Journal of Intelligent System and Telecommunications

Journal homepage: https://journal.unesa.ac.id/index.php/jistel/index

Application of Principal Component Analysis (PCA) for Identifying Dominant Factors Affecting Energy Efficiency in a House

Dyah Pranesti Shafira Fitri[1], Mhd. Ali Hanafiah[2]*

[1]Technical and Vocational Education, Postgraduate School, Universitas Negeri Surabaya, Surabaya, East Java, Indonesia

[2]Public Vocational High School 1 Mojokerto, East Java, Indonesia

*Correspondence: dyah.23002@mhs.unesa.ac.id

ARTICLE INFO

Article History:

Submitted/Received 17 December 2024
First Revised 31 December 2024
Accepted 31 December 2024
First Available Online 31 December 2024
Publication Date 31 December 2024

Keyword:

Principal Component Analysis, PCA, Energy efficiency.

ABSTRACT

Energy efficiency in residential buildings has become an important field of study given the increasing global energy demand and environmental concerns. Residential buildings account for the majority of global energy consumption, making it an important focus for strategies aimed at reducing carbon footprint and improving energy utilization. There are several factors that affect energy efficiency, including Relative Compactness, Surface Area, Wall Area, Roof Area, Overall Height, Orientation, Glazing Area, Glazing Area Distribution. This study aims to simplify and reduce these factors so as to obtain the dominant factors that affect the energy efficiency of residential buildings using the Principal Component Analysis (PCA) analysis method. Based on the description of the results of the study, 3 factors were obtained that most influenced, namely: The first factor (X8) is the most dominant factor with an eigenvalue of 1556.39648. The second factor (X7) consists of an eigenvalue of 99.2431641. The third factor (X6) with an eigenvalue of 0.8. The three factors (X8 to X6) can be assumed to be the most dominant factors that affect energy efficiency in the house. So that from the eight variables analyzed using Principal Component Analysis (PCA), 3 variables were obtained that became the dominant factors that affect the energy efficiency of residential houses, namely Glazing Area Distribution, Glazing Area and Orientation.

© 2024 Tim Pengembang Jurnal UNESA

1. INTRODUCTION

Currently, energy consumption is a very important topic around the world, according to *the World Energy Council*, global energy demand is expected to double by 2050 [1]. Energy efficiency in residential buildings has become an important area of study given the increasing global energy demand and environmental concerns [2]. Residential buildings account for the majority of global energy consumption, making them an important focus for strategies aimed at reducing carbon footprints and improving energy utilization [3]. Understanding the factors that affect energy efficiency is essential for designing effective interventions and policies to achieve sustainability goals [4]. The complexity of energy efficiency in housing lies in the many interrelated factors that affect it. These factors include building design, materials, insulation, HVAC systems, occupant behavior, and local climatic conditions [5]. Analyzing these variables comprehensively poses a challenge, as they often exhibit multicollinearity and high dimensions, making traditional statistical approaches less effective in capturing the impact of their combinations.[6]

Major Component Analysis (PCA) is a multivariate statistical technique that linearly transforms the shape of a group of original variables into a smaller, uncorrelated set of variables that can represent information from the original set of variables [7]. PCA will identify patterns in the dataset, find similarities and differences between each attribute as it serves as a powerful model for analyzing data [8]. A covariance matrix is calculated where the results are used in calculating the eigenvector and eigenvalue and the eigenvector with the highest eigenvalue is chosen as the main component of the established data because it shows the most significant relationship between the attributes of the dataset [9].

In the previous study researched by Wangge [10] by reducing the factors that affect the length of completion of the thesis of Mathematics Study Program students. The data used in the study was primary data from interviews with 50 students with 13 variables and the results of 10 selected variables were obtained. Khikmah [11] has also conducted research using the PCA method in determining the dominant weather factors for the spread of the Covid-19 outbreak in Surabaya. The results of the study are that there are 3 components that affect the spread of the virus, namely temperature, humidity, and duration of sunlight. Further research related to the PCA method was also carried [12] in reducing the variables that affect the improvement of rat kidney function from 8 variables to 3 variables.

Principal Component Analysis (PCA) has emerged as a powerful statistical tool to address these challenges. Converting a large set of correlated variables into a small number of principal components that are not correlated, PCA allows researchers to identify the most dominant factors influencing a particular outcome [13]. In the context of energy efficiency, PCA provides a systematic approach to reducing data complexity while retaining the most important information, thus facilitating better decision-making.

In addition to data with a large number of variables, this PCA method can also be used on data that has a large correlation between variables [14]. So it is hoped that the regression equation obtained can be free from the problem of correlation between variables without eliminating important information in the data used. Based on this, the purpose of writing this journal is to find out the stages in the application of the PCA method in dealing with violations of the multicollinearity assumption in multiple linear regression analysis without having to reduce the number of variables used in the research as well as knowing the dominant factors of residential energy efficiency [15].

This study applies PCA which aims to identify the dominant factors that affect energy efficiency in residential buildings.

DOI: http://dx.doi.org/10.xxxxx/jistel.vXiX
e- ISSN xxxx-xxxx

2. METHODS

2.1 Material

This type of research is quantitative research using the factor analysis method, namely Principal Component Analysis (PCA). The purpose of using this method is to simplify and eliminate less dominant screening factors or indicators without compromising the intent and purpose of the original data [16]. The analysis stage is carried out by collecting factors that affect energy efficiency in house buildings. There are 8 variables obtained from the UCI Repository Dataset used in this study, namely, Relative Compactness, Surface Area, Wall Area, Roof Area, Overall Height, Orientation, Glazing Area, Glazing Area Distribution [17][18].

The collected data was analyzed using the Principal Component Analysis (PCA) method. [10]. The PCA analysis process is carried out in several stages, namely correlation matrix testing, PCA analysis, and interpretation of PCA results [19].

2.1. Methods

A conceptual schematic illustrated of how PCA can help to simplify the dimensions of the data through the dataset hypothesis totaling m variables can be shown in **Figure 1** [20] :

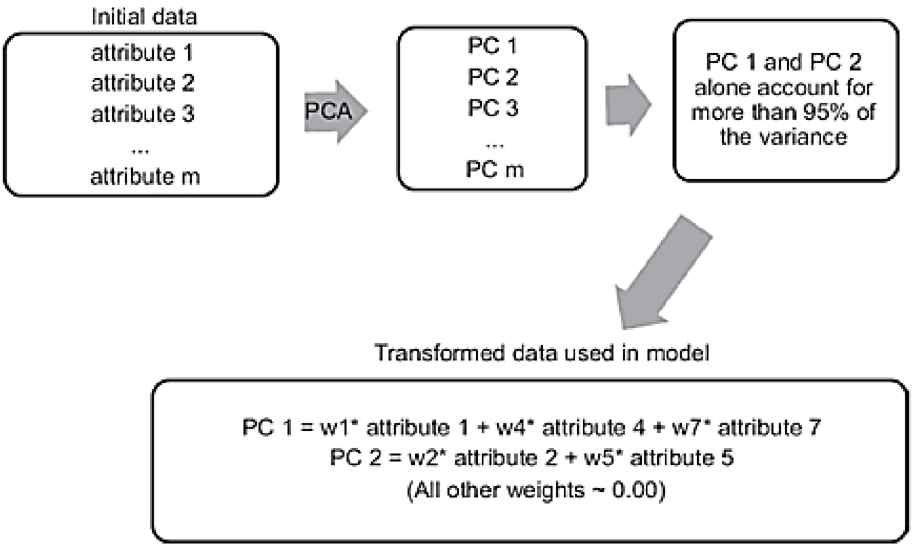

Figure 1. PCA conceptual model for the feature selection stage [20]

The following are the stages of factor analysis using PCA which can be described as follows:

1. Calculate the covariance matrix from the observation data.
   Variance (Var(x)) is calculated to find the data dispersion in the energy efficacy factors dataset to determine the data deviation in the sample dataset [8]. The Covariance Covariance Matrix (x,y) is a matrix in which the covariance values in each cell are obtained from a sample , suppose x and y are random variables.

DOI: http://dx.doi.org/10.xxxxx/jistel.vXiX
e- ISSN xxxx-xxxx

$$Var(x) = \sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(z_{ij} - \mu_j)^2 \quad \text{... (1)}$$

$$Cov(x.y) = \frac{1}{n-1}\sum_{i=1}^{n}(x_{ij} - \mu_{xj})(y_{xj} - \mu_{yj}) \quad \text{... (2)}$$

Where µx and µy are the mean of the samples from the x and y variables, where xi and yi are the ith observation values of the x and y variables. From the value data used, a covariance matrix measuring n x n was obtained.

2. Finding *the eigenvalue* and eigenvector of the covariance matrix that has been obtained, namely: The eigenvalue and eigenvector for the covariance matrix are calculated. The computed eigenvalues are then transformed (orthogonal varimax rotation) using the eq (3) [21].

$$Det\ (A - \lambda I) = 0 \quad \text{... (3)}$$

where:
A = matrik n x n
λ = exit of the eigenvalue
I = identity matrix (a square matrix with the main diagonal element is worth 1 while the other elements are worth 0

3. Conducting PCA factor analysis by identifying variables, including: X1 = Relative Compactness, X2 = Surface Area, X3 = Wall Area, X4 = Roof Area, X5 = Overall Height, X6 = Orientation, X7 = Glazing Area, X8 = Glazing Area Distribution

4. Make conclusions in the form of analysis results.

3. RESULTS AND DISCUSSION

The Z-score *process* is that each data on a variable is subtracted by the *mean* of the variable and divided by its standard deviation (in other words, each row per column minus the mean of the column, divided by the standard deviation of the same column) [22]. In this study, the attributes of the preprocessing results are represented in the form of labels (X1, X2, X3... X8) which represents the sequence of data that corresponds to the data being tested. The results of the variable representation in the form of labels: X1 = Relative Compactness; X2 = Surface Area; X3 = Wall Area; X4 = Roof Area; X5 = Overall Height; X6 = Orientation; X7 = Glazing Area; X8 = Glazing Area Distribution

The results of data normalization from raw data are explained in **Table 1**. **Table 1** contains from x1 to x8, where each variable is substracted by the average value of the data. This stage is the initial step in calculating the covariance of all data. The mean of the data for each variable is highlighted in yellow.

Table 1. Results of Normalized Data (Z-Score) of Energy Efficiency Factor

DOI: http://dx.doi.org/10.xxxxx/jistel.vXiX
e- ISSN xxxx-xxxx

No	X1	X1-(x̄)	X2	X2-(x̄)	X3	X3-(x̄)	X4	X4-(x̄)	X5	X5-(x̄)	X6	X6-(x̄)	X7	X7-(x̄)	X8	X8-(x̄)
1	0.98	0.22	514.50	-157.21	294.00	-24.50	110.25	-66.35	7.00	7.00	2	-1.50	0.00	-0.23	0	-2.81
2	0.98	0.22	514.50	-157.21	294.00	-24.50	110.25	-66.35	7.00	7.00	3	-0.50	0.00	-0.23	0	-2.81
3	0.98	0.22	514.50	-157.21	294.00	-24.50	110.25	-66.35	7.00	7.00	4	0.50	0.00	-0.23	0	-2.81
4	0.98	0.22	514.50	-157.21	294.00	-24.50	110.25	-66.35	7.00	7.00	5	1.50	0.00	-0.23	0	-2.81
5	0.90	0.14	563.50	-108.21	318.50	0.00	122.50	54.10	7.00	7.00	2	-1.50	0.00	-0.23	0	-2.81
6	0.90	0.14	563.50	-108.21	318.50	0.00	122.50	54.10	7.00	7.00	3	-0.50	0.00	-0.23	0	-2.81
7	0.90	0.14	563.50	-108.21	318.50	0.00	122.50	54.10	7.00	7.00	4	0.50	0.00	-0.23	0	-2.81
8	0.90	0.14	563.50	-108.21	318.50	0.00	122.50	54.10	7.00	7.00	5	1.50	0.00	-0.23	0	-2.81
9	0.86	0.10	588.00	-83.71	294.00	-24.50	147.00	29.60	7.00	7.00	2	-1.50	0.00	-0.23	0	- 2.81
10	0.86	0.10	588.00	-83.71	294.00	-24.50	147.00	29.60	7.00	7.00	3	-0.50	0.00	-0.23	0	-2.81
11	0.86	0.10	588.00	-83.71	294.00	-24.50	147.00	29.60	7.00	7.00	4	0.50	0.00	-0.23	0	-2.81
12	0.86	0.10	588.00	-83.71	294.00	-24.50	147.00	29.60	7.00	7.00	5	1.50	0.00	-0.23	0	-2.81
13	0.82	0.06	612.50	-59.21	318.50	0.00	147.00	29.60	7.00	7.00	2	-1.50	0.00	-0.23	0	-2.81
14	0.82	0.06	612.50	-59.21	318.50	0.00	147.00	29.60	7.00	7.00	3	-0.50	0.00	-0.23	0	-2.81
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
769	0.62	-0.14	808.50	136.79	367.50	49.00	220.50	43.90	3.50	3.50	5	1.50	0.40	0.17	5	2.19
Mean	0.76		671.71		318.50		176.60		5.25		3.50		0.23		2.81	

Covariance Matrix (Rapidminer) is explained in **Table 2**. **Table 2** shows that each variable has a different value. Based on this table, the highest total covariance value is in X2 and the lowest value is in variable X5. The results of this process are then used for the next process.

Table 2. Covariance Matrix (Rapidminer) Calculation Results

COVARIANS	X1	X2	X3	X4	X5	X6	X7	X8
X1	0.01117	-9.23003	-0.93917	-4.14543	0.15313	0.00000	0.00000	0.00000
X2	-9.23003	7749.06076	750.31250	3499.37413	-132.19792	0.00000	0.00000	0.00000
X3	-0.93917	750.31250	1900.79167	-575.23958	21.43750	0.00000	0.00000	0.00000
X4	-4.14543	3499.37413	-575.23958	2037.30686	-76.81771	0.00000	0.00000	0.00000
X5	0.15313	-132.19792	21.43750	-76.81771	3.06250	0.00000	0.00000	0.00000
X6	0.00000	0.00000	0.00000	0.00000	0.00000	1.25000	0.00000	0.00000
X7	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.01772	0.04395
X8	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.04395	2.40234
Total	-14.15	11857.32	2096.36	4880.48	-184.36	1.25	0.02	0.04

Table 3 is the results of calculation of priority matrix and eigenvalue. **Table 3** shows that each variable has a different value. Based on this table, the highest total covariance value is in X2 and the lowest value is in variable X5. The results of this process are then used for the next process.

DOI: http://dx.doi.org/10.xxxxx/jistel.vXiX
e- ISSN xxxx-xxxx

Table 3. Results of Calculation of Priority Matrix and Eigenvalue

	X1	**X2**	**X3**	**X4**	**X5**	**X6**	**X7**	**X8**	**Total**	**Priority**	**Eigen Value**
X1	-0.00079	-0.00078	-0.00045	-0.00085	-0.00083	0.00000	0.00000	7.22801E-17	-0.00370	0.000261208	-0.003696094
X2	0.65230	0.65353	0.35791	0.71701	0.71706	0.00000	0.00000	-1.2138E-12	3.09782	0.00026125	3.097815127
X3	0.06637	0.06328	0.90671	-0.11787	-0.11628	0.00000	0.00000	0	0.80222	0.00038267	0.802215128
X4	0.29296	0.29512	-0.27440	0.41744	0.41667	0.00000	0.00000	-1.8874E-13	1.14780	0.00023518	1.147799999
X5	-0.01082	-0.01115	0.01023	-0.01574	-0.01661	0.00000	0.00000	0	-0.04410	0.00023918	-0.044095858
X6	0.00000	0.00000	0.00000	0.00000	0.00000	1.00000	0.00000	0	1.00000	0.8	1
X7	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.88623	1.098632812	1.98486	99.2431640	1.984863281
X8	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	2.19727	60.05859375	62.25586	1556.39648	62.25585938
TOTAL	1.00002	1.00000	1.00000	1.00000	1.00001	1.00000	3.08350	61.15723	70.24076	1656.44103	70.24076

The final results of the factors that influence energy efficiency on a house scale are explained in **Table 4**, and depicted in **Figure 2**. Based on **Table 4**, it shows that the most influential factors are glazing area distribution, glazing area, and orientation.

Table 4. Factors Affecting Energy Efficiency (PCA Final Results)

No	Variabel	Label	Eigenvalue
1	X8	Glazing Area Distribution	1556.39648
2	X7	Glazing Area	99.2431641
3	X6	Orientation	0.8
4	X3	Wall Area	0.00038267
5	X5	Overall Height	0.00023918
6	X4	Roof Area	0.00023518
7	X2	Surface Area	0.00026126
8	X1	Relative Compactness	0.00026121

The main purpose of this analysis is to explain as many variants of the initial data as possible with a few key components called factors. Where the factors formed are a linear combination of the variables studied and can explain the diversity of data to the maximum so that it is easier to analyze the data [23]. Based on the description of the research results, calculation data for each PCA analysis parameter was obtained. Eight variables already observed and 3 factors were obtained that most influenced, namely: the first factor (X8) is the most dominant factor with an eigenvalue of 1556.396, the second factor (X7) consists of an eigenvalue of 99.243, and he third factor (X6) with an eigenvalue of 0.8. The three factors (X8 to X6) can be assumed to be the most dominant factors that affect energy efficiency in the house. So that from the eight variables analyzed using PCA, three variables were obtained that became the dominant factors that affect the energy efficiency of residential houses, namely glazing area distribution, glazing area and orientation.

DOI: http://dx.doi.org/10.xxxxx/jistel.vXiX
e- ISSN xxxx-xxxx

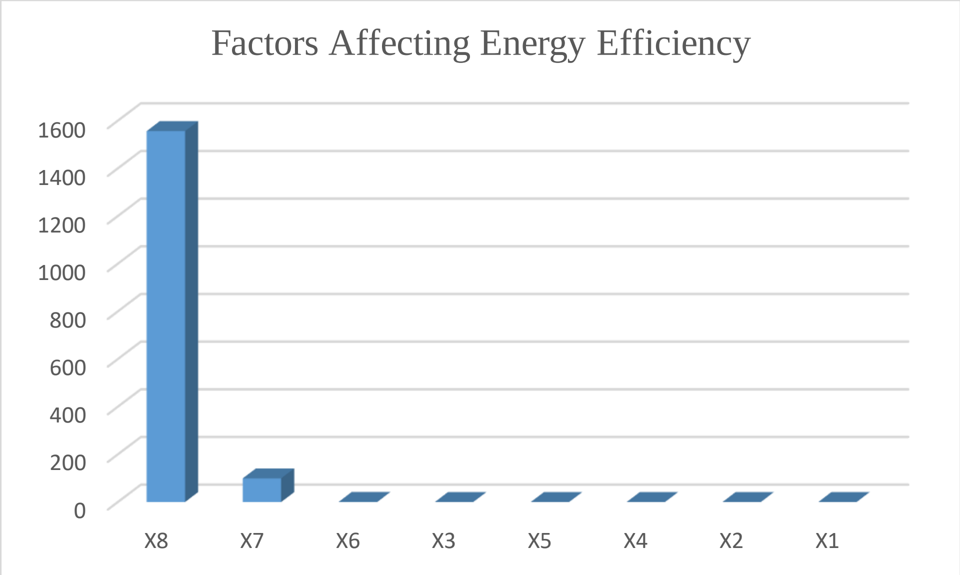

Figure 2. Ranking results of factors influencing energy efficiency using PCA

There is one of the studies by Halida [24] using the PCA method in her research in determining the factors of poverty cases. In their research, they used a method to reduce 4 correlated variables to 2 new factors that did not correlate. In addition, the research conducted by Badri & Sari, [25] using the same method, namely PCA in his research in determining the factors that affect students' attitudes to choose the mathematics study program of UIN Malang. In his research, Badri used the PCA method to reduce 20 variables that correlate into 5 factors that affect students' attitudes to continue their studies at PTAI.

So from this study and some of the results of previous studies, it can be interpreted that the PCA method can be used to simplify and eliminate less dominant screening factors or indicators without reducing the purpose and purpose of the original data and determining the most dominant factors of certain variables or indicators.

4. CONCLUSION

The results of our test use factor analysis (PCA) by reducing 8 variables that correlate there are 3 factors that are energy efficiency of residential houses. The first factor (X8) is the most dominant factor having an eigenvalue of 1556.396 . The second factor (X7) consists of an eigenvalue of 99.243. The third factor (X6) with an eigenvalue of 0.8. The three factors (X8 to X6) can be assumed to be the most dominant factors that affect energy efficiency in the house. The three factors are Glazing Area Distribution, Glazing Area and Orientation. So these three factors are assumed to be the dominant factors that affect the energy efficiency of residential houses.

5. ACKNOWLEDGMENT

We would like to thank UCI Repository which has provided data in this study.

6. AUTHORS' NOTE

The authors declare that there is no conflict of interest regarding the publication of this article. The authors confirmed that the paper was free of plagiarism.

7. AUTHORS' CONSTRIBUTION/ROLE

Dyah Pranesti Shafira Fitri: Conceptualization, data curation, formal analysis, funding acquisituion, investigation, mrthodology; Mhd. Ali Hanafiah: project administration, resources, software, supervision, validation, visualization, writing-original draft, writing-review and editing.

8. REFERENCES

[1] N. Mendes, E. and Mendes, "'An instructional design for building energy simulation elearning: an interdisciplinary approach," *J. Build. Perform. Simul.*, pp. 1–17, 2019, doi: doi: 1.

[2] G. Di Foggia, "Energy efficiency measures in buildings for achieving sustainable development goals," *Heliyon*, vol. 4, no. 11, 2018, doi: doi.org/10.1016/j.heliyon.2018.e00953.

[3] A. E. Fenner, "The carbon footprint of buildings: A review of methodologies and applications," *Renew. Sustain. Energy Rev.*, vol. 94, pp. 1142–1152, 2018, doi: https://doi.org/10.1016/j.rser.2018.07.012.

[4] F. Belaid, M. E. Sanin, and N. Lazaric, "Energy efficiency and residential energy consumption: current challenges, policies, and emerging trends," *Front. Energy Res.*, vol. 11, no. September, pp. 1–3, 2023, doi: 10.3389/fenrg.2023.1218527.

[5] V. A. Arowoiya, A. O. Onososen, R. C. Moehler, and Y. Fang, "Influence of Thermal Comfort on Energy Consumption for Building Occupants: The Current State of the Art," *Buildings*, vol. 14, no. 5, 2024, doi: 10.3390/buildings14051310.

[6] T. Rahmawati, "Penerapan Metode Elman Recurrent Neural Network Dan Principal Component Analysis (PCA) Untuk Peramalan Konsumsi Listrik the," *Teknomatika*, vol. 53, no. 9, pp. 1689–1699, 2013.

[7] S. Manullang, D. Aryani, and H. Rusydah, "Analisis Principal Component Analysis (PCA) dalam Penentuan Faktor Kepuasan Pengunjung terhadap Layanan Perpustakaan Digilib," *Edumatic J. Pendidik. Inform.*, vol. 7, no. 1, pp. 123–130, 2023, doi: 10.29408/edumatic.v7i1.14839.

[8] I. . Jolliffe, *Principal Component Analysis. 2nd Edition.* New York: Springer-Verlag, 2002.

[9] D. . Johnson, W.A. & Wichern, *Applied Multivariate Statistical Analysis. 6th Edition*. New Jersey: Pearson Prentice Hall, 2007.

[10] M. Wangge, "Penerapan Metode Principal Component Analysis (PCA) Terhadap Faktor-faktor yang Mempengaruhi Lamanya Penyelesaian Skripsi Mahasiswa Program Studi Pendidikan Matematika FKIP UNDANA," *J. Cendekia J. Pendidik. Mat.*, vol. 5, no. 2, pp. 974–988, 2021, doi: 10.31004/cendekia.v5i2.465.

[11] K. N. Khikmah, "Penerapan Principal Component Analysis dalam Penentuan Faktor Dominan Cuaca Terhadap Penyebaran Covid-19 di Surabaya," *ESTIMASI J. Stat. Its Appl*, vol. 2, no. 1, pp. 11–18, 2021, doi: doi: 10.20956/ejsa.v2i1.11943, 2021.

[12] F. F. and S. Sugiyarto, "Implementasi Analisa Komponen Utama untuk Mereduksi Variabel yang Mempengaruhi Perbaikan pada Fungsi Ginjal Tikus," *" J. Ilm. Mat*, vol. 6, no. 2, p. 62, 2019, doi: doi: 10.26555/konvergensi.v6i2.19549, 2019.

[13] M. Ahsan, "Evaluasi Performa dari Diagram Kontrol Multivariat berbasis," vol. 1, no. December, pp. 89–92, 2018.

[14] M. Z. Nasution, "Face Recognition based Feature Extraction using Principal Component Analysis (PCA)," *J. Informatics Telecommun. Eng.*, vol. 3, no. 2, pp. 182–191, 2020, doi: 10.31289/jite.v3i2.3132.

[15] D. R. P. Sari, "Metode Principal Component Analysis (PCA) Sebagai Penanganan

Asumsi Multikolinearitas (Studi Kasus: Data Produksi Tapioka)," vol. 02, no. 02, pp. 115–124, 2023.
[16] M. Z. Nasution, "Penerapan Principal Component Analysis (PCA) Dalam Penentuan Faktor Dominan Yang Mempengaruhi Prestasi Belajar Siswa (Studi Kasus : SMK Raksana 2 Medan)," *J. Teknol. Inf.*, vol. 3, no. 1, p. 41, 2019, doi: 10.36294/jurti.v3i1.686.
[17] A. T. and A. Xifara, "Energy Efficiency," UCI Machine Learning Repository.
[18] A. Tsanas, A., & Xifara, "Accurate quantitative estimation of energy performance of residential buildings using statistical machine learning tools.," *Energy Build.*, vol. 49, pp. 560–567, 2012, doi: DOI:10.1016/J.ENBUILD.2012.03.003.
[19] Ilmaniati & Putro, "Analisis Komponen Utama Faktor-Faktor Pendahulu Mikro , Kecil , Dan Menengah (Umkm) Di Indonesia," *J. Teknol.*, vol. 11, no. 1, pp. 67–78, 2019, [Online]. Available: https://jurnal.umj.ac.id/index.php/jurtek/article/view/2652%0Ahttps://jurnal.umj.ac.id/index.php/jurtek/article/download/2652/2912
[20] B. Kotu, V. & Deshpande, *Predictive Analytics and Data Mining Concepts and Practice with Rapid Miner*. USA: Elsevier, 2015.
[21] M. Z. Nasution, A. A. Nababan, K. U. Syaliman, M. S. Novelan, and M. Jannah, "Penerapan Principal Component Analysis (PCA) Dalam Penentuan Faktor Dominan Yang Mempengaruhi Pengidap Kanker Serviks (Studi Kasus : Cervical Cancer Dataset)," *J. Mantik Penusa*, vol. 3, no. 1, pp. 204–210, 2019.
[22] A. P. Sari, "Teori Principal Components Analysis (PCA)," 2019.
[23] A. Maćkiewicz and W. Ratajczak, "Principal components analysis (PCA)," *Comput. Geosci.*, vol. 19, no. 3, pp. 303–342, 1993, doi: https://doi.org/10.1016/0098-3004(93)90090-R.
[24] A. Halida, N. F. Pradita, Y. Sukmawaty, and N. Farah Pradita, "Dengan Pendekatan Principal Component Analysis," *J. Al Qardh*, vol. 5, no. 2, pp. 134–141, 2020, [Online]. Available: http://e-journal.iain-palangkaraya.ac.id/index.php/qardh
[25] F. Badri and S. U. R. Sari, "Penerapan Metode Principal Component Analysis (PCA) Untuk Identifikasi Faktor-Faktor yang Mempengaruhi Sikap Mahasiswa Memilih Melanjutkan Studi ke Kota Malang," *Build. Informatics, Technol. Sci.*, vol. 3, no. 3, pp. 426–431, 2021, doi: 10.47065/bits.v3i3.1139.