JTAM (Jurnal Teori dan Aplikasi Matematika)
http://journal.ummat.ac.id/index.php/jtam
p-ISSN 2597-7512 | e-ISSN 2614-1175
Vol. 4, No. 1, April 2020, pp. 39-48

Comparison of Logistic Regression Model and MARS Using Multicollinearity Data Simulation

Ananto Wibowo[1], M. Rismawan Ridha[2]
[1]Badan Pusat Statistik Kabupaten Cianjur, Indonesia
[2]Badan Pusat Statistik Kabupaten Maluku Tengah, Indonesia
[1]ananto.wibowo@bps.go.id, [2]rismawan.ridha@bps.go.id

Article History:

Received : 12-02-2020
Revised : 23-03-2020
Accepted : 25-03-2020
Online : 02-04-2020

Keyword:

Logistic Regression; Multivariate Adaptive Regression Splines; Principal Component Analysis.

ABSTRACT

There are several statistical methods used to model the effect of predictor variables on categorical response variables, namely logistic regression and Multivariate Adaptive Regression Splines (MARS). However, neither MARS nor logistic regression allows multicollinearity on any predictor variables. This study applies the use of both methods to the simulation data with principal component analysis as an improvement in multicollinearity to find out which regression has better performance. The result of the analysis shows that MARS is very powerful in modeling research simulation data. Besides, based on the criteria of the number of significant major components, accuracy, sensitivity, and specificity values, MARS has more appropriate performance than logistic regression.

https://doi.org/10.31764/jtam.v4i1.1801

A. INTRODUCTION

There are several statistical methods used to model the effect of predictor variables on categorical response variables, namely logistic regression and Multivariate Adaptive Regression Splines (MARS). Both of these methods have significant differences where logistic regression is a parametric approach with the estimated parameter used is the Maximum Likelihood Estimator (Diop et al., 2011), while the MARS method is a non-parametric statistical method based on partitioning data sets into several separate slopes (splines) with different gradients (Zhang and Goh, 2016) .

Neither MARS nor logistic regression allows the multicollinearity of each predictor variable (Adityawardhani et al., 2017). Multicollinearity is a condition in which predictor variables have strong correlations in the model. This state is an unusual phenomenon for logistic regression because there will be many covariates. Multicollinearity also causes parameter estimation to be unstable because of its inaccurate variance that affects the confidence interval in the hypothesis test, and the conclusion drawn will result in misleading information (Midi, 2013).

In line with logistic regression, if the predictor variables in the MARS method are correlated, MARS could make biased estimations in the case of multicollinearities in the model (Kayri, 2010). Therefore, the predictor variables that have multicollinearity problems must address before proceeding with MARS modeling and logistic regression. One appropriate method for handling multicollinearity cases is Principal Component Analysis (PCA) (Gwelo, 2019). The use of PCA will generate new variables, which are a linear combination of the independent variables, and the origin of this new intra-variable is independent. The new variables are called principal components that then regressed to the dependent variable (Rahayu et al., 2017) .

There are many cases of previous studies, whether comparing or using only one of the two methods (logistic regression and MARS) without evaluating the data set that is possible for multicollinearities, such as conducted by Mina & Barrios (2010), Kılınç et al., (2017), and Zewude et al., (2016). These will make the conclusions in the research considerably mislead caused by ignoring the checking procedure and the improvement of the presence of multicollinearity.

Therefore, the purpose of this study is to compare the use of the MARS method and logistic regression in simulation data with the Principal Components Analysis as a multicollinearity improvement to find out which regression has better performance. A comparison of the two methods measures by the number of significant components, accuracy, sensitivity, and specificity values.

B. METHODS

The MARS model is a nonlinear and nonparametric regression method that combines classic linear regression, the mathematical construction of splines, binary recursive partitioning, and brute and intelligent algorithms, in which no assumption is made regarding the functional relationship between the dependent and predictor variables (Felicísimo et al., 2013) . The MARS model predicts a function using linear combinations and interactions of the adaptive piecewise linear regression known as the "basic function (BF)" (Park et al., 2017). The MARS model can be used for both dependent and binary dependent variables (y). The MARS model estimator used for binary responses can be written as follows:

$$\ln\left[\frac{\pi(x)}{1-\pi(x)}\right]=a_0+\sum_{m=1}^{M} a_m \prod_{k=1}^{K_m}\left[s_{km}\left(x_{v(k,m)}-t_{km}\right)\right] \quad (1)$$

Where a_0 is a constant and $\pi(x)$ is the probability of success for an observation. The MARS method uses the stepwise (forward and backward) algorithm in selecting a model with a minimum Generalized Cross-Validation (GCV) value (Koc and Bozdogan, 2015). The GCV formula is expressed in the following form:

$$GCV(M)=\frac{\frac{1}{N}\sum_{i=1}^{N}\left[y_i-\hat{f}_M(x_i)\right]^2}{\left[1-\frac{\tilde{C}(M)}{N}\right]^2} \quad (2)$$

The procedure to get the best MARS model is done by a process of trial and error of a combination of the number of Basic Function (BF), Maximum Interaction (MI) and Minimum Observation (MO). Wahyuningrum (2009) uses the Basic Function two to four times of its

predictor variables. Whereas Maximum Interaction is one to three times of its predictor variables because it will be more difficult to model interpretation when there is not sufficient Maximum Interaction. The Minimum Observation used in this study is 0 and 1. Another method is the logistic regression model describing the response variable in the form of categorical with the predictor variable either categorical or continuous (Hosmer et al., 2013). Logistic regression used in this study is a response variable in the form of two categories or often referred to as binary logistic regression (Adityawardhani et al, 2017). (Hosmer et al., 2013) categorizes the response variable $Y = 1$ as the success category and $Y = 0$ for the failure category. The $E(Y|x)$ can be denoted as $\pi(x)$ as the chance of a successful event occurring at the value of the predictor variable x.

$$\pi(x) = \frac{e^{(\beta_0+\beta_1 x_1+\ldots+\beta_p x_p)}}{1+e^{(\beta_0+\beta_1 x_1+\ldots+\beta_p x_p)}} \tag{3}$$

Where β_p is the parameter value to $-j$, where j = 0.1,.., p, and p is the number of variables. To avoid the problem of multicollinearity in logistic regression and MARS, principal component analysis is used as an appropriate method. Johnson and Wichern (2013) state that principal component analysis explains the structure of variance-covariance through several linear combinations of a set of variables. The linear combination resulting from the analysis of main components has a maximum variance value without significantly reducing the characteristics of the data.

Geometrically, principal component analysis transforms the linear data to form a new coordinate system, denoted as Y where $X_1, X_2, \ldots X_p$ as the axis. This new coordinate is the direction with maximum variability and provides simpler covariance.

Let the random vector $\boldsymbol{X}' = [X_1, X_2, \ldots X_p]$ have the covariance matrix $\boldsymbol{\Sigma}$ with eigenvalues $\lambda_1 \geq \lambda_2 \ldots \geq \lambda_p \geq 0$. Consider the combinations:

$$\begin{aligned} Y_1 &= \boldsymbol{a'_1 X} = a_{11}X_1 + a_{12}X_2 + \cdots + a_{1p}X_p \\ Y_2 &= \boldsymbol{a'_2 X} = a_{21}X_1 + a_{22}X_2 + \cdots + a_{2p}X_p \\ &\vdots \\ Y_p &= \boldsymbol{a'_p X} = a_{p1}X_1 + a_{p2}X_2 + \cdots + a_{pp}X_p \end{aligned}$$

We then obtain,

$$Total\ variance = \lambda_1 + \lambda_2 + \cdots + \lambda_p \tag{4}$$

The principal components are those uncorrelated linear combination$Y_1, Y_2, \ldots, Y_p$ Whose total variance as large as possible. And consequently, the proportion of total variance due to (explained by) the principal component is:

$$The\ proportion\ of\ total\ variance\ due\ to\ k-th\ principal\ component = \frac{\lambda_k}{\lambda_1+\lambda_2+\cdots+\lambda_p} \tag{5}$$

Where $k = 1,2,3, \ldots, p$. The value of this new variable Y will be used for new data as a comparison of uncorrelated data between the MARS method and logistic regression. The source of research data for the predictor variable x is 25 observed data collected from the

website https://academic.uprm.edu (University de Puerto Rico Mayaguez), where the interrelated predictor variables combined with the response variable y in the form of simulation data. These data used as one representative that allows the existence of multicollinearity to meet the research goals. The correlation test is done by Pearson correlation test (Mukaka, 2012) with the formula:

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})\sum_{i=1}^{n}(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}} \quad (6)$$

Where x_i and y_i are the values of x and y for the i individual. Moreover, the comparison of the MARS method and logistic regression will be evaluated by the number of significant variables in each method, the value of accuracy, sensitivity, and specificity. Accuracy is a measure that shows the closeness of the results to its true value. Sensitivity is a percentage of occurrences correctly predicted. In health science, sensitivity demonstrates the ability of a test to correctly identify a sick person in all who suffer. Specificity is a percentage of nonoccurrences correctly predicted (Osetljivosti et al, 2014) .

Table 1. The Classification Table

No	Actual Class	Predicted Class	
		y_1	y_2
1	y_1	y_{11}	y_{12}
2	y_2	y_{21}	y_{22}

Based on Table 1, each formula can be written as follows:

$$Accuracy = \frac{y_{11+}y_{22}}{y_{11} + y_{12} + y_{21} + y_{22}} x100\% \quad (7)$$

$$Sensitivity = \frac{y_{11}}{y_{11} + y_{12}} x100\% \quad (8)$$

$$Specificity = \frac{y_{22}}{y_{21} + y_{22}} x100\% \quad (9)$$

This study uses the SPSS 20 program to calculate the Pearson correlation, and Logistic Regression, MARS 2.0 for the use of the MARS methods and MINITAB 16 to form the principal component. The researcher's use of different software is quite reasonable because there is currently no program package that accommodates the MARS method, logistic regression, and principal component analysis in one single package.

C. RESULT AND DISCUSSION

As a first step, multicollinearity detection is performed on the simulation data between each predictor variable X_1, $X_2, \ldots X_5$. Schober et al (2018) state that multicollinearity will occur if there is a strong correlation between two variables with a minimum value of 0.7. The output results in Table 2 show that there are strong correlations for all variables except X_2 with X_4 with a correlation coefficient of 0.695. Therefore, it is necessary to analyze the main components before modeling is done to compare the MARS method and logistic regression in Table 2 below.

Table 2. Coefficient of correlation among predictor variables

No	Correlation	X_1	X_2	X_3	X_4	X_5
1	X_1	1	0.806	0.754	0.733	0.758
2	X_2	0.806	1	0.774	0.695	0.715
3	X_3	0.754	0.774	1	0.841	0.838
4	X_4	0.733	0.695	0.841	1	0.785
5	X_5	0.758	0.715	0.838	0.785	1

Furthermore, based on the results of the analysis of the main components in Table 3, the value of PC1 has a variance (or eigenvalue) that is greater than 1, which is 4.081. This first component can explain 81.6 percent of the total variance. The scores from the formed principal components can be calculated by looking at the coefficient values for each variable in Table 3 below.

Table 3. Principal Component Analysis Result

No	Variables	PC_1	PC_2	PC_3	PC_4	PC_5
1	X_1	0.443	0.473	-0.352	-0.615	-0.279
2	X_2	0.437	0.616	0.368	0.424	0.339
3	X_3	0.462	-0.276	0.210	0.338	-0.473
4	X_4	0.445	-0.470	0.470	-0.476	0.367
5	X_5	0.449	-0.317	-0.690	0.318	0.347
6	Eigenvalue	4.081	0.374	0.217	0.204	0.123
7	Proportion	0.816	0.075	0.043	0.041	0.025
8	Cumulative	0.816	0.891	0.934	0.975	1.000

It is very subjective at deciding the numbers of principal components that should be used in the model. The researcher chooses only three principal components which are PC1, PC2, and PC3 that represented 93.4 percent of the total variance. It is because they combine for more than 90 percent, which is considered to have captured the overall data structure in the model, while the other two components have a small proportion of the total variance so that it can be considered insignificant in the model.
The principal component of PC1, PC2, and PC3 can be written as follows:

$$PC_1 = 0.443X_1 + 0.437X_2 + 0.462X_3 + 0.445X_4 + 0.449X_5$$
$$PC_2 = 0.473X_1 + 0.616X_2 - 0.276X_3 - 0.470X_4 - 0.317X_5$$
$$PC_3 = -0.352 + 0.368X_2 + 0.210X_3 + 0.470X_4 - 0.690X_5$$

The results of these three main components still need to be checked whether there is still multicollinearity or not. Based on the output shown in Table 4, it turns out that there is no strong correlation between components that exceed the value of 0.8 or it can be concluded that there is no multicollinearity among the principal components. Therefore, MARS modeling and logistic regression using the principal components PC1, PC2, and PC3 as predictor variables can be continued in Table 4 below.

Table 4. Coefficient of correlation among principal components

No	Correlation	PC_1	PC_2	PC_3
1	PC_1	1	0.219	-0.608
2	PC_2	0.219	1	-0.136
3	PC_3	-0.608	-0.136	1

1. Logistic Regression Method

The results of the logistic regression modeling output are shown by the following equation:

$$\pi(x) = \frac{e^{10.699-0.119PC_1+0.045PC_2+0.027PC_3}}{1+e^{10.699-0.119PC_1+0.045PC_2+0.027PC_3}}$$

In particular, we transform the model into a natural log of the odds ratio as follows:

$$\ln\left(\frac{\pi(x)}{1-\pi(x)}\right) = 10.699 - 0.119PC_1 + 0.045PC_2 + 0.027PC_3$$

Based on the simultaneous Likelihood Ratio Test concludes that there is at least one component that affects the response variable. This can be shown by the value of G with a significance level (sig) 0.002 which means that H_0 is rejected. The Wald Chi-Square statistics, which tests the unique contribution of each predictor. Based on the test shows that out of the three components, only one principal component significantly affects the response variable. This can be seen from the significance column of Table 5, where PC1 has a significance value of 0.044 or less than 0.05 standard for statistical significance. Let us now interpret the odds ratio: The 0.887 odds ratio for PC1 indicates that for each one-point increase on PC1, the odds of success of the response variable increasing by a multiplicative factor of 0.887.

Table 5. Logistic Regression Output

No	Variable	*B*	Wald Test	Sig.	Exp(B)
1	PC_1	-0.119	4.042	0.044	0.887
2	PC_2	0.045	0.400	0.527	1.046
3	PC_3	0.027	0.083	0.774	1.028
4	Constant	10.699	2.983	0.084	44,312.815

From Table 5, the classification table shows the accuracy values of the logistic regression model is 80 percent. The miss classification of all the observations is 20 percent. Respectively, the sensitivity and specificity values were 87.5 percent and 66.67 percent.

2. MARS Model

The procedure to get the best MARS model is done by a process of trial and error of a combination of the number of Basic Function (BF), Maximum Interaction (MI) and Minimum Observation (MO). In this study, we use the Basic Function 2 to 3 times of the principal component used in the model. The MI that is used in the model is 1 to 3 and the number of MO per knot is 0 and 1.

Table 6. Summary of MARS Model

No	BF	MI	MO	GCV	Number of significance Principal Components	*Accuracy*
1.	6	1	0	0.1601	1	84%
2.	6	1	1	0.1091	3	92%
3.	6	2	0	0.1745	1	84%
4.	6	2	1	0.1225	3	92%
5.	6	3	0	0.1745	1	84%
6.	6	3	1	0.1225	3	92%
7.	9	1	0	0.1601	1	84%
8.	9	1	1	0.1091	3	92%
9.	9	2	0	0.1745	1	84%
10.	9	2	1	0.1225	3	92%

No	BF	MI	MO	GCV	Number of significance Principal Components	*Accuracy*
11.	9	3	0	0.1745	1	84%
12.	9	3	1	0.1225	3	92%
13.	12	1	0	0.1601	1	84%
14.	12	1	1	0.1091	3	92%
15.	12	2	0	0.1745	1	84%
16.	12	2	1	0.1225	3	92%
17.	12	3	0	0.1745	1	84%
18.	12	3	1	0.1225	3	92%

Based on the MARS program output summarized in Table 6, it can be seen that the value of MO = 1 both in interactions 1 and 2, and BF values of 6, 9 and 12 always make MARS modeling have a maximum accuracy value. Besides, the results show that the best MARS model is in the combination of BF = 6, MI = 1, and MO = 1 because it has the lowest GCV value, maximum accuracy and the most significant number of the principal components. This equation can be written:

$$\ln\left(\frac{\pi(x)}{1-\pi(x)}\right) = -0.037 + 0.012BF2 + 0.101 * BF3 + 0.093 * BF5$$

Where:

$$BF2 = max(0, 117.982 - PC_1);$$
$$BF3 = max(0, PC_3 - 14.056);$$
$$BF5 = max(0, PC_2 - 10.686)$$

Let us interpret the MARS Model:

a. $BF2$: If PC1 has a value of less than 117,982, the odds of success of the response variable are $e^{0,012}$ times compared to PC1 which has a value of more than 117,982.
b. $BF3$: If PC3 has a value of less than 14,056, the odds of success of the response variable are $e^{0,101}$ times compared to PC3 which has a value of more than 14,056.
c. $BF5$: If PC2 has a value of less than 10,686, the odds of success of the response variable are $e^{0,093}$ times compared to PC1 which has a value of more than 10,686.

The MARS method also has other advantages that can capture the level of importance in each component based on the -GCV in the model. Based on Attachment 1 (Relative and Variable Importance), it can be seen that PC1, PC3, and PC2 have an interest level of 100%, 78,943%, and 50,047%, respectively. The results of MARS modeling stated that the sensitivity and specificity values produced in MARS modeling were 100% and 77.80%.

3. Comparison of Logistics Regression and MARS Model

A comparison of the two methods performed using the criteria for the number of significant principal components, as well as the accuracy, sensitivity, and specificity values are shown in Table 7.

Table 7. The Comparison of Logistics Regression and MARS Model

No	Method	Number of significance Principal Components	*Accuracy*	*Sensitivity*	*Specificity*
1.	Logistic Regression	1	80 %	80.75%	66.67%
2.	MARS	3	92 %	100%	77.80%

The higher the four criteria, it can be ensured that the better the model formed. Based on this comparability, the MARS method has more appropriate performance than logistic regression because the differences in the four criteria are highly noticeable. Besides, it can be said that the MARS method is very powerful in modeling research simulation data.

D. CONCLUSION AND SUGGESTIONS

From the preceding discussion, the data set is plagued by multicollinearity problems. Thus, we ran the Principal Component Analysis to address the problem to get better comparison results between logistics regression and MARS. The study results show that MARS performs better than logistics regression based on four criteria, which include numbers of significant Principal Component (MARS has three significant; logistics regression has one significant), accuracy (MARS is 92%; logistics regression is 80%), sensitivity (MARS is 100%; logistics regression 80.75%), and specificity (MARS is 77.80 %; logistics regression is 66.67%). Also, we can conclude that the MARS method is very powerful in modeling research simulation data.

REFERENCES

Adityawardhani et al. (2017). Comparison of Logistic Regression Model and Mars Classification Results on Binary Response for Teknisi Ahli BBPLK Serang Training Graduates Status. *DOARJ: International Journal of Humanities, Religion and Social Science*, *2*(1), 14-20

Diop et al. (2011). Maximum Likelihood Estimation in the Logistic Regression Model in a Cure Fraction. *Electronic Journal of Statistics*, 5(1), 460-483

Felicísimo et al. (2013). Mapping landslide susceptibility with logistic regression, multiple adaptive regression splines, classification and regression trees, and maximum entropy methods: A comparative study. *Landslides*. 10(1), 175-189

Gwelo, Abubakari S. (2019). Principal Components To Overcome Multicollinearity Problem. *Oradea Journal of Business and Economics*, 4(1), 79-91

Hosmer, D., Lemeshow, S & Sturdivant X, R. (2013). *Applied logistic regression Third Edition*. New York: John Wiley & Sons,Inc.

Johnson RA, Wichern DW (2013). *Applied Multivariate Statistical Analysis*. Pearson New International Sixth Edition Paperback.

Kayri, M. (2010). The Analysis of Internet Addiction Scale Using Multivariate Adaptive Regression Splines. *Iranian Journal of Public Health, 39*(2), 51-63

Kilinç et al. (2017). Using Multivariate Adaptive Regression Splines to Estimate Pollution in Soil. *International Journal of Advanced and Applied Sciences*, 4(2), 10-16

Koc, E. K. & Bozdogan, H. (2014). Model selection in multivariate adaptive regression splines (MARS) using information complexity as the fitness function. Springer, June 2014, 35-58

Midi et al. (2013). Collinearity diagnostics of binary logistic regression model. *Journal of Interdisciplinary Mathematics,* Mei 2013, 253-267.

Mina, C. D., & Barrios, E. B. (2010). Profiling poverty with multivariate adaptive regression splines. Phillipine Journal of Development, 37(2): 55-97

Mukaka, M.M. (2012). Statistics Corner: A guide to appropriate use of Correlation coefficient in medical research. *Malawi Medical Journal, 24*(3), 69-71

Osetljivosti et al. (2014). Understanding sensitivity, specificity and predictive values. Vojnosanitetski pregled. *Military-medical and Pharmaceutical Review,* 71(11), 1062-1065.

Park et al. (2017). Evaluation of Logistic Regression and Multivariate Adaptive Regression Spline Models for Groundwater Potential Mapping Using R and GI. *Sustainability, 9*(7), 1-20

Rahayu et al. (2017). Application of Principal Component Analysis (PCA) to Reduce Multicollinearity Exchange Rate Currency of Some Countries in Asia Period 2004-2014. *International Journal of Educational Methodology*, 3(2), 75-83.

Schober et al. (2018). Correlation Coefficients: Appropriate Use and Interpretation. *Anesthesia and Analgesia*, *126* (5), 1763-1768

Wahyuningrum, S. (2009). *Mars Classification For Indigent Village/Sub-District at East Kalimantan in 2005*. Surabaya: Institut Tekhnologi Sepuluh November.

Zewude et al. (2016). Binary Logistic Regression Analysis in Assessment and Identifying Factors That Influence Students' Academic Achievement: The Case of College of Natural and Computational Science, Wolaita Sodo University, Ethiopia. *Journal of Education and Practice*, 7(25), 3-7.

Zhang, W & Goh, A. T. C. (2016). Multivariate adaptive regression splines and neural network model for prediction of pile drivability. *Geoscience Frontiers, 7*(1), 45-52.

ATTACHMENT

Attachment 1

```
Relative Variable Importance
============================

     Variable       Importance         -gcv
-------------------------------------------
   1      PC1          100.000        0.156
   3      PC3           78.943        0.138
   2      PC2           50.046        0.121

ORDINARY LEAST SQUARES RESULTS
==============================

N: 25.000                                R-SQUARED: 0.754
MEAN DEP VAR: 0.360                  ADJ R-SQUARED: 0.719
                  UNCENTERED R-SQUARED = R-0 SQUARED: 0.843

   PARAMETER                  ESTIMATE         S.E.      T-RATIO      P-VALUE
-------------------------------------------------------------------------------
Constant                |       -0.037        0.076       -0.485        0.632
Basis Function 2        |        0.012        0.003        4.008  .636737E-03
Basis Function 3        |        0.101        0.029        3.445        0.002
Basis Function 5        |        0.093        0.033        2.778        0.011
-------------------------------------------------------------------------------
F-STATISTIC =  21.512                     S.E. OF REGRESSION =  0.260
    P-VALUE =  .131443E-05             RESIDUAL SUM OF SQUARES =  1.414
  [MDF,NDF] = [ 3, 21 ]             REGRESSION SUM OF SQUARES =  4.346
-------------------------------------------------------------------------------
The Following Graphics Are Piecewise Linear
Basis Functions
===============

BF2 = max(0, 117.982 - PC1 );
BF3 = max(0, PC3 - 14.056);
BF5 = max(0, PC2 - 10.686);

Y = -0.037 + 0.012 * BF2 + 0.101 * BF3 + 0.093 * BF5;

model Y = BF2 BF3 BF5;
====================================
LEARNING SAMPLE CLASSIFICATION TABLE
====================================
      Actual         Predicted Class         Actual
       Class          0            1          Total
-----------------------------------------------------
           0       16.000        0.000       16.000
           1        2.000        7.000        9.000
-----------------------------------------------------
Pred. Tot.         18.000        7.000       25.000
Correct             1.000        0.778
Success Ind.        0.360        0.418
Tot. Correct        0.920

    Sensitivity:        1.000     Specificity:        0.778
False Reference:        0.111  False Response:        0.000
Reference = Class 0, Response = Class 1
-----------------------------------------------------------
```