Available online at http://scik.org
Commun. Math. Biol. Neurosci. 2021, 2021:28
https://doi.org/10.28919/cmbn/5353
ISSN: 2052-2541

ESTIMATION OF NONPARAMETRIC ORDINAL LOGISTIC REGRESSION MODEL USING LOCAL MAXIMUM LIKELIHOOD ESTIMATION

MARISA RIFADA[1,3], NUR CHAMIDAH[1,3,*], VITA RATNASARI[2], PURHADI[2]

[1]Department of Mathematics, Faculty of Science and Technology, Airlangga University, Surabaya, Indonesia

[2]Department of Statistics, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia

[3]Research Group of Statistical Modeling in Life Science, Fac. Sci. & Tech., Airlangga University, Indonesia

Abstract: Ordinal logistic regression is a statistical method used to analyze the ordinal response variable with three or more categories and predictor variables that are categorical or continuous. The parametric models for ordinal response variable assume that the predictor is given by a linear form of covariates. In this study, the parametric models are extended to include smooth components based on nonparametric approach. The covariates are modeled as unspecified but smooth functions. Estimation is based on local maximum likelihood estimation (LMLE).

Keywords: ordinal logistic; nonparametric; local maximum likelihood estimation.

2010 AMS Subject Classification: 62G05, 62G08, 65D15

1. INTRODUCTION

The logistic regression model is used to analyze the relationship between response variables that are categorical and predictor variables that are categorical or continuous. If the response variable consists of two categories, then it is called a binary logistic regression model. While if the response variable is divided into more than two categories with ordinal scale then it is called the

[*]Corresponding author
E-mail address: nur-c@fst.unair.ac.id
Received December 26, 2020

ordinal logistic regression model [1]. There are two ways to approach the regression model, i.e. parametric and nonparametric approach. While semiparametric is a combination both of them.

Several studies on logistic regression based on parametric approach have been discussed by [2-3] who examined parameter estimation using the maximum likelihood estimation (MLE) method, numerical approaches for parameter estimation and diagnosis in logistic regression models. Researchers [3] studied parameter estimation, hypothesis testing and selection of the best model in the logistic regression model. Next, computational algorithms for multinomial logistic regression modeling was examined by [4].

The researches that studied logistic regression using semiparametric approach have been carried out by [5] who examined parameter estimation using the penalized likelihood estimation method in the semiparametric multinomial logit model, and by [6] who discussed semiparametric ordinal logistic model. Parameter estimation based on kernel estimator in semiparametric multinomial logit model was examined by [7].

Furthermore, researches on logistic regression by using nonparametric approach have been carried out by [8] who studied parameter estimation using the local likelihood estimation (LLE) method based on local scoring algorithm, by [9] who examined parameters estimation of the multinomial logistic regression based on kernel estimators, and by [10-12] who studied the binary logistic regression based on local likelihood logit estimation (LLLE). Some researchers studied nonparametric binary logistic regression based on generalized additive model (GAM), for examples, by using local linear estimator [13-14], by using penalized spline estimator [15-17] and by using local polynomial estimator [18]. In this study, we develop local maximum likelihood estimation (LMLE) method for estimating nonparametric ordinal logistic regression model approach and provide algorithm for estimating the parameters.

2. PRELIMINARIES

The nonparametric ordinal logistic regression model is an extension of the ordinal logistic regression model with a nonparametric approach. According to [8], the general form of the nonparametric ordinal logistic regression model is

(1) $$\text{logit}\left[\Pr\left(Y \le g\,|\mathbf{x}\right)\right] = \alpha_g + \sum_{j=1}^{p} f_j(x_j) \quad , g = 1, ..., G\text{-}1$$

where g is the category of response variable, p is the number of predictor variables, $\mathbf{x}=\begin{bmatrix}1 & x_1 & x_2 & \dots & x_p\end{bmatrix}^T$ is a vector of predictor variables, α is the intercept parameter, and $f_j(\cdot)$ is a regression function whose unknown form of the *k*-th predictor variable which will be estimated by nonparametric approach.

According to [19], if x_j is around x_{0j} then the function $f_j(\cdot)$ can be approximated by Taylor expansion for d = 1 which can be expressed as follows:

(2) $$f_j\left(x_j\right) \approx f_j\left(x_{0j}\right)+\left(x-x_{0j}\right) f_j'\left(x_{0j}\right) \equiv \beta_0\left(x_{0j}\right)+\beta_1\left(x_{0j}\right)\left(x_j-x_{0j}\right)$$

Based on equation (2) then equation (1) can be written as

(3) $$\operatorname{logit}\left[P\left(Y \le g \mid \mathbf{x}\right)\right]=\alpha_g+\tilde{\mathbf{x}}^T\left(x_0\right)\boldsymbol{\beta}\left(x_0\right)$$

where $\boldsymbol{\beta}\left(x_0\right)=\begin{bmatrix}\beta_0 & \beta_1\left(x_{01}\right) & \beta_2\left(x_{02}\right) & \dots & \beta_p\left(x_{0p}\right)\end{bmatrix}^T$

$$\tilde{\mathbf{x}}\left(x_0\right)=\begin{bmatrix}1 & x_1-x_{01} & x_2-x_{02} & \dots & x_p-x_{0p}\end{bmatrix}^T$$

Based on equations (1) and (3), the cumulative probability of the g-th response category can be expressed as

(4) $$P\left(Y \le g \mid \mathbf{x}\right)=\frac{\exp\left(\alpha_g+\tilde{\mathbf{x}}^T\left(x_0\right)\boldsymbol{\beta}\left(x_0\right)\right)}{1+\exp\left(\alpha_g+\tilde{\mathbf{x}}^T\left(x_0\right)\boldsymbol{\beta}\left(x_0\right)\right)},\ g=1,2,\dots,G-1$$

Suppose $\pi_g\left(\mathbf{x}\right)=P\left(Y=g \mid \mathbf{x}\right)$ that the probability of the response variable has the g-th category, then we get

(5) $$\pi_g\left(\mathbf{x}\right)=\frac{\exp\left(\alpha_g+\tilde{\mathbf{x}}^T\left(x_0\right)\boldsymbol{\beta}\left(x_0\right)\right)}{1+\exp\left(\alpha_g+\tilde{\mathbf{x}}^T\left(x_0\right)\boldsymbol{\beta}\left(x_0\right)\right)}-\frac{\exp\left(\alpha_{g-1}+\tilde{\mathbf{x}}^T\left(x_0\right)\boldsymbol{\beta}\left(x_0\right)\right)}{1+\exp\left(\alpha_{g-1}+\tilde{\mathbf{x}}^T\left(x_0\right)\boldsymbol{\beta}\left(x_0\right)\right)}$$

where $\frac{\exp\left(\alpha_0+\tilde{\mathbf{x}}^T\left(x_0\right)\boldsymbol{\beta}\left(x_0\right)\right)}{1+\exp\left(\alpha_0+\tilde{\mathbf{x}}^T\left(x_0\right)\boldsymbol{\beta}\left(x_0\right)\right)}=0$ and $\frac{\exp\left(\alpha_G+\tilde{\mathbf{x}}^T\left(x_0\right)\boldsymbol{\beta}\left(x_0\right)\right)}{1+\exp\left(\alpha_G+\tilde{\mathbf{x}}^T\left(x_0\right)\boldsymbol{\beta}\left(x_0\right)\right)}=1$.

The probability values for each response category are used as a guideline for classification. An observation will fall into the g-th category response based on the greatest probability value.

3. MAIN RESULTS

In this section we describe the main results of the research including estimating parameters of model and algorithm for estimating the parameters.

3.1 Estimating Parameters of Model

The parameter estimation of the nonparametric ordinal logistic regression model is performed by using LMLE method. The initial step of estimating this parameter starts with taking *n* random samples that are mutually independent. For example $\left(x_{i1}, x_{i2},, ..., x_{ip}, y_i\right)$ is data paired with $\left(x_{i1}, x_{i2},, ..., x_{ip}\right)$ is data from as many as p predictor variables and y_i is data from categorical response variables with ordinal scale consisting of G categories with the probability of the outcome in the g-category is $\pi_g\left(\mathbf{x}_i\right)$ then $\boldsymbol{Y}_i = \left(y_{i1}, y_{i2}, ..., y_{i,G-1}\right) \sim \text{Multinomial}\left(1; \pi_1\left(\mathbf{x}_i\right), \pi_2\left(\mathbf{x}_i\right), ..., \pi_{G-1}\left(\mathbf{x}_i\right)\right)$, then form the local likelihood function is [10]:

$$L\left(\boldsymbol{\theta}_i\right) = \prod_{i=1}^{n} f\left(y_i\right)^{\prod_{j=1}^{p} K_{h_j}\left(x_{ij} - x_{0j}\right)}$$

(6) $$= \prod_{i=1}^{n} \left\{ \prod_{g=1}^{G} \left(\frac{\exp\left(\alpha_g + \tilde{\mathbf{x}}_i^T\left(x_0\right)\boldsymbol{\beta}\left(x_0\right)\right)}{1 + \exp\left(\alpha_g + \tilde{\mathbf{x}}_i^T\left(x_0\right)\boldsymbol{\beta}\left(x_0\right)\right)} - \frac{\exp\left(\alpha_{g-1} + \tilde{\mathbf{x}}_i^T\left(x_0\right)\boldsymbol{\beta}\left(x_0\right)\right)}{1 + \exp\left(\alpha_{g-1} + \tilde{\mathbf{x}}_i^T\left(x_0\right)\boldsymbol{\beta}\left(x_0\right)\right)} \right)^{y_{ig}} \right\}^{\prod_{j=1}^{p} K_{h_j}\left(x_{ij} - x_{0j}\right)}$$

where $\boldsymbol{\theta}_i = \left[\alpha_{1i} \quad \alpha_{2i} \quad \cdots \quad \alpha_{G-1,i} \quad \boldsymbol{\beta}_i\left(x_0\right)\right]^T$ is the local parameter of the nonparametric ordinal logistic regression model in the *i*-th observation and $K_h\left(u\right)$ is the Kernel weighting function. According to [20], Kernel functions are defined as

(7) $$K_h\left(u\right) = \frac{1}{h} K\left(\frac{u}{h}\right) \text{ for } -\infty < u < \infty \text{ , } h > 0$$

where h is the optimum bandwidth value.

Based on the LMLE method, the parameter estimation of the nonparametric ordinal logistic regression model can be obtained by maximizing the local likelihood function. To simplify the calculation, the ln-transformation is carried out on the local likelihood function so that the ln-local likelihood function is formed as follows:

(8) $$\ell\left(\boldsymbol{\theta}_i\right) = \sum_{i=1}^{n} \left[\prod_{j=1}^{p} K_{h_j}\left(x_{ij} - x_{0j}\right) \sum_{g=1}^{G} \left\{ y_{ig} \ln\left(\frac{\exp\left(\alpha_g + \tilde{\mathbf{x}}_i^T\left(x_0\right)\boldsymbol{\beta}\left(x_0\right)\right)}{1 + \exp\left(\alpha_g + \tilde{\mathbf{x}}_i^T\left(x_0\right)\boldsymbol{\beta}\left(x_0\right)\right)} - \frac{\exp\left(\alpha_{g-1} + \tilde{\mathbf{x}}_i^T\left(x_0\right)\boldsymbol{\beta}\left(x_0\right)\right)}{1 + \exp\left(\alpha_{g-1} + \tilde{\mathbf{x}}_i^T\left(x_0\right)\boldsymbol{\beta}\left(x_0\right)\right)} \right) \right\} \right]$$

In this study, we estimate the parameters for a model that has a response variable consisting of 3 categories. Therefore, the ln-local likelihood function that is formed is as follows:

(9) $$\ell(\boldsymbol{\theta}_i)=\sum_{i=1}^{n}\left[\prod_{j=1}^{p}K_{h_j}\left(x_{ij}-x_{0j}\right)\left\{y_{i1}\left(\alpha_1+\tilde{\mathbf{x}}_i^T\left(x_0\right)\boldsymbol{\beta}\left(x_0\right)\right)-\left(y_{i1}+y_{i2}\right)\ln\left(1+\exp\left(\alpha_1+\tilde{\mathbf{x}}_i^T\left(x_0\right)\boldsymbol{\beta}\left(x_0\right)\right)\right)\right.\right.$$

$$\left.\left.+y_{i2}\ln\left(\exp\left(\alpha_2+\tilde{\mathbf{x}}_i^T\left(x_0\right)\boldsymbol{\beta}\left(x_0\right)\right)-\exp\left(\alpha_1+\tilde{\mathbf{x}}_i^T\left(x_0\right)\boldsymbol{\beta}\left(x_0\right)\right)\right)+\left(y_{i1}-1\right)\ln\left(1+\exp\left(\alpha_2+\tilde{\mathbf{x}}_i^T\left(x_0\right)\boldsymbol{\beta}\left(x_0\right)\right)\right)\right\}\right]$$

Parameter estimation is done by performing the first partial derivative of equation (9) on the parameter to be estimated and then equalized to zero, therefore we obtain

(10) $$\frac{\partial\ell(\boldsymbol{\theta}_i)}{\partial\alpha_1}=\sum_{i=1}^{n}\left[\prod_{j=1}^{p}K_{h_j}\left(x_{ij}-x_{0j}\right)\left\{y_{i1}-\left(y_{i1}+y_{i2}\right)\frac{e_1}{1+e_1}-y_{i2}\frac{e_1}{e_2-e_1}\right\}\right]=0$$

(11) $$\frac{\partial\ell(\boldsymbol{\theta}_i)}{\partial\alpha_2}=\sum_{i=1}^{n}\left[\prod_{j=1}^{p}K_{h_j}\left(x_{ij}-x_{0j}\right)\left\{y_{i2}\frac{e_2}{e_2-e_1}+\left(y_{i1}-1\right)\frac{e_2}{1+e_2}\right\}\right]=0$$

(12) $$\frac{\partial\ell(\boldsymbol{\theta}_i)}{\partial\boldsymbol{\beta}(x_0)}=\sum_{i=1}^{n}\left[\prod_{j=1}^{p}K_{h_j}\left(x_{ij}-x_{0j}\right)\left\{\left(y_{i1}+y_{i2}\right)\tilde{\mathbf{x}}_i^T\left(x_0\right)\left(\frac{1}{1+e_1}\right)+\left(y_{i1}-1\right)\tilde{\mathbf{x}}_i^T\left(x_0\right)\left(\frac{e_2}{1+e_2}\right)\right\}\right]=0$$

where $e_1=\exp\left(\alpha_1+\tilde{\mathbf{x}}_i^T\left(x_0\right)\boldsymbol{\beta}\left(x_0\right)\right)$ and $e_2=\exp\left(\alpha_2+\tilde{\mathbf{x}}_i^T\left(x_0\right)\boldsymbol{\beta}\left(x_0\right)\right)$.

The result of the first partial derivatives are not closed-form, therefore a numerical optimization such as the Newton-Raphson method is needed to obtain the local-maximum likelihood estimators. The equation used in the Newton-Raphson iteration process is [21]:

(13) $$\boldsymbol{\theta}_i^{(t+1)}=\boldsymbol{\theta}_i^{(t)}-\left[\boldsymbol{H}\left(\boldsymbol{\theta}_i^{(t)}\right)\right]^{-1}\boldsymbol{q}\left(\boldsymbol{\theta}_i^{(t)}\right)$$

where t is the number of iterations (t = 0,1,2,…), $\boldsymbol{q}\left(\boldsymbol{\theta}_i\right)$ is the gradient vector with the vector elements being the first partial derivative of the ln-local likelihood function against the parameter to be estimated and $\boldsymbol{H}\left(\boldsymbol{\theta}_i\right)$ is a Hessian matrix with the elements of the matrix being the second partial derivative of the ln-local likelihood function for the parameter to be estimated, hence required the second partial derivative of the ln-local likelihood function against the parameter to be estimated. The results of the second partial derivative obtained are as follows:

(14) $$\frac{\partial^2\ell(\boldsymbol{\theta}_i)}{\partial\alpha_1^2}=\sum_{i=1}^{n}\left[\prod_{j=1}^{p}K_{h_j}\left(x_{ij}-x_{0j}\right)\left\{-\left(y_{i1}+y_{i2}\right)\frac{e_1}{\left[1+e_1\right]^2}-y_{i2}\frac{e_1\ e_2}{\left[e_2-e_1\right]^2}\right\}\right]$$

(15) $$\frac{\partial^2 \ell(\boldsymbol{\theta}_i)}{\partial\alpha_1 \partial\alpha_2} = \sum_{i=1}^{n}\left[\prod_{j=1}^{p} K_{h_j}\left(x_{ij}-x_{0j}\right)\left\{y_{i2}\frac{e_1\, e_2}{\left[e_2-e_1\right]^2}\right\}\right]$$

(16) $$\frac{\partial^2 \ell(\boldsymbol{\theta}_i)}{\partial\alpha_1 \partial\boldsymbol{\beta}^T(x_0)} = \sum_{i=1}^{n}\left[\prod_{j=1}^{p} K_{h_j}\left(x_{ij}-x_{0j}\right)\left\{-\left(y_{i1}+y_{i2}\right)\tilde{\mathbf{x}}_i^T(x_0)\frac{e_1}{\left[1+e_1\right]^2}\right\}\right]$$

(17) $$\frac{\partial^2 \ell(\boldsymbol{\theta}_i)}{\partial\alpha_2^2} = \sum_{i=1}^{n}\left[\prod_{j=1}^{p} K_{h_j}\left(x_{ij}-x_{0j}\right)\left\{-y_{i2}\frac{e_1\, e_2}{\left[e_2-e_1\right]^2}+\left(y_{i1}-1\right)\frac{e_2}{\left[1+e_2\right]^2}\right\}\right]$$

(18) $$\frac{\partial^2 \ell(\boldsymbol{\theta}_i)}{\partial\alpha_2 \partial\boldsymbol{\beta}^T(x_0)} = \sum_{i=1}^{n}\left[\prod_{j=1}^{p} K_{h_j}\left(x_{ij}-x_{0j}\right)\left\{\left(y_{i1}-1\right)\tilde{\mathbf{x}}_i^T(x_0)\frac{e_2}{\left[1+e_2\right]^2}\right\}\right]$$

(19) $$\frac{\partial^2 \ell(\boldsymbol{\theta}_i)}{\partial\boldsymbol{\beta}(x_0) \partial\boldsymbol{\beta}^T(x_0)} = \sum_{i=1}^{n}\left[\prod_{j=1}^{p} K_{h_j}\left(x_{ij}-x_{0j}\right)\left\{-\left(y_{i1}+y_{i2}\right)\tilde{\mathbf{x}}_i^T(x_0)\tilde{\mathbf{x}}_i(x_0)\frac{e_1}{\left[1+e_1\right]^2}\right.\right.$$
$$\left.\left.+\left(y_{i1}-1\right)\tilde{\mathbf{x}}_i^T(x_0)\tilde{\mathbf{x}}_i(x_0)\frac{e_2}{\left[1+e_2\right]^2}\right\}\right]$$

The Newton-Raphson iteration process will stop if the convergent conditions, where $\left\|\boldsymbol{\theta}^{(t+1)}-\boldsymbol{\theta}^{(t)}\right\| \leq \varepsilon$, which ε is a very small number. The estimation results obtained are $\boldsymbol{\theta}^{(t+1)}$ at the time of the last iteration. This iteration procedure is repeated for every *i*-th observation, therefore a local parameter estimator will be obtained from a nonparametric ordinal logistic regression model based on a linear local estimator.

3.2 Algorithm for Estimating Parameters

The algorithm for estimating parameter of nonparametric ordinal logistic regression model using the local maximum likelihood estimation (LMLE) method are as follows:

1. Determining the indicator variable of Y, i.e

   $y_{ig}=1$, if y_i has a *g*-th category $\left(i=1,2,\ldots,n\,;\,g=1,2,3\right)$

   $y_{ig}=0$, for others

2. Determining the initial value of bandwidth (*h*) for each predictor variable

3. Estimating $\boldsymbol{\theta}_i$ (for $i=1,2,...,n$) using the Newton-Raphson algorithm as follows:
   a. Determining the initial value of the parameter estimator $\boldsymbol{\theta}_i^{(0)}$ using a parametric ordinal logistic regression model
   b. Forming the gradient vector $\boldsymbol{q}\left(\boldsymbol{\theta}_i^{(t)}\right)$ and the Hessian matrix $\boldsymbol{H}\left(\boldsymbol{\theta}_i^{(t)}\right)$
   c. Calculating

   $$\boldsymbol{\theta}_i^{(t+1)} = \boldsymbol{\theta}_i^{(t)} - \left[\boldsymbol{H}\left(\boldsymbol{\theta}_i^{(t)}\right)\right]^{-1} \boldsymbol{q}\left(\boldsymbol{\theta}_i^{(t)}\right)$$

   d. If value of $\left\|\boldsymbol{\theta}_i^{(t+1)} - \boldsymbol{\theta}_i^{(t)}\right\| \le \varepsilon$ for $\varepsilon = 10^{-6}$, then go to step (e), but if it has not converged then return to step (b) for the iteration to-$t = t+1$
   e. Get an estimator $\hat{\boldsymbol{\theta}}_i$
4. Calculating the value of $\hat{\pi}_{i,g}$
5. Repeating step (3) to (4) for $i = 2,...,n$
6. Calculating $CV(h)$ value with the formula:

   $$CV(h) = \sum_{i=1}^{n}\sum_{g=1}^{3}\left(y_{i,g} - \hat{\pi}_{i,g}(h)\right)^2$$

7. Repeating step (3) to (6) for different bandwidth values (*h*) so that the minimum value of $CV(h)$ is obtained
8. Get the optimal bandwidth value (h_{opt}) based on the value of h that minimizes $CV(h)$ value
9. Estimating the parameter vector $\boldsymbol{\theta}_i$ using the Newton-Raphson iteration based on the optimal bandwidth value (h_{opt}), for $i = 2,...,n$
10. Calculating the classification accuracy of the nonparametric ordinal logistic regression model.

4. CONCLUSIONS

We developed a local regression model, called a nonparametric ordinal logistic regression model by using local linear estimator. We have shown the step-by-step procedure to obtain the estimation of the parameters based on LMLE and the algorithm for estimating parameter of nonparametric ordinal logistic regression model. It is easy to apply the same procedure presented in this article for developing other models based on different distributions or the other kinds of estimators.

ACKNOWLEDGMENTS

We wish to thank Universitas Airlangga for funding this research with a contract number 824/UN3.14/PT/2020. We also thank to editors and peer reviewers who have given useful suggestions and criticisms for the improvement of this paper.

CONFLICT OF INTERESTS

The authors declare that there is no conflict of interests.

REFERENCES

[1] D.W. Hosmer, S. Lemeshow, Applied Logistic Regression, 2nd ed. John Wiley & Sons, New York, (2000), 260-288.

[2] D. Pregibon, Logistic Regression Diagnostics. Ann. Stat. 9 (1981), 705-724.

[3] P. Andriani, N. Chamidah, Modelling of Hypertension Risk Factors Using Logistic Regression to Prevent Hypertension in Indonesia, J. Phys.: Conf. Ser. 1306 (2019) 012027.

[3] D.W. Hosmer, B. Jovanovic, S. Lemeshow, Best Subsets Logistic Regression, Biometrics. 45 (1989), 1265-1270.

[4] D. Bohning, Multinomial Logistic Regression Algorithm. Ann. Inst. Stat. Math. 44 (1992),197-200.

[5] G. Tutz, T. Scholz, Semiparametric Modelling of Multicategorical Data. J. Stat. Comput. Simul. 74 (2004), 183-200.

[6] G. Tutz, Generalized Semiparametrically Structured Ordinal Models. Biometrics, 59 (2003), 263-273.

[7] R. Langrock, N.-B. Heidenreich, S. Sperlich, Kernel-based semiparametric multinomial logit modelling of political party preferences, Stat. Methods Appl. 23 (2014) 435–449.

[8] T. Hastie, R. Tibshirani, Non-Parametric Logistic and Proportional Odds Regression, Appl. Stat. 36 (1987), 260-276.

[9] H. Okumura, K. Naito, Non-parametric Kernel Regression for Multinomial Data, J. Multivar. Anal. 97 (2006), 2009-2022.

[10] M. Frölich, Non-parametric regression for binary dependent variables, Econometrics J. 9 (2006), 511–540.

[11] M. Rifada, Suliyanto, E. Tjahjono, A. Kesumawati, The Logistic Regression Analysis with Nonparametric Approach based on Local Scoring Algorithm (Case Study: Diabetes Mellitus Type II Cases in Surabaya of Indonesia), Int. J. Adv. Soft Comput. Appl. 10 (2018), 167-178.

[12] Suliyanto, M. Rifada, 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), AIP Conf. Proc. 2264 (2020), 030007.

[13] E. Ana, N. Chamidah, P. Andriani, B. Lestari, Modeling of hypertension risk factors using local linear of additive nonparametric logistic regression, J. Phys.: Conf. Ser. 1397 (2019), 012067.

[14] A. Puspitawati, N. Chamidah, Choroidal Neovascularisation Classification on Fundus Retinal Images Using Local Linear Estimator, IOP Conf. Ser.: Mater. Sci. Eng. 546 (2019), 052056.

[15] T. Adiwati, N. Chamidah, Modelling of Hypertension Risk Factors Using Penalized Spline to Prevent Hypertension in Indonesia, IOP Conf. Ser.: Mater. Sci. Eng. 546 (2019), 052003.

[16] A.E. Anwar, N. Chamidah, Glaucoma Identification on Fundus Retinal Images Using Statistical Modelling Approach, IOP Conf. Ser.: Mater. Sci. Eng. 546 (2019), 052010.

[17] N. Rizka, N. Chamidah, Lung Tumor Classification on Human Chest X-Ray Using Statistical Modelling Approach, IOP Conf. Ser.: Mater. Sci. Eng. 546 (2019), 052065.

[18] N. Chamidah, K. H. Gusti, E. Tjahjono, B. Lestari, Improving of Classification Accuracy of Cyst and Tumor Using Local Polynomial Estimator, TELKOMNIKA, 17(3) (2019),1492-1500.

[19] J. Fan, I. Gijbels, Local Polynomial Modelling and Its Applications, 1st ed. Chapman & Hall, London, (1996).

[20] W. Hardle, Applied Nonparametric Regression, Cambridge University Press, New York, (1990).

[21] A. Agresti, Categorical Data Analysis, John Wiley & Sons, New Jersey, (2002).