{(t)} - \left(H(\beta)^{(t)}\right)^{-1} g(\beta)^{(t)}, \tag{12}$$

where $\beta^{(t)}$ is the $\beta$ of the $t$-th iteration, $t = 1,2,\dots$, is convergent.

$$\beta^{(t)} = ({\beta_1}^{(t)} \quad {\beta_2}^{(t)} \quad \cdots \quad {\beta_p}^{(t)}).$$

$\boldsymbol{g}(\boldsymbol{\beta})$ is the gradient vector of $\beta$

$$\boldsymbol{g}(\boldsymbol{\beta}) = \left(\frac{\partial L(\boldsymbol{\beta})}{\partial \beta_1}, \frac{\partial L(\boldsymbol{\beta})}{\partial \beta_2}, \cdots, \frac{\partial L(\boldsymbol{\beta})}{\partial \beta_p}\right)^{\boldsymbol{T}}, \tag{13}$$

and $\boldsymbol{H}(\boldsymbol{\beta})$ is the Hessian matrix of $\beta$ in Eq. (12), with the following equation.

$$\boldsymbol{H}(\boldsymbol{\beta}) = \begin{bmatrix} \frac{\partial^2 L(\boldsymbol{\beta})}{\partial {\beta_1}^2} & \frac{\partial^2 L(\boldsymbol{\beta})}{\partial \beta_1 \partial \beta_2} & \cdots & \frac{\partial^2 L(\boldsymbol{\beta})}{\partial \beta_1 \partial \beta_p} \\ \frac{\partial^2 L(\boldsymbol{\beta})}{\partial \beta_2 \partial \beta_1} & \frac{\partial^2 L(\boldsymbol{\beta})}{\partial {\beta_2}^2} & \cdots & \frac{\partial^2 L(\boldsymbol{\beta})}{\partial \beta_2 \partial \beta_p} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 L(\boldsymbol{\beta})}{\partial \beta_p \partial \beta_1} & \frac{\partial^2 L(\boldsymbol{\beta})}{\partial \beta_p \partial \beta_2} & \cdots & \frac{\partial^2 L(\boldsymbol{\beta})}{\partial {\beta_p}^2} \end{bmatrix}. \tag{14}$$

### 2.2.6 Estimator $\widehat{\boldsymbol{\beta}}$

From the Newton-Raphson iteration equation, $\hat{\beta}$ will be obtained when

$$\left|\beta^{(t+1)} - \beta^{(t)}\right| < \varepsilon,\ \varepsilon = 0.000001. \tag{15}$$

Thus, the estimator $\hat{\beta}$ is given by

$$\hat{\beta} = (\hat{\beta}_1 \quad \hat{\beta}_2 \quad \cdots \quad \hat{\beta}_p).$$

Based on the result of the estimator $\hat{\beta}$, BLR model Eq. (16) can be written:

$$\hat{\pi}(\boldsymbol{x_i}) = \frac{\exp(\hat{\beta}_0 + \hat{\beta}_1 x_{1i} + \hat{\beta}_2 x_{2i} + \cdots + \hat{\beta}_p x_{pi})}{1 + \exp(\hat{\beta}_0 + \hat{\beta}_1 x_{1i} + \hat{\beta}_2 x_{2i} + \cdots + \hat{\beta}_p x_{pi})}, \tag{16}$$

where, $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_p$ are the estimator model of the logit function and $p$ is the number of predictor variables.

## 2.3 BPR Estimator

### 2.3.1 Probit Function

Eq. (1) can be expressed as a probit function ($\Phi$)

$$\Phi(y) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2\sigma^2}(y-\mu)^2\right]. \tag{17}$$

### 2.3.2 Probit Transformation Model

The probit transformation model is defined as follows.

$$\Phi^{-1}(x_i) = x_i^T \gamma, i = 1, 2, \dots, n; \tag{18}$$

$$z = \Phi^{-1}(\pi(x_i)) = x_i^T \gamma; \tag{19}$$

$$\pi(x_i) = \Phi(x_i^T \gamma); \tag{20}$$

$$z = \Phi^{-1}(\pi(x_i)) = \beta_0 + \beta_1 x_1 + \cdots + \beta_p x_p. \tag{21}$$

### 2.3.3 Likelihood Function $\boldsymbol{l}(\boldsymbol{\gamma})$

The form of the likelihood function $l(\gamma)$

where

$$\gamma = (\gamma_0 \quad \gamma_1 \quad \cdots \quad \gamma_p), \tag{22}$$

is obtained using the Maximum Likelihood Estimation (MLE) method.

$$l(\gamma) = \prod_{i=1}^{n} \mathrm{P}\,(y_i = 1|x_i)^{y_i}[1 - P(y_i = 1|x_i)]^{1-y_i}, \tag{23}$$

$$l(\gamma) = \prod_{i=1}^{n} \Phi\,(x_i^T\gamma)^{y_i}\left[1 - \Phi(x_i^T\gamma)\right]^{1-y_i}. \tag{24}$$

### 2.3.4 Log-Likelihood Function $L(\gamma)$

The likelihood function Eq. (24) can be easily maximized as follows $\ln l(\gamma)$.

$$L(\gamma) = \ln l(\gamma) = \sum_{i=1}^{n} [y_i \ln(\Phi) + (1 - y_i)\ln(1 - \Phi)]\,. \tag{25}$$

### 2.3.5 Newton-Raphson Iteration

$$\gamma^{(t+1)} = \gamma^{(t)} - \left(H(\gamma)^{(t)}\right)^{-1} g(\gamma)^{(t)}, \tag{26}$$

where $\gamma^{(t)}$ is the $\gamma$ of the $t$-th iteration, $t = 1,2,\ldots$, is convergent.

$$\gamma^{(t)} = \begin{pmatrix}\gamma_1{}^{(t)} & \gamma_2{}^{(t)} & \cdots & \gamma_p{}^{(t)}\end{pmatrix}.$$

$\boldsymbol{g(\gamma)}$ is the gradient vector of $\gamma$

$$\boldsymbol{g(\gamma)} = \left(\frac{\partial \boldsymbol{L(\gamma)}}{\partial \boldsymbol{\beta_1}}, \frac{\partial \boldsymbol{L(\gamma)}}{\partial \boldsymbol{\beta_2}}, \ldots, \frac{\partial \boldsymbol{L(\gamma)}}{\partial \boldsymbol{\beta_p}}\right)^{\boldsymbol{T}}, \tag{27}$$

and $\boldsymbol{H(\gamma)}$ is the Hessian matrix of $\gamma$ in Eq. (28), with the following equation.

$$\boldsymbol{H(\gamma)} = \begin{bmatrix} \frac{\partial^2 L(\gamma)}{\partial \gamma_1{}^2} & \frac{\partial^2 L(\gamma)}{\partial \gamma_1 \partial \gamma_2} & \cdots & \frac{\partial^2 L(\gamma)}{\partial \gamma_1 \partial \gamma_p} \\ \frac{\partial^2 L(\gamma)}{\partial \gamma_2 \partial \gamma_1} & \frac{\partial^2 L(\gamma)}{\partial \gamma_2{}^2} & \cdots & \frac{\partial^2 L(\gamma)}{\partial \gamma_2 \partial \gamma_p} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 L(\gamma)}{\partial \gamma_p \partial \gamma_1} & \frac{\partial^2 L(\gamma)}{\partial \gamma_p \partial \gamma_2} & \cdots & \frac{\partial^2 L(\gamma)}{\partial \gamma_p{}^2} \end{bmatrix}. \tag{28}$$

### 2.3.6 Estimator $\hat{\gamma}$

From the Newton-Raphson iteration equation, $\hat{\gamma}$ will be obtained when

$$\left|\gamma^{(t+1)} - \gamma^{(t)}\right| < \varepsilon,\ \varepsilon = 0.000001. \tag{29}$$

Thus, the estimator $\hat{\gamma}$ is given by

$$\hat{\gamma} = \begin{pmatrix}\hat{\gamma}_1 & \hat{\gamma}_2 & \cdots & \hat{\gamma}_p\end{pmatrix}.$$

Based on the result of the estimator $\hat{\gamma}$, BPR model Eq. (30) can be written:

$$P(y_i = 1) = \Phi\left(\hat{\gamma}_0 + \hat{\gamma}_1 x_1 + \hat{\gamma}_2 x_2 + \cdots + \hat{\gamma}_p x_p\right), i = 1, 2, \ldots, n;$$

$$P(y_i = 0) = 1 - \Phi\left(\hat{\gamma}_0 + \hat{\gamma}_1 x_1 + \hat{\gamma}_2 x_2 + \cdots + \hat{\gamma}_p x_p\right), i = 1, 2, \ldots, n; \tag{30}$$

where, $\hat{\gamma}_0, \hat{\gamma}_1, \hat{\gamma}_2, \ldots, \hat{\gamma}_p$ are the estimator model of the probit function and $p$ is the number of predictor variables.

## 2.4 FSNBLR Estimator

### 2.4.1 Logit Transformation Model

The logit transformation model is defined as follows.

$$\ln\left(\frac{\pi(\boldsymbol{x_i})}{1-\pi(\boldsymbol{x_i})}\right) = f(x_{1i}, \ \dots \ , x_{pi}), \tag{31}$$

where $f$ is a regression equation or regression function (regression curve) that follows an additive model. Since

$f(x_{1i}, \ \dots \ , x_{pi})$ in Eq. (31) can be approximated by a multivariable Fourier Series function as follows.

$$\ln\left(\frac{\pi(\boldsymbol{x_i})}{1-\pi(\boldsymbol{x_i})}\right) = \sum_{j=1}^{p}\left(b_j x_{ji} + \frac{1}{2}a_{0j} + \sum_{k=1}^{K} a_{kj}\cos kx_{ji}\right); \ i = 1,2,\dots n. \tag{32}$$

By using Eq. (32), FSNBLR model is obtained as follows.

$$\pi(x_i) = \frac{\exp\sum_{j=1}^{p}\left(b_j x_{ji} + \frac{1}{2}a_{0j} + \sum_{k=1}^{K} a_{kj}\cos kx_{ji}\right)}{1+\exp\sum_{j=1}^{p}\left(b_j x_{ji} + \frac{1}{2}a_{0j} + \sum_{k=1}^{K} a_{kj}\cos kx_{ji}\right)}; \ i = 1,2,\dots,n; \tag{33}$$

where, $b_j, a_{0j}$ and $a_{kj}$ , $j = 1,2,\dots,p$ , $k = 1,2,\dots,K$ are the model parameters of the Fourier Series function.

### 2.4.2 Likelihood Function $l(\boldsymbol{\beta})$

The form of the likelihood function $l(\beta)$

where

$$\beta = (b_1 \quad a_{01} \quad a_{11} \quad \dots \quad a_{K1} \ \vdots \ \cdots \ \vdots \ b_p \quad a_{0p} \quad a_{1p} \quad \dots \quad a_{Kp}),$$

is obtained using the Maximum Likelihood Estimation (MLE) method.

$$l(\beta) = \prod_{i=1}^{n} P(Y_i = \ y_i) = \pi(x_i)^{\sum_{i=1}^{n} y_i}\left(1-\pi(x_i)\right)^{n-\sum_{i=1}^{n} y_i}. \tag{34}$$

Parameter estimation in logistic regression can be performed using the MLE method by maximizing the first derivative of the log likelihood function. The likelihood function Eq. (34) can be easily maximized as follows $\ln l(\beta)$.

### 2.4.3 Log-Likelihood Function $L(\boldsymbol{\beta})$

$$\begin{aligned} L(\beta) = \ln[l(\beta)] &= \sum\nolimits_{i=1}^{n} y_i \ln[\pi(x_i)\,] + \sum\nolimits_{i=1}^{n} (1-y_i)\ln[1-\pi(x_i)\,] \\ &= \sum\nolimits_{i=1}^{n}\left\{y_i\left(f(x_{1i}, \ \dots \ , x_{pi})\right) - \ln[1+\exp(f(x_{1i}, \ \dots \ , x_{pi}))]\right\}. \end{aligned} \tag{35}$$

The estimator $\hat{\beta}$ is obtained by partially deriving function Eq. (35) relative to $b_j, a_{0j}, a_{kj}$ and then equating to 0

$$\frac{\partial L(\beta)}{\partial b_j} = 0 \ ; \ j = 1,2,\dots,p;$$

$$\frac{\partial L(\beta)}{\partial a_{0j}} = 0 \ ; \ j = 1,2,\dots,p;$$

$$\frac{\partial L(\beta)}{\partial a_{kj}} = 0 \ ; \ k = 1,2,\dots,K \ ; \ j = 1,2,\dots,p.$$

The estimator $\hat{b}$ will be obtained using Eq. (36).

$$\sum\nolimits_{i=1}^{n}\{(y_i - \ \pi(x_i))x_{ji}\} = 0. \tag{36}$$

The estimator $\hat{a}_0$ will be obtained using Eq. (37).

$$\sum\nolimits_{i=1}^{n}\left\{\frac{1}{2}(y_i - \ \pi(x_i))\right\} = 0. \tag{37}$$

The estimator $\hat{a}_k$ will be obtained using Eq. (38).

$$\sum_{i=1}^{n}\left\{\sum_{k=1}^{K}\cos kx_{ji}\left(y_i - \pi(x_i)\right)\right\} = 0. \tag{38}$$

### 2.4.4 Newton-Raphson Iteration

The derivative of $L(\beta)$ Eq. (35) against $b_j, a_{0j}, a_{kj}$ that has been made in the implicit equation, gives results that are not closed form, so it is necessary to continue with the numerical iteration method using the Newton–Raphson method.

$$\beta^{(t+1)} = \beta^{(t)} - \left(H(\beta)^{(t)}\right)^{-1} g(\beta)^{(t)}, \tag{39}$$

where $\beta^{(t)}$ is the $\beta$ of the $t$-th iteration, $t = 1,2,\dots$, is convergent.

$$\beta^{(t)} = \left(b_1{}^{(t)} \quad a_{01}{}^{(t)} \quad a_{11}{}^{(t)} \quad \dots \quad a_{K1}{}^{(t)} \quad \vdots \quad \cdots \quad \vdots \quad b_p{}^{(t)} \quad a_{0p}{}^{(t)} \quad a_{1p}{}^{(t)} \quad \dots \quad a_{Kp}{}^{(t)}\right).$$

$\boldsymbol{g(\beta)}$ is the gradient vector of $\beta$

$$\boldsymbol{g(\beta)} = \left(\frac{\partial L(\boldsymbol{\beta})}{\partial b_1}, \frac{\partial L(\boldsymbol{\beta})}{\partial a_{01}}, \frac{\partial L(\boldsymbol{\beta})}{\partial a_{11}}, \cdots, \frac{\partial L(\boldsymbol{\beta})}{\partial a_{K1}}, \cdots, \frac{\partial L(\boldsymbol{\beta})}{\partial b_p}, \frac{\partial L(\boldsymbol{\beta})}{\partial a_{0p}}, \frac{\partial L(\boldsymbol{\beta})}{\partial a_{1p}}, \cdots, \frac{\partial L(\boldsymbol{\beta})}{\partial a_{Kp}}\right)^{T}, \tag{40}$$

and $\boldsymbol{H(\beta)}$ is the Hessian matrix of $\beta$ in Eq. (40), with the following equation.

$$\boldsymbol{H(\beta)} = \begin{bmatrix} \frac{\partial^2 L(\boldsymbol{\beta})}{\partial b_1{}^2} & \frac{\partial^2 L(\boldsymbol{\beta})}{\partial b_1 \partial a_{01}} & \cdots & \frac{\partial^2 L(\boldsymbol{\beta})}{\partial b_1 \partial a_{Kp}} \\ \frac{\partial^2 L(\boldsymbol{\beta})}{\partial a_{01} \partial b_1} & \frac{\partial^2 L(\boldsymbol{\beta})}{\partial a_{01}{}^2} & \cdots & \frac{\partial^2 L(\boldsymbol{\beta})}{\partial a_{01} \partial a_{Kp}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 L(\boldsymbol{\beta})}{\partial a_{Kp} \partial b_1} & \frac{\partial^2 L(\boldsymbol{\beta})}{\partial a_{Kp} \partial a_{01}} & \cdots & \frac{\partial^2 L(\boldsymbol{\beta})}{\partial a_{Kp}{}^2} \end{bmatrix}. \tag{41}$$

The elements of vector $\boldsymbol{g(\beta)}$ Eq. (40) are obtained from the first derivative of function $L(\beta)$ with respect to $b_j, a_{0j}, a_{kj}$, while elements of matrix $\boldsymbol{H(\beta)}$ Eq. (41) are obtained from the second derivative of function $L(\beta)$ with respect to $b_u, a_{0u}, a_{ku}$.

Second derivative of $L(\beta)$ function with respect to $b_u$

$$\frac{\partial^2 L(\beta)}{\partial b_u \partial b_j} = -\sum_{i=1}^{n} x_{ji}x_{ui}\, \pi(x_i)\left(1 - \pi(x_i)\right). \tag{42}$$

In the same manner as Eq. (42), the second derivative of the parameter combination is obtained as follows.

$$\frac{\partial^2 L(\beta)}{\partial a_{ku} \partial b_j} = -\sum_{i=1}^{n}\sum_{k=1}^{K} \pi(x_i)\left(1 - \pi(x_i)\right)x_{ji} \cos kx_{ui}\,. \tag{43}$$

Second derivative of $L(\beta)$ function with respect to $a_{0u}$

$$\frac{\partial^2 L(\beta)}{\partial a_{0u} \partial a_{0j}} = -\frac{1}{4}\sum_{i=1}^{n} \pi(x_i)\left(1 - \pi(x_i)\right). \tag{44}$$

In the same manner as Eq. (44), the second derivative of the parameter combination is obtained as follows.

$$\frac{\partial^2 L(\beta)}{\partial a_{ku} \partial a_{0j}} = -\frac{1}{2}\sum_{i=1}^{n}\sum_{k=1}^{K} \pi(x_i)\left(1 - \pi(x_i)\right) \cos kx_{ui}\,. \tag{45}$$

Second derivative of $L(\beta)$ function with respect to $a_{ku}$

$$\frac{\partial^2 L(\beta)}{\partial a_{ku} \partial a_{kj}} = -\sum_{i=1}^{n}\sum_{k=1}^{K} \cos kx_{ji} \sum_{k=1}^{K} \cos kx_{ui}\, \pi(\boldsymbol{x_i})\left(1 - \pi(\boldsymbol{x_i})\right). \tag{46}$$

In the same manner as Eq. (46), the second derivative of the parameter combination is obtained as follows.

$$\frac{\partial^2 L(\boldsymbol{\beta})}{\partial a_{0u}\partial a_{kj}} = -\frac{1}{2}\sum_{i=1}^{n}\sum_{k=1}^{K}\pi(\boldsymbol{x_i})\big(1-\pi(\boldsymbol{x_i})\big)\cos kx_{ji}\,. \tag{47}$$

where, $b_j, a_{0j}$ and $a_{kj}$ , $j, u = 1,2, \dots, p$ , $j \neq u$, $k = 1,2, \dots, K$ are the model parameters of the Fourier Series function.

### 2.4.5 Estimator $\hat{\boldsymbol{\beta}}$

From the Newton-Raphson iteration equation, $\hat{\beta}$ will be obtained when

$$\left|\beta^{(t+1)} - \beta^{(t)}\right| < \varepsilon,\ \varepsilon = 0.000001 \tag{48}$$

Thus, the estimator $\hat{\beta}$ is given by

$$\hat{\beta} = \begin{pmatrix}\hat{b}_1 & \hat{a}_{0_1} & \hat{a}_{1_1} & \dots & \hat{a}_{K_1} & \vdots & \cdots & \vdots & \hat{b}_p & \hat{a}_{0_p} & \hat{a}_{1_p} & \dots & \hat{a}_{K_p}\end{pmatrix}.$$

Based on the result of the estimator $\hat{\beta}$, FSNBLR model Eq. (49) can be written:

$$\hat{\pi}(x_i) = \frac{\exp\left(\hat{b}_1x_{1i} + \frac{1}{2}\hat{a}_{01} + \hat{a}_{11}\cos x_{1i} + \cdots + \hat{a}_{K1}\cos Kx_{1i} + \cdots + \hat{b}_px_{pi} + \frac{1}{2}\hat{a}_{0p} + \hat{a}_{1p}\cos x_{pi} + \cdots + \hat{a}_{Kp}\cos Kx_{pi}\right)}{1 + \exp\left(\hat{b}_1x_{1i} + \frac{1}{2}\hat{a}_{01} + \hat{a}_{11}\cos x_{1i} + \cdots + \hat{a}_{K1}\cos Kx_{1i} + \cdots + \hat{b}_px_{pi} + \frac{1}{2}\hat{a}_{0p} + \hat{a}_{1p}\cos x_{pi} + \cdots + \hat{a}_{Kp}\cos Kx_{pi}\right)}, \tag{49}$$

where, $\hat{b}_1, \hat{a}_{01}$ and $\hat{a}_{k1}$ are the estimator model of the Fourier Series function for predictor variable $x_1$, while $\hat{b}_p, \hat{a}_{0p}$ and $\hat{a}_{kp}$ are for predictor variable $x_p$, $K$ is the number of oscillation parameters and $p$ is the number of predictor variables.

## 2.5 Hypothesis Test for Parameter Model

Hypothesis test for parameter model consists of simultaneous and partial tests. Hypothesis test for simultaneous uses the Likelihood Ratio Test (LRT) and hypothesis test for partial uses the Wald test.

### 2.5.1 Simultaneous

The simultaneous test is conducted to determine the significance of parameter $\theta$ as a whole or simultaneously, where $\theta$ is parameters for the model.

Hypothesis:

$H_0 :\ \theta_1 = \theta_2 = \theta_3 = \cdots = \theta_j = \cdots = \theta_k = 0,$
$H_1 :$ There is at least one $\theta_j \neq 0.$

Statistics test for simultaneous test Eq. (50):

$$G^2 = -2\sum_{i=1}^{n}\left[y_i \ln\left(\frac{\hat{\pi}(x_i)}{y_i}\right) + (1-y_i)\ln\left(\frac{1-\hat{\pi}(x_i)}{1-y_i}\right)\right]. \tag{50}$$

Decision:
Reject $H_0$ when $G^2 > \chi^2_{(v,a)}$ or $p - value < a$.

### 2.5.2 Partial

Hypothesis:

$H_0 :\ \theta_j = 0\ , j = 1,2, \dots, k;$
$H_1 :\ \theta_j \neq 0.$

Statistics test for partial test Eq. (51):

$$W = \frac{\hat{\theta}_j}{\widehat{SE}(\hat{\theta}_j)}. \tag{51}$$

Decision:
Reject $H_0$ when $W > \chi^2_{(v,a)}$ or $p - value < a$.

## 3. RESULTS AND DISCUSSION

In applying the BLR, BPR, and FSNBLR methods, we use application data the status of diabetes mellitus. The data used is secondary data sourced from Internal Medicine Clinic of Hajj General Hospital Surabaya which was carried out during August 2018. The data consists of 1 response variable $(y)$ and 3 predictor variables $(x)$. The variables are detailed in Table 1.

**Table 1. Variable Description**

| Variable | Notation | Description | Unit | Scale |
|---|---|---|---|---|
| Response | $y$ | Status of Type 2 Diabetes Mellitus | 0 = Doesn't have Diabetes Mellitus<br>1 = Has Diabetes Mellitus | Nominal |
| Predictor | $x_1$ | Age | Year | Ratio |
| | $x_2$ | Body Mass Index | $kg/m^2$ | Ratio |
| | $x_3$ | Abdominal Circumference | cm | Ratio |

Based on Table 1, these variables were selected based on medical relevance and availability in patient records at Hajj General Hospital Surabaya. They are used to model the probability of having Type 2 Diabetes Mellitus using BLR, BPR, and FSNBLR methods. 60 patients consist of 39 patients diagnosed with diabetes mellitus and 21 non-patients without diabetes mellitus which as shown in Fig. 1.

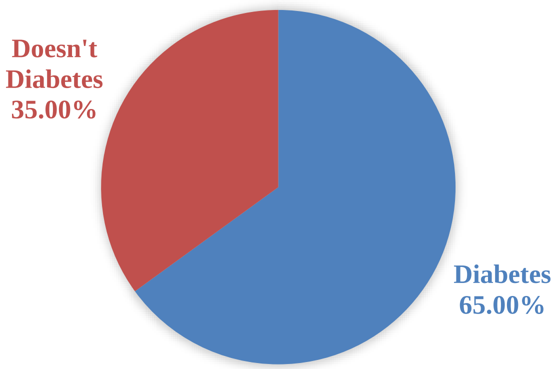


**Figure 1**. Status of Type 2 Diabetes Mellitus

### 3.1 Descriptive Analytics

Descriptive analysis is used to determine the characteristics of the data for each variable as follows Table 2.

**Table 2**. Descriptive Statistics of Research Variables

| Category | Variable | Mean | StDev | Minimum | Maximum |
|---|---|---|---|---|---|
| Doesn't have Diabetes Mellitus | Age | 47.0952 | 16.4405 | 17 | 71 |
| | IMT | 22.3681 | 4.47414 | 16.02 | 31.25 |
| | Abdominal Mass | 84.9524 | 13.1851 | 64 | 119 |
| Has Diabetes Mellitus | Age | 62.9487 | 8.72066 | 51 | 83 |
| | IMT | 25.5018 | 3.54516 | 18.49 | 33.78 |
| | Abdominal Mass | 93.6410 | 8.39952 | 82 | 115 |

Table 2 provides information about the characteristics of the variables, which is the status of diabetes mellitus. In addition, it is obtained that each variable does not have missing values and there is no multicollinearity between predictor variables. The conceptual predictor variable used in this study is as follows in Fig. 2.

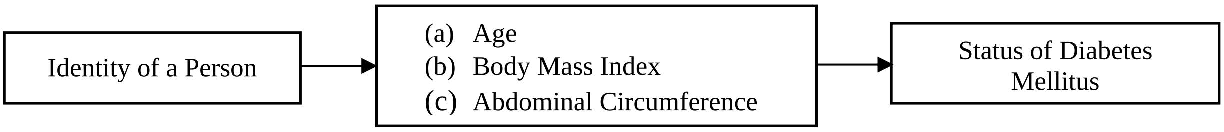


**Figure 2**. Conceptual Diagram of Variables

Based on Fig. 2, diabetes mellitus is a problem that covers many aspects. One of them is identity of a person. Based on Central Bureau of Statistics, identity of a person consists of age, body mass index, and abdominal circumference.

We created a scatterplot for each predictor variable that was built into several groups versus the presentation of the number of has diabetes mellitus $(y = 1)$ in each group to identify the relationship. The presentation represents the proportion of patients diagnosed with diabetes mellitus relative to the total number of patients in each group. The scatterplot is presented in Fig. 3 as follows.

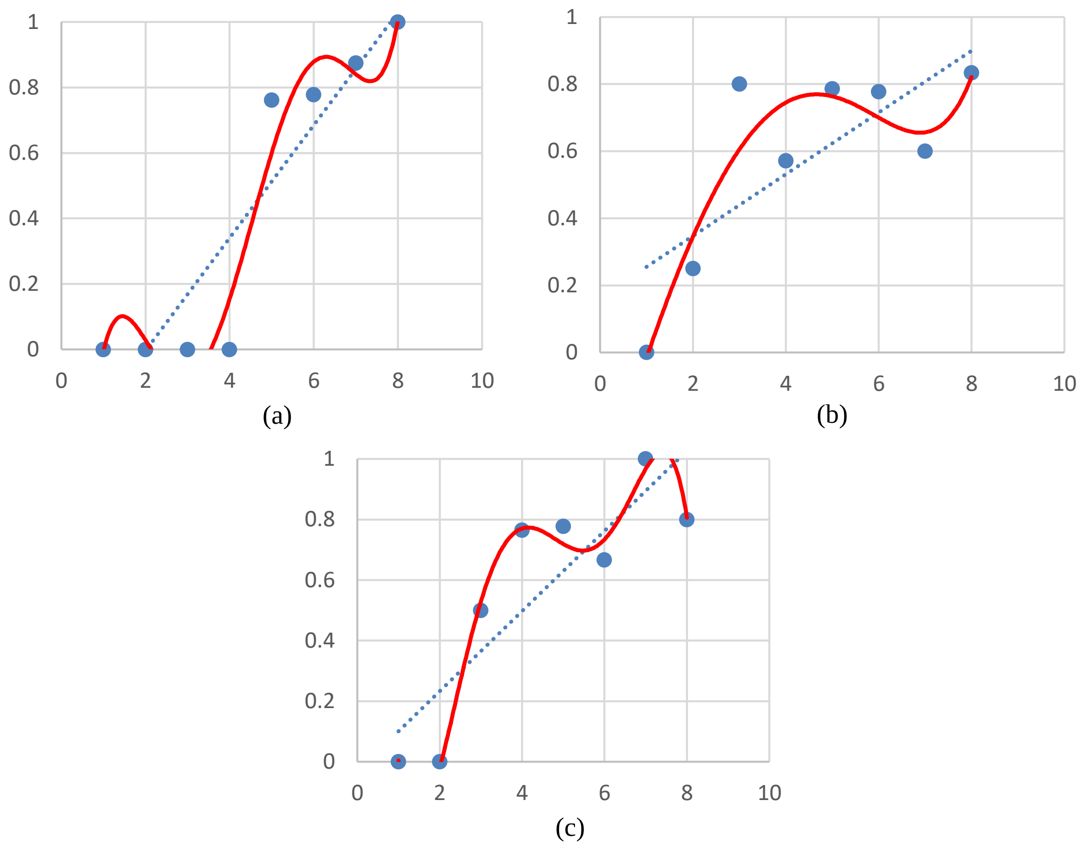


**Figure 3**. Scatterplots of Several Data Groups Versus the Number of Has Diabetes Mellitus in the Group (a) Age, (b) Body Mass Index, (c) Abdominal Circumference

Based on Fig. 3, The probability of a has diabetes mellitus $(y = 1)$ for variable $x_1$, $x_2$, and $x_3$ have a repeating pattern and follows a upward trend line. Thus, the logit function that assumes a linear pattern does not describe the pattern formed in this case.

For modeling status of diabetes mellitus, we use BLR, BPR, and FSNBLR methods. Parameter estimation and significant parameter in the model result can be seen as follows.

### 3.2 BLR Model

The BLR model follows Eq. (9) as follows.

$$\pi(x_i) = \frac{\exp \sum_{j=1}^{p}(\beta_0+\beta_j x_{ji})}{1+\exp \sum_{j=1}^{p}(\beta_0+\beta_j x_{ji})} \; ; \; i = 1,2,\ldots,n$$

where, $\beta_0$ and $\beta_j$ , $j = 1,2,\dots,p$ are the model parameters of the logit function.

### 3.2.1 Parameter Estimation in BLR Model Results

Based on BLR model in Eq. (16), the results of parameter estimation in the BLR model for data on diabetes mellitus are as follows.

$$\hat{\pi}(\boldsymbol{x_i}) = \frac{\exp(-11.781 + 0.118x_{1i} + 0.132x_{2i} + 0.026x_{3i})}{1 + \exp(-11.781 + 0.118x_{1i} + 0.132x_{2i} + 0.026x_{3i})}$$

### 3.2.2 Significant Parameter in BLR Model Results

Based on parameter estimation BLR model, the results of significant parameter in the BLR model for data on diabetes mellitus can be seen in Table 3.

**Table 3**. Significant Parameter in BLR Model

| | **Estimate** | **Std. Error** | **z value** | **Pr(>\|z\|)** |
|---|---|---|---|---|
| Intercept | -11.781 | 4.40281 | -2.676 | 0.00745 |
| $x_1$ | 0.118 | 0.04147 | 2.847 | 0.00441 |
| $x_2$ | 0.132 | 0.11749 | 1.124 | 0.26102 |
| $x_3$ | 0.026 | 0.05101 | 0.52 | 0.60282 |

Table 3 results show that only the age variable is significant in the model. Therefore, age affects a person's diabetes mellitus status.

### 3.3 BPR Model

The BPR model is as follows Eq. (18).

$$\Phi(y) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left[-\frac{1}{2\sigma^2}(y-\mu)^2\right]$$

### 3.3.1 Parameter Estimation in BPR Model Results

Based on BPR model in Eq. (30), parameter estimation results on the BPR model for data on diabetes mellitus are as follows.

$$P(y_i = 1) = \Phi(6.997 + 0.071x_1 + 0.078x_2 + 0.015x_3)$$
$$P(y_i = 0) = 1 - \Phi(6.997 + 0.071x_1 + 0.078x_2 + 0.015x_3)$$

### 3.3.2 Significant Parameter in BPR Model Results

Based on parameter estimation BPR model, the results of significant parameter in the BPR model for data on diabetes mellitus can be seen in Table 4.

**Table 4**. Significant Parameter in BPR Model

| | **Estimate** | **Std. Error** | **z value** | **Pr(>\|z\|)** |
|---|---|---|---|---|
| Intercept | 6.997 | 2.42416 | -2.886 | 0.0039 |
| $x_1$ | 0.071 | 0.02263 | 3.13 | 0.00175 |
| $x_2$ | 0.078 | 0.06947 | 1.133 | 0.25738 |
| $x_3$ | 0.015 | 0.02947 | 0.516 | 0.60551 |

Table 4 results show that only variables age are significant in the model. So age affects a person's diabetes mellitus status.

### 3.4 FSNBLR Model

### 3.4.1 Selecting Optimal Oscillation Parameters

The oscillation parameters in the FSNBLR model were selected based on the smallest AIC value. The number of oscillation parameters used in this study was limited to produce a model that is not too complicated and provides appropriate significance results. With the help of the R algorithm, the AIC results for each combination of oscillation parameters in the model are given in Table 5.

**Table 5**. Minimum AIC Results for Each Number of Oscillation Parameter

| Number of Oscillation Parameter | Oscillation Parameter Combination | | | AIC (K) |
|---|---|---|---|---|
| | $x_1$ | $x_2$ | $x_3$ | |
| $K=1$ | 1 | 1 | 1 | 61.543 |
| $K=2$ | 1 | 2 | 1 | 57.837 |
| $K=3$ | 1 | 2 | 1 | **57.837** |

Based on Table 5, the model with a combination of oscillation parameters $x_1=1, x_2=2, x_3=1$ is the FSNBLR model with optimal oscillation parameters because it has the smallest AIC value. This combination was identified by evaluating multiple parameter settings and selecting the one that yielded the lowest AIC value at each level of oscillation complexity, ensuring the most optimal balance between model fit and complexity.

### 3.4.2 Parameter Estimation in FSNBLR Model Results

Based on the FSNBLR model in Eq. (33), the results of parameter estimation in the FSNBLR model Eq. (49) for data on diabetes mellitus are as follows.

$$\hat{\pi}(x_i)=\frac{\exp(-15.56+0.16x_{1i}+0.81\cos x_{1i}+0.22x_{2i}-0.65\cos x_{2i}+1.40\cos 2x_{2i}+0.01x_{3i}-1.40\cos x_{3i})}{1+\exp(-15.56+0.16x_{1i}+0.81\cos x_{1i}+0.22x_{2i}-0.65\cos x_{2i}+1.40\cos 2x_{2i}+0.01x_{3i}-1.40\cos x_{3i})}$$

More details can be seen in Table 6.

**Table 6**. Parameter Estimation in FSNBLR

| Parameters | Estimations |
|---|---|
| $\beta_0$ | -15.564 |
| $b_1$ | 0.165 |
| $a_{1,1}$ | 0.813 |
| $b_2$ | 0.229 |
| $a_{1,2}$ | -0.659 |
| $a_{2,2}$ | 1.408 |
| $b_3$ | 0.016 |
| $a_{1,3}$ | -1.401 |

### 3.4.3 Significant Parameter in FSNBLR Model Results

Based on parameter estimation FSNBLR model, the results of significant parameter in the FSNBLR model for data on diabetes mellitus can be seen in Table 7.

**Table 7**. Significant Parameter in FSNBLR Model

| | Estimate | Std. Error | z value | Pr(>\|z\|) |
|---|---|---|---|---|
| Intercept | -15.565 | 5.9572 | -2.613 | 0.00898 |
| $x_1$ | 0.16555 | 0.05548 | 2.984 | 0.00285 |
| $x_2$ | 0.81312 | 0.70369 | 1.156 | 0.24788 |
| $x_3$ | 0.2293 | 0.15428 | 1.486 | 0.13721 |
| $x_4$ | -0.6596 | 0.66471 | -0.992 | 0.32105 |
| $x_5$ | 1.40888 | 0.65126 | 2.163 | 0.03052 |
| $x_6$ | 0.01656 | 0.05883 | 0.281 | 0.77836 |
| $x_7$ | -1.40055 | 0.66829 | -2.096 | 0.03611 |

Table 7 results show that variables age, body mass index, and abdominal circumference are significant in the model. Thus, age affects a person's diabetes mellitus status.

## 3.5 Comparison of BLR, GWBLR, FSNBLR

### 3.5.1 Getting the Best Model Based on Deviance Value

The regression model chosen is the model that has the smallest deviance value. Using the deviance statistical test, the following results are obtained in Table 8.

**Table 8**. Comparison of Deviance Values

| Methods | Deviance Values |
|---|---|
| BLR | 53.007 |
| BPR | 52.728 |
| FSNBLR | **41.837** |

Based on Table 8, the deviance value for the FSNBLR (41.837) was smaller than that for the BLR (53.007) and BPR (52.728). Therefore, the FSNBLR model is the best model for data on the status of diabetes mellitus because has the smallest deviance value.

### Getting the Best Classification Based on AUC & Press's Q Value

The selected FSNBLR model demonstrated the highest AUC or the smallest Press's Q. Using the classification test, the following results are obtained in Table 9.

**Table 9**. Comparison of AUC and Press's Q

| Methods | Accuracy | Sensitivity | Specificity | AUC | Press's Q | Chi Square |
|---|---|---|---|---|---|---|
| BLR | 73.33% | 42.85% | 89.74% | 66.30% | 13.067 | 51.829 |
| BPR | 73.33% | 42.85% | 89.74% | 66.30% | 13.067 | 51.623 |
| FSNBLR | 85% | 71.42% | 92.31% | 81.86% | 29.400 | 49.879 |

Based on Table 9, case 1 shows that the AUC value of FSNBLR (81.86%) is higher than BLR (66.30%) and BPR (66.30%). In addition, a larger Press's Q value for FSNBLR (29.400) indicates that the FSNBLR model can classify well and has a greater chance of rejecting H0 or Press's Q > Chi Square. These results are obtained using diabetes mellitus data from patients at Hajj General Hospital Surabaya. The comparison through the plot is shown in Fig. 4.

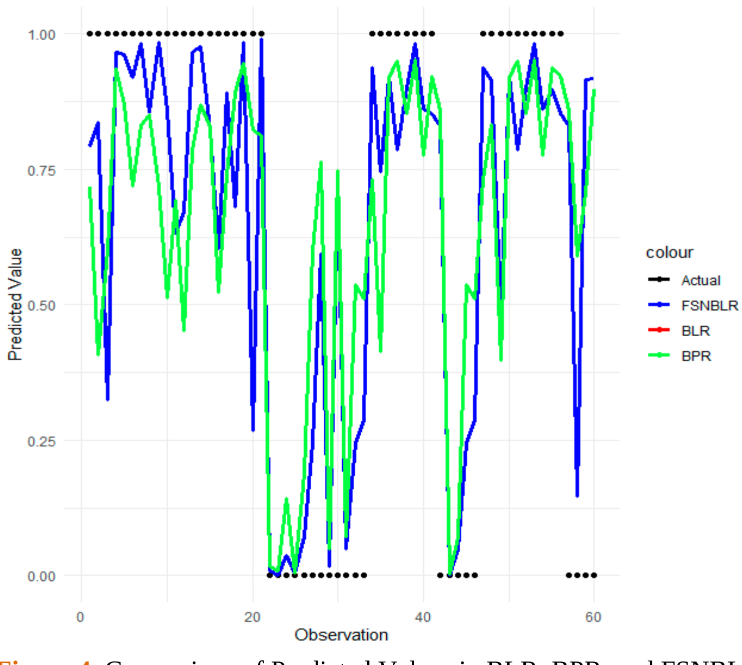


**Figure 4**. Comparison of Predicted Values in BLR, BPR, and FSNBLR

Based on Fig. 4, the plots show that the predicted value of the three methods fluctuate and do not necessarily indicate which method is better. However, for certain cases such as the one used in this article FSNBLR tends to perform better than BLR and BPR. The FSNBLR is superior as it provides odds estimates that are close to the actual values in almost all selected observation.

## 4. CONCLUSION

Based on the discussion that has been described, these findings align with the theoretical advantage of the FSNBLR model, which incorporates oscillatory components (e.g., cosine functions) to capture nonlinear and repeating patterns in the data. This makes FSNBLR especially effective for modeling categorical response variables influenced by predictors with non-monotonic or cyclical relationships—common in complex medical data such as diabetes mellitus risk factors.

Based on the data that has been used in this research, the FSNBLR is the best model for status diabetes mellitus is as follows. The FSNBLR model for categorical data is as follows:

$$\hat{\pi}(x_i) = \frac{\exp(-15.56 + 0.16x_{1i} + 0.81\cos x_{1i} + 0.22x_{2i} - 0.65\cos x_{2i} + 1.40\cos 2x_{2i} + 0.01x_{3i} - 1.40\cos x_{3i})}{1 + \exp(-15.56 + 0.16x_{1i} + 0.81\cos x_{1i} + 0.22x_{2i} - 0.65\cos x_{2i} + 1.40\cos 2x_{2i} + 0.01x_{3i} - 1.40\cos x_{3i})}$$

The deviance value for the FSNBLR model (41.837) is also lower than that of the BLR (53.007) and BPR (52.728), which further confirms that the FSNBLR model fits the data more effectively. The estimation of the FSNBLR model has a higher AUC value of 81.86%, so it can be concluded that the FSNBLR provides a better estimate than BLR and BPR. The accuracy, sensitivity, and specificity value of the FSNBLR model has a higher value than BLR and BPR. This indicates that the performance of the FSNBLR model is better.

## Author Contributions

Bambang Widjanarko Otok: Formal Analysis, Funding Acquisition, Writing - Review and Editing. Muhammad Zulfadhli: Investigation, Resources, Validation, Writing - Review and Editing; Riwi Dyah Pangesti: Methodology, Visualization, Writing - Review and Editing. Muhammad Idham Kurniawan: Software, Data Curation, Writing - Original Draft. Albertus Eka Putra Haryanto: Investigation, Project Administration, Writing – Original Draft. Darwis: Investigation, Resources, Validation. Iwan Kurniawan: Resources, Writing - Original Draft. All authors discussed the results and contributed to the final manuscript.

## Funding Statement

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

## Acknowledgment

The authors would like to thank Institut Teknologi Sepuluh Nopember, Bengkulu University, Regional Economic Development Institute (REDI), STAIN Majene, Politeknik STIA LAN Bandung for their technical support, laboratory facilities, and other non-authorial contributions that supported the completion of this research.

## Declarations

The authors declare no conflicts of interest to report study.

## REFERENCES

[1] F. E. Harrell Jr., *REGRESSION MODELING STRATEGIES: WITH APPLICATIONS TO LINEAR MODELS, LOGISTIC AND ORDINAL REGRESSION, AND SURVIVAL ANALYSIS*, 2nd ed., Springer, 2015. doi: https://doi.org/10.1198/tech.2003.s158.

[2] J. J. Faraway, *EXTENDING THE LINEAR MODEL WITH R: GENERALIZED LINEAR, MIXED EFFECTS AND NONPARAMETRIC REGRESSION MODELS*, 2nd ed., Chapman & Hall/CRC, 2016. doi: https://doi.org/10.1201/9781315382722

[3] M. M. Panja and B. N. Mandal, *WAVELET BASED APPROXIMATION SCHEMES FOR SINGULAR INTEGRAL EQUATIONS*. 2020. doi: https://doi.org/10.4324/9780429244070.

[4] Y. Farida, I. Purwanti, and N. Ulinnuha, “COMPARING GAUSSIAN AND EPANECHNIKOV KERNEL OF NONPARAMETRIC REGRESSION IN FORECASTING ISSI (INDONESIA SHARIA STOCK INDEX),” *BAREKENG: Journal of Mathematics and Applications*, vol. 16, no. 1, pp. 323-332, 2022. doi: https://doi.org/10.30598/barekengvol16iss1pp321-330.

[5] B. Pratama, A. Suryono, N. Auliyah, and N. Chamidah, “COMPARISON OF LOCAL POLYNOMIAL REGRESSION AND ARIMA IN PREDICTING THE NUMBER OF FOREIGN TOURIST VISITS TO INDONESIA,” *BAREKENG: Journal of Mathematics and Applications*, vol. 18, no. 1, pp. 53-64, 2024. doi: https://doi.org/10.30598/barekengvol18iss1pp0043-0052

[6] M. D. Cattaneo, M. Jansson, and X. Ma, “LOCAL REGRESSION DISTRIBUTION ESTIMATORS,” *Journal of Econometrics*, vol. 240, no. 2, 2024. doi: https://doi.org/10.1016/j.jeconom.2021.01.006

[7] S. D. P. Yasmirullah, B.W. Otok, J. D. T. Purnomo, and D. D. Prastyo, “PARAMETER ESTIMATION OF MULTIVARIATE ADAPTIVE REGRESSION SPLINE (MARS) WITH STEPWISE APPROACH TO MULTI DRUG-

RESISTANT TUBERCULOSIS (MDR-TB) MODELING IN LAMONGAN REGENCY," *Journal of Physics: Conference Series*, vol. 1752, IOP Publishing, 2021. doi: https://doi.org/10.1088/1742-6596/1752/1/012017

[8] S. D. P. Yasmirullah, B. W. Otok, J. D. T. Purnomo, and D. D. Prastyo, "MODIFICATION OF MULTIVARIATE ADAPTIVE REGRESSION SPLINE (MARS)," *Journal of Physics: Conference Series*, vol. 1863, IOP Publishing, 2021. doi: https://doi.org/10.1088/1742-6596/1863/1/012078.

[9] P. A. Morettin and R. F. Porto, "ESTIMATION OF NONPARAMETRIC REGRESSION MODELS BY WAVELETS," *São Paulo Journal of Mathematical Sciences*, vol. 16, pp. 539–565, 2022. doi: https://doi.org/10.1007/s40863-021-00240-5.

[10] M. D. Pasarella, Sifriyani, and F. D. T. Amijaya, "NONPARAMETRIC REGRESSION MODEL ESTIMATION WITH THE FOURIER SERIES APPROACH AND ITS APPLICATION TO THE ACCUMULATIVE COVID-19 DATA IN INDONESIA," *BAREKENG: Journal of Mathematics and Its Application*, vol. 16, no. 4, pp. 1167–1174, 2022. doi: https://doi.org/10.30598/barekengvol16iss4pp1167-1174.

[11] B. B. Jena, S. K. Paikray, and M. Mursaleen, "ON THE DEGREE OF APPROXIMATION OF FOURIER SERIES BASED ON A CERTAIN CLASS OF PRODUCT DEFERRED SUMMABILITY MEANS," *Journal of Inequalities and Applications*, vol. 2023, no. 18, 2023. doi: https://doi.org/10.1186/s13660-023-02927-z.

[12] D. S. Sheiso, "APPROXIMATION OF FUNCTIONS USING FOURIER SERIES AND ITS APPLICATION TO THE SOLUTION OF PARTIAL DIFFERENTIAL EQUATIONS," *Science Journal of Applied Mathematics and Statistics*, vol. 10, no. 4, pp. 57-84, 2022. doi: https://doi.org/10.2139/ssrn.4775015.

[13] W. Gazali, "THE ORIGIN OF THE BASIC FORMULA OF THE FOURIER SERIES," *Engineering, Mathematics and Computer Science Journal (EMACS)*, vol. 5, no. 1, 2023. doi: https://doi.org/10.21512/emacsjournal.v5i1.9398.

[14] M. L. Maslakov and V. V. Egorov, "FOR THE PROBLEM OF PHASE PROBABILITY DENSITY FUNCTION ESTIMATION," *Numerical Analysis and Applications*, vol. 15, no. 2, pp. 125–137, 2022. doi: https://doi.org/10.15372/SJNM20220205.

[15] Kuzairi, Miswanto, and M. F. F. Mardianto, "SEMIPARAMETRIC REGRESSION BASED ON FOURIER SERIES FOR LONGITUDINAL DATA WITH WEIGHTED LEAST SQUARE (WLS) OPTIMIZATION," *Journal of Physics: Conference Series*, vol. 1836, no. 1, 2021. doi: https://doi.org/10.1088/1742-6596/1836/1/012038.

[16] H. Husain, I. N. Budiantara, and I. Zain, "MIXED ESTIMATOR OF SPLINE TRUNCATED, FOURIER SERIES, AND KERNEL IN BIRESPONSE SEMIPARAMETRIC REGRESSION MODEL," *IOP Conf. Ser. Earth Environ. Sci.*, vol. 880, no. 1, 2021. doi: https://doi.org/10.1088/1755-1315/880/1/012046.

[17] I. W. Sudiarsa, I. N. Budiantara, S. Suhartono, and S. W. Purnami, "COMBINED ESTIMATOR FOURIER SERIES AND SPLINE TRUNCATED IN MULTIVARIABLE NONPARAMETRIC REGRESSION," *Appl. Math. Sci.*, vol. 9, no. 97–100, pp. 4997–5010, 2015. doi: https://doi.org/10.12988/ams.2015.55394.

[18] N. P. A. M. Mariati, I. N. Budiantara, and V. Ratnasari, "COMBINATION ESTIMATION OF SMOOTHING SPLINE AND FOURIER SERIES IN NONPARAMETRIC REGRESSION," *J. Math.*, vol. 2020, 2020. doi: https://doi.org/10.1155/2020/4712531.

[19] I. N. Budiantara *et al.*, "MODELING PERCENTAGE OF POOR PEOPLE IN INDONESIA USING KERNEL AND FOURIER SERIES MIXED ESTIMATOR IN NONPARAMETRIC REGRESSION," *Investigacion Operacional*, vol. 40, pp. 538-551, 2019.

[20] K. Nisa, I. N. Budiantara, and A. T. Rumiati, "MULTIVARIABLE SEMIPARAMETRIC REGRESSION MODEL WITH COMBINED ESTIMATOR OF FOURIER SERIES AND KERNEL," in *IOP Conference Series: Earth and Environmental Science*, IOP Publishing, 2017. doi: https://doi.org/10.1088/1755-1315/58/1/012028

[21] A. T. Ampa, I. N. Budiantara, and I. Zain, "MODELING THE LEVEL OF DRINKING WATER CLARITY IN SURABAYA CITY DRINKING WATER REGIONAL COMPANY USING COMBINED ESTIMATION OF MULTIVARIABLE FOURIER SERIES AND KERNEL," *Sustainability*, vol. 14, 2022. doi: https://doi.org/10.3390/su142013663.

[22] M. Ramli, I. N. Budiantara, and V. Ratnasari, "A METHOD FOR PARAMETER HYPOTHESIS TESTING IN NONPARAMETRIC REGRESSION WITH FOURIER SERIES APPROACH," *MethodsX*, vol. 11, 2023. doi: https://doi.org/10.1016/j.mex.2023.102468.

[23] P. H. Jou and S. H. Mirhashemi, "FREQUENCY ANALYSIS OF EXTREME DAILY RAINFALL OVER AN ARID ZONE OF IRAN USING FOURIER SERIES METHOD," *Applied Water Science*, vol. 13, p. 16, 2023. doi: https://doi.org/10.1007/s13201-022-01823-z.

[24] L. Laome, I. N. Budiantara, and V. Ratnasari, "POVERTY MODELING WITH SPLINE TRUNCATED, FOURIER SERIES, AND MIXED ESTIMATOR GEOGRAPHICALLY WEIGHTED NONPARAMETRIC REGRESSION," in *AIP Conference Proceedings*, AIP Publishing, 2024. doi: https://doi.org/10.1063/5.0206173.

[25] S. Suliyanto, M. Rifada, and E. Tjahjono, "ESTIMATION OF NONPARAMETRIC BINARY LOGISTIC REGRESSION MODEL WITH LOCAL LIKELIHOOD LOGIT ESTIMATION METHOD (CASE STUDY OF DIABETES MELLITUS PATIENTS AT SURABAYA HAJJ GENERAL HOSPITAL)," in *Symposium on Biomathematics 2019*, pp. 1551-7616, AIP Conference Proceedings, Bali, 2020. doi: https://doi.org/10.1063/5.0025807.

[26] H. Hamie, A. Hoayek, B. El-Ghoul, and M. Khalifeh, "APPLICATION OF NON-PARAMETRIC STATISTICAL METHODS TO PREDICT PUMPABILITY OF GEOPOLYMERS FOR WELL CEMENTING," *Journal of Petroleum Science and Engineering*, vol. 212, 2022. doi: https://doi.org/10.1016/j.petrol.2022.110333.

[27] T. Wang, W. Tang, Y. Lin, and W. Su, "SEMI-SUPERVISED INFERENCE FOR NONPARAMETRIC LOGISTIC REGRESSION," *Statistics in Medicine*, vol. 42, pp. 2573–2589, 2023. doi: https://doi.org/10.1002/sim.9737.

[28] M. Zulfadhli, I.N. Budiantara, and V. Ratnasari, "NONPARAMETRIC REGRESSION ESTIMATOR OF MULTIVARIABLE FOURIER SERIES FOR CATEGORICAL DATA," *MethodsX*, 2024. doi: https://doi.org/10.1016/j.mex.2024.102983.

[29] V. Ratnasari, S. H. Utama, and A. T. R. Dani, "TOWARD SUSTAINABLE DEVELOPMENT GOALS (SDGS) WITH STATISTICAL MODELING: RECURSIVE BIVARIATE BINARY PROBIT," *IAENG Int. J. Appl. Math.*, vol. 54, no. 8, pp. 1515–1521, 2024.