BAREKENG: Journal of Mathematics and Its Applications
March 2024 Volume 18 Issue 1 Page 0373–0382
P-ISSN: 1978-7227 E-ISSN: 2615-3017
https://doi.org/10.30598/barekengvol18iss1pp0373-0382

NONLINEAR PRINCIPAL COMPONENT ANALYSIS IN PATH ANALYSIS WITH LATENT VARIABLES MIXED DATA

Rindu Hardianti[1*], Solimun[2], Nurjannah[3]

[1,2,3]*Departement of Statistics, Faculty of Mathematics & Natural Sciences, University of Brawijaya*
Veteran Ketawanggede Street, Lowokwaru, Malang, 65145, Indonesia

Corresponding author's e-mail: ***rinduhardianti@student.ub.ac.id***

ABSTRACT

Article History:

Received: 3rd September 2023
Revised: 8th December 2023
Accepted: 7th January 2024

Keywords:

Latent Variables;
Mixed Data;
Nonlinear Principal Component Analysis;
Path Analysis.

This study aims to obtain the main component score of the ability to pay latent variable, determine the strongest indicators forming the ability to pay on a mixed scale based on defined indicators, and model the ability to pay on time mediated by fear of paying using path analysis. The data used in this study is secondary data from mortgage-paying customers with a sample size of 100. The method used is nonlinear principal component analysis with path analysis modeling. The results of this study indicate that the eleven variables formed by PC1 or X1 are able to store diversity or information by 32.50%, while 67.50% of diversity or other information is not stored (wasted). The credit term is the strongest indicator that forms the ability to pay variable. The variable ability to pay mortgages has a significant effect on payments by mediating the fear of paying late with a coefficient of determination of 80.40%.

How to cite this article:

R. Hardianti, Solimun and Nurjannah., "NONLINEAR PRINCIPAL COMPONENT ANALYSIS IN PATH ANALYSIS WITH LATENT VARIABLES MIXED DATA" *BAREKENG: J. Math. & App.*, vol. 18, iss. 1, pp. 0373-0382, March, 2024.

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

Home Ownership Credit (mortgages) is one of the most common credit facilities provided by banks in Indonesia and many other countries. According to [1], a mortgage is a type of loan provided by a particular bank to individual customers either to buy a house or to repair a house. The main purpose of mortgages is to enable customers to own a house or property without having to pay a large amount of cash at once. Mortgages allow regular payments within a certain period. Each bank has different requirements for mortgages. However, in general, prospective customers need to meet requirements such as having an adequate income, having a permanent job, or a stable business. Before a bank provides a mortgage to a customer, it is important for the bank to assess the prospective customer. So that the risk of congestion in the process of returning customer mortgages to the bank can be minimized. According to [2], the risk of congestion in the mortgage repayment process is characterized by a high level of credit repayment by customers who are not on time (non-performing loans). To reduce the risk of traffic jams in the process of returning a customer's mortgage to the bank is to examine the customer's good faith and ability to pay off their debt and analyze the relationship between the customer's ability to pay off the debt and the timeliness of paying the customer. Through the fear of paying late, it is hoped that the bank will gain trust. in providing financing or credit in question.

Path analysis is a multivariate analysis method used to test models of relationships between variables in the form of cause and effect. Path analysis was first developed by Wright in 1934 and was used to determine the direct and indirect effects between variables [3]. Variables are characteristics of a subject or object that are relevant to the problem being researched, where there are various variables viewed from various points of view. Based on the measurement process, variables are divided into manifest variables (observable) and latent variables (unobservable) [4].

According to [5], latent variables are generally defined as variables that cannot be measured directly, but must be measured through indicators that reflect or structure them. Latent variables can be classified into variables in the form of psychological attributes, such as satisfaction in the form of conceptual variables, and can also be classified into latent variables which are factual in nature, such as the variable ability to pay house installments, which will be examined in this research. The variable of ability to pay mortgages can be measured by measuring the indicators that compose and form this variable. According to [6], indicators whose nature is to form or compose variables are called formative indicator models, where in this indicator model, the indicators that compose them are not required to have a common factor. To measure latent variable data with a formative indicator model, the principal component assessment method obtained from the Principal Component Analysis (PCA) is used.

The PCA is a multivariate analysis introduced by Karl Pearson [7]. The PCA is a technique used to simplify data, by means of linear transformation to form a new coordinate system with maximum variation. The PCA can be used to reduce the dimensions of data without significantly reducing the characteristics of the data. Data reduction is carried out by means of linear transformation so that a new coordinate system is formed with maximum variation without significantly reducing the characteristics of the data [8].

The PCA does not allow the analysis of data with nonlinear relationships between variables or with indicators of mixed scale (metric and non-metric) latent variables. However, in its application, not all indicators of latent variables are metric data. According to [9], based on the scale of measurement, data can be divided into nominal, ordinal, interval, and ratio data. Nominal and ordinal data scales are qualitative (non-metric), so the data must be converted into numerical data by giving a score. Interval and ratio scales belong to the category of quantitative (metric) data.

According to [10], when analyzing data that are still on a mixed scale consisting of a nonmetric scale, nonlinear PCA can be carried out with nonlinear transformation, namely by optimal scaling or optimal transformation from a qualitative scale to quantitative values. Nonlinear PCA is used to analyze variables with nominal or ordinal data scales when these variables include "categorical" variables and is also used to resolve variables that have a nonlinear relationship.

Based on the previous explanation, from the eleven indicators that form the ability to pay mortgage variable, the researcher aims to determine the most significant indicator in forming the ability to pay a mortgage variable, utilizing nonlinear PCA. This is achieved by transforming the original independent variables into new uncorrelated variables, thus avoiding any correlation between them. After obtaining several components from the PCA results that are free of multicollinearity, these components are transformed into new independent variables. They are then analyzed to determine their impact and ability to determine time to pay. Path analysis is used to investigate the mediation of fear of paying and to explain the level of data diversity.

2. RESEARCH METHODS

2.1 Data

The data in the research is secondary data obtained through distributing questionnaires with a mixed data scale, namely a Likert scale for the variables of fear of being late in paying and paying mortgages on time. Meanwhile, the variable ability to pay mortgages uses an ordinal data scale. The sampling technique used was purposive sampling. Purposive sampling is a sampling technique that is based on certain characteristics or conditions that are the same as the characteristics of the population. The number of samples used in the research was 100 customers. This study uses two analytical methods with a mixed method approach, namely Nonlinear PCA with path analysis.

2.2 Research Variables

This research is designed to answer the problems that have been formulated, as well as to achieve research objectives by involving hypothesis testing to determine the effect between research variables. **Figure 1** is a research model used in this research.

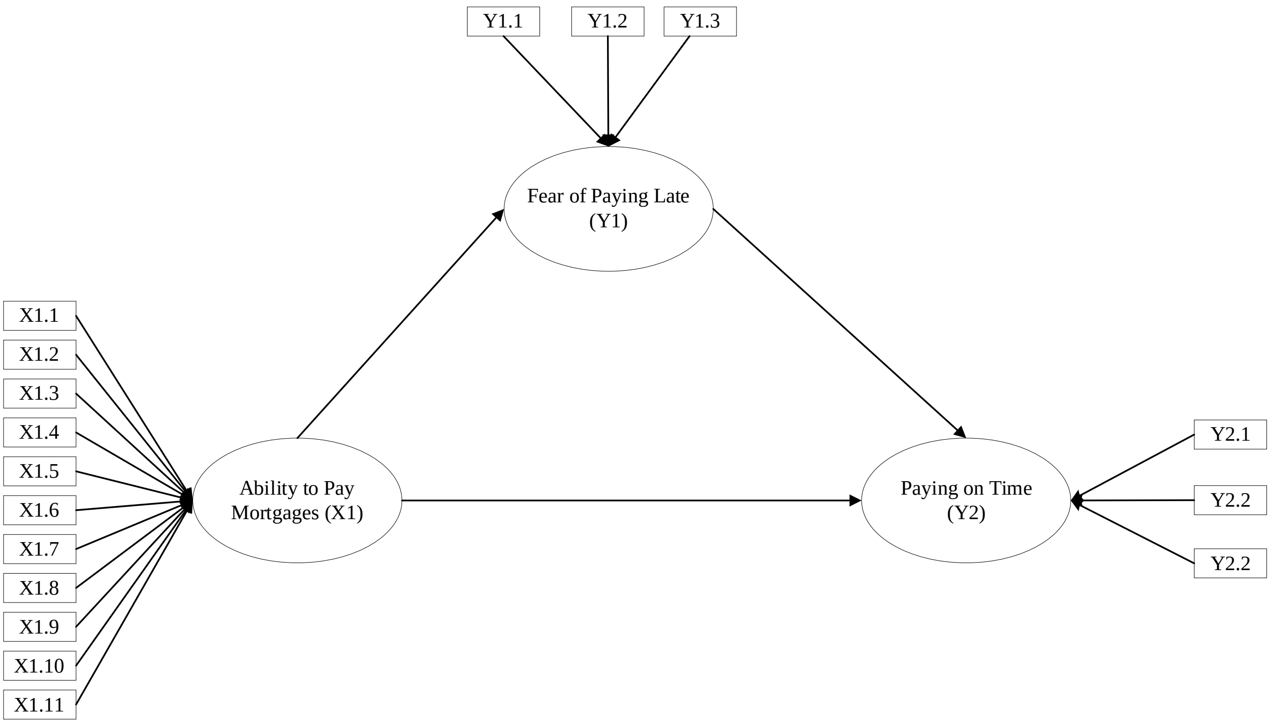

Figure 1. Research Model

The following is an indicator of the variable ability to pay mortgages, which are analyzed using nonlinear PCA in **Tabel 1** below.

Table 1. Variable Indicators of Ability to Pay Mortgages

Indicator	Data Scale	Category	
Guarantee Document (X1.1)	Ordinal	Other	1
		Parent SHM / Main SHGB	2
		SHGB	3
		SHM	4
Education (X1.2)	Ordinal	SD, SMP, SMA	1
		Diploma	2
		S1,S2,S3	3
Collectability Status (X1.3)	Ordinal	In Special Attention	1
		Fluent	2
Work Experience (X1.4)	Ordinal	$\leq$ 3Year	1
		$> 3 - \leq 6$ Year	2
		$> 6 - \leq 12$ Year	3
		$> 12 - \leq 18$ Year	4
		$> 18 - \leq 20$ Year	5
		> 20 Year	6
Joint Income (X1.5)	Ordinal	Non-joint income	1
		Have joint income	2
RPA (Instalment Income Ratio) (X1.6)	Ordinal	$\leq$ 1,5 Million	1
		$> 1{,}5 - \leq 2$ Million	2
		$> 2 - \leq 2{,}5$ Million	3
		$> 2{,}5 - \leq 3$ Million	4
		> 3 Million	5
Form of Business Entity (X1.7)	Ordinal	Other	1
		PERUM, PERSERO	2
		PT Non Tbk	3
Credit Term (X1.8)	Ordinal	> 240 Month	1
		$> 180 - \leq 240$ Month	2
		$> 120 - \leq 180$ Month	3
		$> 60 - \leq 120$ Month	4
		$> 48 - \leq 60$ Month	5
		$> 36 - \leq 48$ Month	6
		≤ 36 Month	7
Number of Family Responsibilities (X1.9)	Ordinal	> 2 Person	1
		2 Persons	2
		1 Person	3
		0 Person	4
Savings Ownership (X1.10)	Ordinal	Do not have	1
		Have Savings in Other Banks	2
		Have Savings	3
Loan to Value (X1.11)	Ordinal	$> 95\%$	1
		$> 90\% - \leq 95\%$	2
		$> 80\% - \leq 90\%$	3
		$> 70\% - \leq 80\%$	4
		$> 60\% - \leq 70\%$	5
		$> 50\% - \leq 60\%$	6
		$\leq 50\%$	7

The variables Fear of Paying (Y1) and Timeliness of Paying (Y2) are endogenous variables on an ordinal scale, namely a Likert scale, so it is necessary to scale the data using SRS (summated rating scale). Data that has been scaled using SRS has an interval scale. With this method, each response to an item is given a score that assumes a normal distribution. The SRS method is carried out by converting questionnaire score data with z scores using a normal distribution. Data scaling is assisted by Microsoft Excel software. Table 2 is an endogenous variable in this research.

Table 2. Variable Indicators of Ability to Pay Mortgages Variables Fear of Paying and Timeliness of Paying

Variable	Indicator	Data Scale
Fear of Paying Late (Y1)	Physical Symptoms (Y1.1)	Likert scale
	Behavioral symptoms (Y1.2)	
	Cognitive Symptoms (Y1.3)	
Timeliness of Payment (Y2)	Timely Desire Pays (Y2.1)	
	Always Ontime Monthly Payment (Y2.2)	
	Timeliness of Payment History (Y2.3)	

2.3 Nonlinear Principal Component Analysis

Linear PCA assumes that the data are quantitative so that Linear PCA cannot be directly used for qualitative data. In Nonlinear PCA, qualitative data with nominal and ordinal (nonmetric) scale variables are done by transforming nonlinear data into quantitative data. Thus, linear PCA using the quantification method (optimal scaling technique) is a Nonlinear PCA **[11].**

This is reinforced by **[12]** in his book, which states that Nonlinear PCA is used to analyze variables with nominal and ordinal data scales, where he calls them "categorical" variables and is also used to solve variables that have nonlinear relationships. Nonlinear PCA converts each category into a numerical value with a quantification process which is called "optimal quantification" or "optimal scaling."

The initial stage in nonlinear PCA is the quantification process. Quantification is assigning a numerical value to a category of a variable. The quantification technique used in this analysis is the optimal scaling technique. For example, qualitative data is formed in a matrix H of size $n \times m$. The matrix and notation that will be used in nonlinear PCA can be seen in **Equation (1)**.

$$\mathbf{H} = (\boldsymbol{h}_{ij}) = \begin{pmatrix} h_{11} & h_{12} & \dots & h_{1m} \\ \vdots & \vdots & \ddots & \vdots \\ h_{n1} & h\_n2 & \dots & h_{nm} \end{pmatrix} \tag{1}$$

$$\text{with } \boldsymbol{h}_{ij} = \begin{pmatrix} h_{11} \\ \vdots \\ h_{n1} \end{pmatrix}$$

where:

$\mathbf{H}$	:	Sized qualitative data matrix $n \times m$
n	:	Many objects of observation, $i = 1,2, \dots, n$
m	:	Many variables, $j = 1,2, \dots, m$
k_j	:	Many categories in the i-th variable j, $r = 1,2, \dots, k_j$
$\boldsymbol{h}_{ij}$	:	Vector of the i-th object i in the i-th category k in the i-th variable j

To quantify the matrix **H,** where $\boldsymbol{h}_j$ is the column vector j of the matrix $\mathbf{H}$ of size $n \times m$, it becomes an indicator matrix $\mathbf{G}_J$ of size $n \times k_j$ as in **Equation (2)**.

$$G_J = (g_{ijk}) = \begin{pmatrix} g_{j11} & \cdots & g_{j1k_j} \\ \vdots & \vdots & \vdots \\ g_{jn1} & \cdots & g_{jnk_j} \end{pmatrix} = (g_{j1} \quad \cdots \quad g_{jk_j}) \tag{2}$$

$$g_{ijk} \begin{cases} 1, & \text{if the i} - \text{th object in the j} - \text{th variable is in the k} - \text{th category} \\ 0, & \text{if the i} - \text{th object is not in the k} - \text{th category in the j} - \text{th} \end{cases}$$

where:

$\mathbf{G}_\mathrm{j}$	:	Indicator matrix of h_j size$n \times k_j$
g_{ijk}	:	Matrix column vector $\mathbf{G}_\mathrm{j}$

2.4 Path Analysis

Path analysis is an extension of the regression model that can be used to test the correlation matrix in causal models compared by researchers. The path analysis model is a development of multiple regression analysis, as in **Equation (3)**.

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \ldots + \beta_p X_{pi} + \varepsilon_i \text{ , } i = 1,2,\ldots,n \tag{3}$$

where:

Y_i	: The value of the variable response to the i-th observation
X_{ki}	: The value of the k-th predictor variable, the i-th observation
β_0	: *Intercepts*
$\beta_1, \beta_2, \ldots, \beta_p$	: *Slopes*
ε_i	: *errors*, $\varepsilon_i \sim$ N $(0,\sigma^2)$
n	: The number of observation units
k	: unit of observation predictor k, $k = 1,2,\ldots,p$
p	: The number of units of the predictor variable
i	: Unit of observation to - i, $i = 1,2,\ldots,n$

Equation (3) does not use *standardize data yet*. Standardization can be done by standard normal transformation which has an average of zero and a variance of one as in **Equation (4) [13]**.

$$z_{x_i} = \frac{x_i - \bar{x}}{s_x} \text{ and } z_{y_i} = \frac{y_i - \bar{y}}{s_y} \tag{4}$$

The path equation for endogenous variables is in **Equation (5)**

$$Y_i = \beta_0 + \beta_{X_1Y} X_{1i} + \beta_{X_2Y} X_{2i} + \varepsilon_i \tag{5}$$

Equation (5) can be standardized to become **Equation (6).**

$$Z_{Y_i} = \beta_{X_1Y} Z_{X_{1i}} + \beta_{X_2Y} Z_{X_{2i}} + \varepsilon_i \tag{6}$$

2.5 Coefficient of Determination

According to **[14]**, coefficient of determination is information used to see the suitability of a model with the symbol R^2. The coefficient of determination value ranges from 0 to 1. If R^2 is close to 1, it can be said that the influence of the independent variable on the dependent variable is large. So, the model used can explain the influence of these variables. Vice versa if R^2 approaches 0. Calculating the coefficient of determination can be done using **Equation (7)**.

$$R_m^2 = 1 - P_{e1}^2 P_{e2}^2 \ldots P_{ep}^2$$
$$P_{ei} = \sqrt{1 - R_i^2} \tag{7}$$

where:

P_{ei}	: The residual effect on each equation ($i = 1,2,\ldots,p$)
R_i^2	: Coefficient of total determination in each equation
R_m^2	: Coefficient of total determination

3. RESULTS AND DISCUSSION

3.1 Nonlinear Principal Component Analysis

There are nine indicators that form the customer's ability to pay mortgage variables. These indicators are analyzed by nonlinear principal components (PC) to obtain PC scores for the variables. According to **[15]**, in determining how many PC should be taken, three methods are used, one of which is by using eigenvalues. The selected component must have an eigenvalue >1. The eigenvalues from the analysis of the main components of the ability to pay mortgages are presented in **Table 3**.

Table 3. Eigen Value of the Ability to Pay Mortgages Variables

	Eigen Value
PC1	2.076
PC2	0.251
PC3	0.212
PC4	0.143

	Eigen Value
PC5	0.094
PC6	0.519
PC7	0.846
PC8	0.322
PC9	0.244
PC10	0.119
PC11	0.658

Based on **Table 3**, the main component used is the main component with the highest eigenvalue or highest diversity, namely the first component or also called PC1 which represents the latent variable being analyzed. The cumulative variance in the first component is 0.3250, meaning that the first main component contains 32.50% of the data from the variable Ability to Pay Mortgage. **Table 4** shows the weight of the first component in the Mortgage Paying Variable.

Table 4. The Weight of the Principal Components of the Ability to Pay Mortgage Variables

Indicator	Weight
Guarantee Document (X1.1)	0.361
Education (X1.2)	0.386
Collectability Status (X1.3)	0.438
Work Experience (X1.4)	0.507
Joint Income (X1.5)	0.522
RPA (Instalment Income Ratio) (X1.6)	0.476
Form of Business Entity (X1.7)	0.446
Credit Term (X1.8)	0.538
Number of Family Responsibilities (X1.9)	0.361
Savings Ownership (X1.10)	0.392
Loan to Value (X1.11)	0.345

Based on the weight of the PC, the indicator that has the greatest weight is the X1.8 indicator, namely the Credit Period. This means that the X18 indicator is able to characterize the Mortgage Paying Ability variable. Based on **Table 4**, the PC linear combination equation can be formed to get the component score, which is the value of the Mortgage Paying Ability variable (X1).

$$X_1 = 0{,}361X_{1.1} + 0{,}386X_{1.2} + 0{,}438X_{1.3} + 0{,}507X_{1.4} + 0{,}522X_{1.5} + 0{,}476X_{1.6} + 0{,}446X_{1.7} + 0{,}538X_{1.8} + 0{,}361X_{1.9} + 0{,}392X_{1.10} + 0{,}345X_{1.11}$$

3.2 Path Analysis

3.2.1. A Path Analysis Assumptions

The results of testing the assumptions of path analysis in this study are as follows.

1) The Relationship between Variables is Linear and Additive

The linearity test was carried out using the RESET method with Rstudio software with the output results shown in **Tabel 5**.

Table 5. Linearity Test Results

Variable Relations	P-Value	Connection
X1 against Y1	0.4377	linear
X1 against Y2	0.5698	linear
Y1 against Y2	0.6714	linear

Based on **Table 3**. it can be seen that there is a relationship between exogenous variables and endogenous variables $p - \text{value} > 0.05$ means accept H_0 so that it can be said that the assumption of linearity has been met.

2) Minimal Endogenous Variable in Interval Measurement Scale

The data used is secondary data in the form of a Likert scale, where the score produced on the Likert scale is data that is close to the interval scale. Then the score that has been obtained is carried out by a scaling process using the Summated Rating Scale (SRS) method. Therefore, the assumption of a minimum endogenous variable measuring interval scale is fulfilled.

3) Normality Assumption

The normality assumption test is used to identify residuals in normally distributed research or not. The regression model can be said to be good if the residuals are normally distributed. Based on the results of the analysis obtained value $p-\text{value}\ (0.2) > \alpha\ (0.05)$, it can be concluded to reject H_0. So, it can be concluded that with a significant level of 5%, the residuals are normally distributed.

4) Models are Recursive

The model in path analysis is said to be recursive if it has a one-way relationship pattern. In Figure 2 it can be seen that each exogenous variable has one-way causality and there is no two-way (reciprocal) relationship so that it can be said that the model is recursive.

5) The Analyzed Model is Correctly Specified Based on Relevant Theories and Concepts

The model form and variables involved in the model are based on expert opinion theory and the results of previous research.

3.2.2. Parameter Estimation and Hypothesis Testing

Parameter estimation in the path analysis is carried out to estimate the path coefficient. This is used to determine the relationship between exogenous variables and endogenous variables, while hypothesis testing is used to test the significance of the path coefficient partially. The hypothesis used is as follows.

H_0: $\rho_{xy} = 0$ (There is no significant effect of exogenous variables on endogenous variables) vs

H_1: $\rho_{xy} \neq 0$ (There is a significant effect of exogenous variables on endogenous variables)

Table 6. Results of Parameter Estimation and Hypothesis Testing

Variable Relations	Path Coefficient	P-Values	Decision
X1 against Y1	0.4181	0.0045	Reject H_0
X1 against Y2	0.4452	0.0127	Reject H_0
Y1 against Y2	0.3939	0.0012	Reject H_0

Based on Table 6, it can be seen that the decision is reject H_0, which means there is a significant influence of exogenous variables on endogenous variables, with the results of estimating the parameters of the path analysis can be formed as follows:

$$Z_{Y1} = 0.4452 Z_{X1}$$

$$Z_{Y2} = 0.4452 Z_{X1} + 0.3939 Z_{Y1}$$

with diagrams and path coefficients as Figure 2:

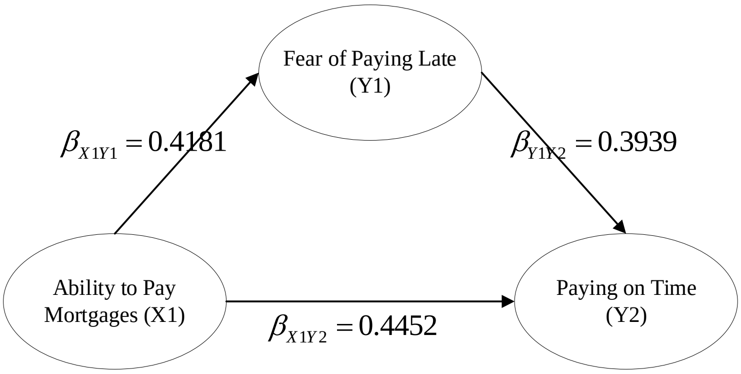

Figure 2. Path Diagram and Coefficient

3.2.3. Path Analysis Model Validity

The validity of the model in path analysis can be known by looking at the total determination coefficient value. The total coefficient of determination is used to explain the diversity of data that can be explained by the model. Calculation of the total determination coefficient obtained from each model that is formed in the path analysis. The coefficient of determination for each model is obtained using the RStudio software. The results of the coefficient of determination can be seen in **Table 7**.

Table 7. Coefficient of Determination

Model	Coefficient of Determination (R^2)
1	0.6123
2	0.5537

Based on **Table 7**, the total determination coefficient is calculated as follows.

1) Model 1

$$R_1^2 = 0.6123$$

$$P_{e1} = \sqrt{1 - R_1^2} = 0.6226$$

2) Model 2

$$R_1^2 = 0.5537$$

$$P_{e2} = \sqrt{1 - R_2^2} = 0.6680$$

The total determination coefficient of model 1 and model 2 is as follows.

$$R_t^2 = 1 - P_{e1}^2 P_{e2}^2$$

$$R_t^2 = 1 - ((0.6226)^2 \times (0.6680)^2)$$

$$R_t^2 = 0.8040$$

The total determination coefficient value of 0.8040 explains that 80.40% of the data diversity can be explained by the research model, while 19.6% of the data diversity is explained by other variables outside the model.

3.3 Discussion

The results showed that the ability to pay mortgages variable. The credit term indicator is the indicator which best characterizes the ability to pay mortgages variable compared to other indicators. This means that the credit term indicator is the most capable indicator to represent the variable of the ability to pay mortgages. The longer the credit period, the lower the monthly installments that must be paid by the customer. This can help ease the burden of monthly payments and give customers more time to repay loans. The longer the credit term, the higher the amount of interest to be paid, so that the overall cost of the loan will increase. Therefore. before taking credit for a certain period of time, it is important to consider the customer's financial capabilities so that they can pay installments consistently and on time.

Based on the research results, the variable ability to pay mortgages has a significant effect on timely payments, which is mediated by the fear of paying late. The ability to pay shows the extent to which the customer has the financial capacity to pay installments according to the agreement. If the customer has a good paying ability, that is, has sufficient income or cash flow to cover credit payments on schedule, this helps ensure that customers can pay debts consistently and don't miss the payments. On the other hand, if customers have problems paying on time, for example, they often miss payments or experience repeated delays, this could have a negative impact on their financial credibility. In some cases, this may lead to additional penalties or fees, and may even result in default, which may negatively affect the ability to obtain credit in the future. Therefore, it is important for customers to have good paying ability and be responsible for maintaining discipline in paying credit installments on time.

4. CONCLUSIONS

Based on the results of the analysis it can be concluded that the Nonlinear PCA was used to obtain latent variables with indicators on a non-metric scale. It was found that the credit term indicator was the indicator that best characterized the ability to pay mortgages compared to other indicators. The results showed that the eleven variables formed by PC1 or X1 were able to store diversity or information by 74.8%, while 25.20% of diversity or other information was not stored (wasted). Based on the results of the path analysis, it can be concluded that the ability to pay mortgages has a significant effect on timely payments, which is mediated by fear of paying, with a total coefficient of determination of 0.8040. It explains that 80.40% of the data diversity can be explained by the research model, while 19.6% of the data diversity is explained by other variables outside the model. Moreover, the path analysis model is obtained as follows:

$$Z_{Y1} = 0.4452 Z_{X1}$$

$$Z_{Y2} = 0.4452 Z_{X1} + 0.3939 Z_{Y1}$$

AKNOWLEDGEMENT

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. The authors would like to thank the editors and anonymous reviewers for their comments that help improve the quality of this work.

REFERENCES

[1] Z. Zefriyenni. “Kebijakan Pemberian Kredit Terhadap Penetapan Jumlah Kredit (Studi Khasus Pada Ued-Sp Amanah Sejahtera Sungai Buluh Kecamatansingingi Hilirkabupaten Kuantan Singingi Propinsi Riau).” *Edik Inform.*. vol. 1. no. 1. pp. 72–80. 2017. doi: 10.22202/ei.2014.v1i1.1443.

[2] N. Wahyuni. “Penerapan Prinsip 5C Dalam Pemberian Kredit Sebagai Perlindungan Bank.” *Lex J. Kaji. Huk. Keadilan*. vol. 1. no. 1. 2017. doi: 10.25139/lex.v1i1.236.

[3] A. A. R. Fernandes and Solimun. “Moderating effects orientation and innovation strategy on the effect of uncertainty on the performance of the business environment.” *Int. J. Law Manag.*. vol. 59. no. 6. pp. 1211–1219. 2017. doi: https://doi.org/10.1108/IJLMA-10-2016-0088.

[4] S. Astutik. Solimun. and Darmanto. *Multivariate Analysis: Theory and Application with SAS*. Brawijaya Press University. 2018.

[5] A. A. R. & N. Solimun. S; Fernandes. *Metode Statistika Multivariat. Pemodelan Persamaan Struktural (SEM)*. Malang: UB Press. 2017.

[6] A. M. Tamonob. A. Saefuddi. and A. H. Wigena. “Nonlinear Principal Component Analysis and Principal Component Analysis With Successive Interval in K-Means Cluster Analysis.” *Forum Stat. Dan Komputasi*. vol. 20. no. 2. pp. 68–77. 2015. [Online]. Available: https://journal.ipb.ac.id/index.php/statistika/article/view/16771

[7] I. Djakaria. S. Guritno. and S. H. Kartiko. “Visualisasi Data Iris Menggunakan Analisis Komponen Utama dan Analisis Komponen Utama Kernel Visualization of Iris Data Using Principal Component Analysis and Kernel Principal Component Analysis.” *J. ILMU DASAR*. vol. 11. no. 1. pp. 31–38. 2017.

[8] W. Härdle and L. Simar. “Applied Multivariate Statistical Analysis.” *Appl. Multivar. Stat. Anal.*. no. October. 2019. doi: 10.1007/978-3-662-05802-2.

[9] I. T. Jolliffe and J. Cadima. “Principal Component Analysis: A Review and Recent Developments.” *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.*. vol. 374. no. 2065. p. 20150202. 2016. doi: https://doi.org/10.1098/rsta.2015.0202.

[10] Y. Mori. M. Kuroda. and N. Makino. *Nonlinear Principal Component Analysis and Its Applications*. Springer Singapore. 2016. doi: https://doi.org/10.1007/978-981-10-0159-8.

[11] C. C. David and D. J. Jacobs. “Principal component analysis: a method for determining the essential dynamics of proteins.” *Protein Dyn. Methods Protoc.*. vol. 1084. pp. 193–226. 2014. doi: https://doi.org/10.1007/978-1-62703-658-0_11.

[12] H. Yu. F. Khan. and G. Ikram. “An alternative formulation of PCA for process monitoring using distance correlation.” *Ind. Eng. Chem. Res.*. vol. 55. no. 3. pp. 656–669. 2015. doi: https://doi.org/10.1021/acs.iecr.5b03397.

[13] N. Kock. “Advantages of nonlinear over segmentation analyses in path models.” *Int. J. E-Collaboration*. vol. 12. no. 4. pp. 1–6. 2016. doi: https://doi.org/10.4018/ijec.2016100101.

[14] M. B. T. Efendi. E.C.L.. Fernandes. A.A.R.. Mitakda. “Modeling of path nonparametric truncated spline linear. quadratic. and cubic in model on time paying bank credit.” *Math. Stat. link is disabled*. vol. 9. no. 6. pp. 947–957. 2021.

[15] H. Katayama. Y. Mori. and M. Kuroda. “Variable Selection in Nonlinear Principal Component Analysis.” in *Advances in Principal Component Analysis*. IntechOpen. 2022. pp. 79–89. doi: http://dx.doi.org/10.5772/intechopen.103758.