BAREKENG: Journal of Mathematics and Its Applications
September 2026 Volume 20 Issue 3 Page 2507-2522
P-ISSN: 1978-7227 E-ISSN: 2615-3017
https://doi.org/10.30598/barekengvol20iss3pp2507-2522

EFFECTIVENESS OF DIMENSIONALITY REDUCTION METHODS ON DATA WITH NON-LINEAR RELATIONSHIPS

Lukmanul Hakim[1], Asep Saefuddin[2*], Kusman Sadik[3], Anwar Fitrianto[4], Bagus Sartono[5]

[1,2,3,4,5]*Statistics and Data Science, School of Data Science, Mathematics, and Informatics, IPB University*
Jln. Meranti Kampus IPB, Babakan, Kec. Dramaga, Bogor, 16680, Indonesia

[1]*Data Science Study Program, Universitas Insan Cita Indonesia*
Jln. Asem Baris Raya No.1, Jakarta, 12750, Indonesia

*Corresponding author's e-mail: * asaefuddin@apps.ipb.ac.id*

Article Info

Article History:

Received: 3rd September 2025
Revised: 27th November 2025
Accepted: 17th March 2026
Published: 8th April 2026

Keywords:

Autoencoders;
Missing value;
Neural Network;
Non Linier;
Outlier;
Principal Component Analysis.

ABSTRACT

The phenomenon of big data poses distinct challenges for analysis, especially when the data contains a very large number of variables. High complexity, potential redundancy, and the risk of overfitting are major issues that must be addressed through dimensionality reduction techniques. Principal Component Analysis (PCA) is a common method effective for data with linear relationships, but it has limitations in identifying nonlinear patterns. This research aims to improve classification performance by introducing an autoencoder to handle nonlinear relationships, data noise, missing values, outliers, and data with varying scales. This study employs a quantitative approach by analyzing simulated and empirical data in the form of the Village Development Index from the Central Statistics Agency, which contains variables with various measurement scales. Both dimensionality reduction methods—PCA and neural network-based autoencoders—are tested across various data scenarios. The evaluation is based on their effectiveness in preserving the data structure and on the Mean Squared Error (MSE) values during the reconstruction process. The results indicate that PCA excels in computational efficiency and accuracy for data with linear relationships. In contrast, the autoencoder demonstrates superior performance in detecting nonlinear patterns, achieving lower Mean Squared Error (MSE) values with stable MSE standard deviations. Additionally, the autoencoder proves more robust to missing values and outliers than PCA. The selection of dimensionality reduction methods highly depends on the characteristics of the analyzed data. Autoencoders are a superior alternative for handling complex, nonlinear data, though they require parameter tuning. Further research is recommended to explore the influence of autoencoder network architecture and training strategies on dimensionality reduction performance.

How to cite this article:

L. Hakim, A. Saefuddin, K. Sadik, A. Fitrianto and B. Sartono, "EFFECTIVENESS OF DIMENSIONALITY REDUCTION METHODS ON DATA WITH NON-LINEAR RELATIONSHIP *BAREKENG: J. Math. & App.,* vol. 20, no. 3, pp. 2507-2522, Sep, 2026.

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

The phenomenon of big data has become a hot topic among researchers and academics. Big data refers to a collection of data that is very large and varied, posing challenges in storage, analysis, and visualization [1]. The complexity of the data itself naturally makes analysis difficult. One common problem encountered in big data is the presence of a very large number of variables. Data sets with more variables generally contain more information, but as the number of variables increases, the likelihood of noise and redundancy also rises [2]. Additionally, this leads to longer computational processes, and some statistical methods are unable to handle such complexity [3]. High data complexity often leads to overfitting [4]. Although the amount of data collected is very large, it does not necessarily mean that the data is evenly distributed in a high-dimensional space [5]. Furthermore, strong correlations among variables can lead some variables to disproportionately influence prediction results, reducing the model's robustness and interpretability [6].

Dimensionality reduction is the process of transforming high-dimensional data into a lower number of dimensions while preserving its essential information [7]. Techniques for dimensionality reduction provide a foundation for effectively analyzing and interpreting high-dimensional datasets [8]. This process enables data representation with fewer features by employing unsupervised methods that uncover relationships among variables and eliminate redundancy. Typically, dimensionality reduction is categorized into two primary methods: feature selection (FS) and feature extraction (FE). FS is regarded as an effective approach for mitigating dimensionality challenges by reducing redundancy, removing superfluous data, and enhancing the interpretability of results. In contrast, FE focuses on identifying a representative and informative feature set, thereby improving processing efficiency [9], [10], [11], [12].

The complexity of data with a large number of variables makes dimensionality reduction an important step in data preprocessing to improve accuracy and reduce computation time [13]. Some advantages of dimensionality reduction include reducing data storage memory, enabling quick data visualization, and decreasing complexity in data analysis [14]. One of the popular and widely used methods for dimensionality reduction is Principal Component Analysis (PCA) [15]. This method was first introduced by Pearson in 1901 and developed by Hotelling in 1933 [16], [17]. In simple terms, PCA is a way to transform data into a new set of coordinates so that the first coordinate captures the most variation, the second the next most, and so on. PCA is an unsupervised learning method that decomposes data into linear combinations [19]. Additionally, PCA assumes that the relationships between variables are linear and that all variables must be measured at the interval or ratio level. Therefore, PCA may not always be the most appropriate analysis method [20]. Some other popular dimensionality reduction methods besides PCA include Linear Discriminant Analysis (LDA) [21], Singular Value Decomposition (SVD) [22], and others. Designed to capture linear patterns, these methods tend to perform poorly on data with non-linear relationships. Similarly, PCA can only capture linear patterns, and its performance will decline if the data exhibit curvature/non-linear patterns [23]. The same thing happens with LDA [24].

Given the phenomenon of big data, methods capable of capturing nonlinear patterns are needed. One such method is autoencoders (AE). This method is a type of neural network that projects high-dimensional data into a low-dimensional latent space [25]. In AE, the input layer and output layer have the same number of nodes [26]. Several applications of the AE method include research by [27], which compared PCA, LDA, LLE, Isomap, and AE methods on the MNIST dataset. The study results indicated that AE performed better than the other methods because it not only reduces dimensionality but also detects repetitive structures. A similar study compared PCA and AE on the MNIST data and found that AE yielded a lower mean squared error (MSE) than PCA [28].

However, most of these studies remain limited to image data such as MNIST, which has very specific characteristics and does not reflect the complexity of other data types in real-world practice, including non-image data, data with nonlinear relationships, or data containing noise. The novelty of this research lies in extending the evaluation of PCA and autoencoders beyond commonly used image datasets, incorporating diverse nonlinear relationships, multiple levels of data disturbances, and empirical data with varied measurement scales, thereby providing a more realistic and comprehensive benchmark for dimensionality reduction in practical applications. Therefore, this study aims to examine and compare the performance of AE and PCA across a broader range of data characteristics, both simulated and empirical. Specifically, the objectives of this research include (1) evaluating the performance of AE and PCA on linearly correlated data with variations in the number of observations and components; (2) comparing the performance of AE and PCA on data with nonlinear relationships (e.g., sinus, cosine, tangent, exponential, logarithm, square root,

Arccos, and arcsin); (3) analyzing the robustness of AE and PCA against data disturbances such as missing values and outliers at various levels (5%, 10%, and 15%); and (4) comparing their performance on empirical data in the form of the Village Development Index from the Central Statistics Agency, with various measurement scales.

2. RESEARCH METHODS

This study uses simulated data with diverse characteristics to assess the effectiveness of AE and PCA methods for identifying patterns. Additionally, the study incorporates empirical data to evaluate the performance of both methods in real-world scenarios. The next section will outline the data simulation design implemented in this study.

2.1 Principal Component Analysis (PCA)

PCA is a classical statistical method utilized for dimensionality reduction [29]. It can be mathematically defined as a technique that transforms data into a new orthogonal coordinate system. In this system, the first coordinate captures the greatest variance of the data, the second coordinate captures the second greatest variance, and so forth [18]. Furthermore, PCA allows for the representation of original variables as integrated factors that reflect the information contained within them [30]. This methodology involves a series of steps that convert the original variables into new orthogonal variables known as principal components, specifically designed to maximize the variance explained by the original data. The stages of the PCA process are as follows [31]:

1. Calculating the variance matrix of observation data

$$Var(x) = \sigma^2 = \frac{1}{n}\Sigma_{1=1}^{n}\left(z_{ij} - \mu_j\right)^2, \tag{1}$$

$$Cov(x,y) = \frac{1}{n-1}\left(x_{ij} - \mu_{xj}\right)\left(y_{ij} - \mu_{yj}\right), \tag{2}$$

with μ_x and μ_y is the sample average of variables x and y, where x_i and y_i are the values of the i observation of variables x and y.

2. Find the eigenvalues and eigenvectors of the covariance matrix that has been obtained.

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

3. Determining the PCA proportion value

$$PC(\%) = \frac{Eigen\ Value}{Variance\ Covariance} x\ 100\%. \tag{4}$$

4. Calculating loading factor [32]

$$Ax = \lambda x \tag{5}$$

Obtained linear combination, namely:

a. $\lambda_1, \lambda_2, \lambda_3, \ldots, \lambda_n$ are eigenvalues of matrix A

b. $x_1, x_2, x_3, \ldots, x_n$ are eigenvectors

The equation of eigen value decomposition is:

$$AX = XD, \tag{6}$$

$$CA = X\ D\ X^{-1}, \tag{7}$$

where:

A : n x n matrix that has n eigen values,

D : eigen value of its eigen vector,

X : eigen vector of matrix A,

X^{-1} : inverse of eigen vector x.

2.2 Autoencoders (AE)

AE is a classical unsupervised neural network architecture that uses the backpropagation algorithm for learning and is commonly applied in dimensionality reduction tasks [33], [34]. This method was first proposed by Le Cun in 1987, with early works on AE focusing on dimensionality reduction or feature learning. AE consists of three main stages: encoding, activation, and decoding. The encoder is the initial layer of the autoencoder that receives input from the original data [35]. The activation function performs nonlinear mapping to transform the encoded coefficients into a specific range [36]. The decoder, which serves as the bottleneck layer, reconstructs the encoded representation back to the original input dimensions [37]. As illustrated in Fig. 1, an AE consists of an encoder network, a bottleneck layer representing the low-dimensional feature space, where σ denotes the activation function, and W and b represent the weights known as latent hidden vectors, followed by a decoder network. For more details, see Fig. 1 below.

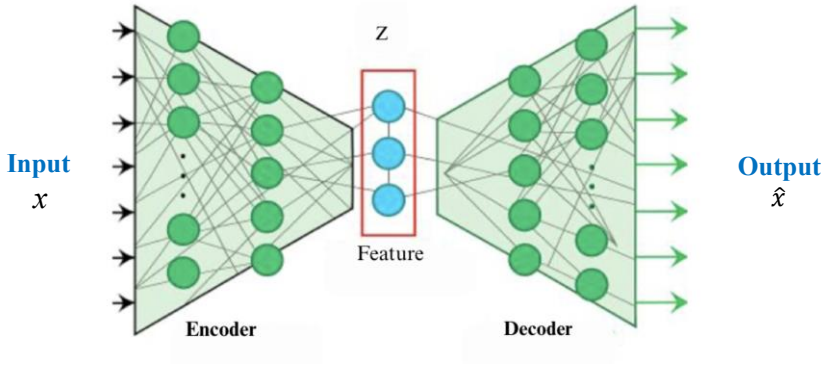


Figure 1. Autoencoders Illustration

The stages of the autoencoder (AE) in performing dimension reduction are outlined as follows [43]:

1. Data input preparation

   Input data is in the form of a vector $x \in R^d$ with dimention d

2. Encoding (Encoder Function f_θ)

   Transforming input data $x \in R^d$ into a latent representation $z \in R^m$, where $m < d$

   $$z = f_\theta(x).$$

3. Decoding (Decoder Function $g_\emptyset$)

   Transforming the latent representation $z \in R^m$ back to the original dimension for output reconstruction $\hat{x} \in R^d$

   $$\hat{x} = g_\emptyset(z). \quad (8)$$

4. Calculating reconstruction loss

   $$\mathcal{L}(x, \hat{x}) = ||x - \hat{x}||^2. \quad (9)$$

2.3 Simulation Approach

At this stage, we generate data based on two specific criteria: linear and nonlinear relationships, with varying numbers of observations and components. For more details, please refer to Table 1 below.

Table 1. Information On Linear and Nonlinear Data Simulation

Step	Relationship	Correlations	Observations	Component
Step 1	Linear	50 Variables with Correlation 70%-100%	500	10
				20
				30
			1000	10
				20
				30
			10000	10
				20
				30

Step 2	Non-Linear	10 Sinus Variables 10 Cosine Variables 10 Tangent Variables 10 Exponential Variables	500 1000 10000 	10 20 30 10 20 30 10 20 30
Step 3	Non-Linear	10 Sinus Variables 10 Cosine Variables 10 Tangent Variables 10 Exponential Variables 10 Logarithmic Variables 10 Square Root Variables 5 Arccos Variables 5 Archin Variables	500 1000 10000 	10 20 30 10 20 30 10 20 30
Step 4	Non-Linear (Missing 5%, 10%, 15%)	10 Sinus Variables 10 Cosine Variables 10 Tangent Variables 10 Exponential Variables	1000	30
Step 5	Non-Linear (Outlier 5%, 10%, 15%)	10 Sinus Variables 10 Cosine Variables 10 Tangent Variables 10 Exponential Variables	1000	30

Table 1 outlines the data generation design implemented in Python to evaluate the performance of the AE and PCA methods. This evaluation considers variations in relationship types, correlation levels, the number of observations, and the number of components. For datasets exhibiting nonlinear relationships, the base data are generated from a uniform distribution ranging from -3 to 3. These data are subsequently transformed into various nonlinear functions, including sine, cosine, tangent, exponential, logarithmic, square root, arccosine, and arcsine. Additionally, disturbances such as noise, missing values, and outliers are introduced at different levels (5%, 10%, and 15%). This process aims to assess the robustness of both methods.

2.4 Empirical Approach

The empirical data were obtained from the Central Statistics Agency and consist of the Village Development Index, which includes 74 variables and 1,199 observations, measured on various scales, including nominal and ratio scales. These variables represent multiple socio-economic and infrastructure characteristics across villages, which naturally lead to heterogeneous relationships, including both linear and nonlinear patterns among variables. This data is used to evaluate the performance of AE and PCA in handling real-world data that are heterogeneous in terms of variable relationships and measurement scales. In particular, the analysis compares how well AE and PCA preserve data structure when reducing dimensionality under these heterogeneous conditions. Table 2 provides information related to the empirical data used.

Table 2. Measurement Scale

Measurement Scale	Number of Variables
Nominal	63
Ratio	11

As shown in Table 2, the total number of variables is 74, comprising 63 nominal variables and 11 ratio variables. The dataset is sourced from the Village Development Index (Indeks Desa Membangun) published by the Central Statistics Agency (BPS), which describes socio-economic, infrastructure, and environmental characteristics at the village level. The 1,199 observations correspond to villages within the selected study area, allowing this dataset to reflect real-world conditions with heterogeneous measurement scales and variable relationships. This makes the dataset suitable for evaluating the performance of dimensionality reduction methods such as AE and PCA.

2.5 Evaluation Models

Model performance evaluation is essential to measure the effectiveness of both methods in retaining important information while reducing data complexity. For this purpose, two commonly used statistical metrics are Mean Squared Error (MSE) and Coefficient of Determination (R^2).

2.5.1 Mean Squared Error (MSE)

MSE is one method used to measure a model's goodness. It measures the average of the squared error between the model's predicted value and the actual value [44]. The smaller the MSE value in a model, the better the model is. In the context of dimension reduction, the formula for MSE is as follows:

$$MSE = \frac{\sum(x_i - \hat{x_i}\ )^2}{n}. \tag{10}$$

2.5.2 Coefficient of Determination (R^2)

R-Squared, also called as co-efficient of determination, is one of the metrics used to measure the goodness of a model. The greater the R^2, the greater the proportion of variation that can be explained by the model. If R^2 approaches 1, it means that the model has a very good fit and if it approaches 0 the model cannot explain the variation of the data well or is not fit [45]. The formula for the coefficient of determination from the application of dimension reduction is as follows:

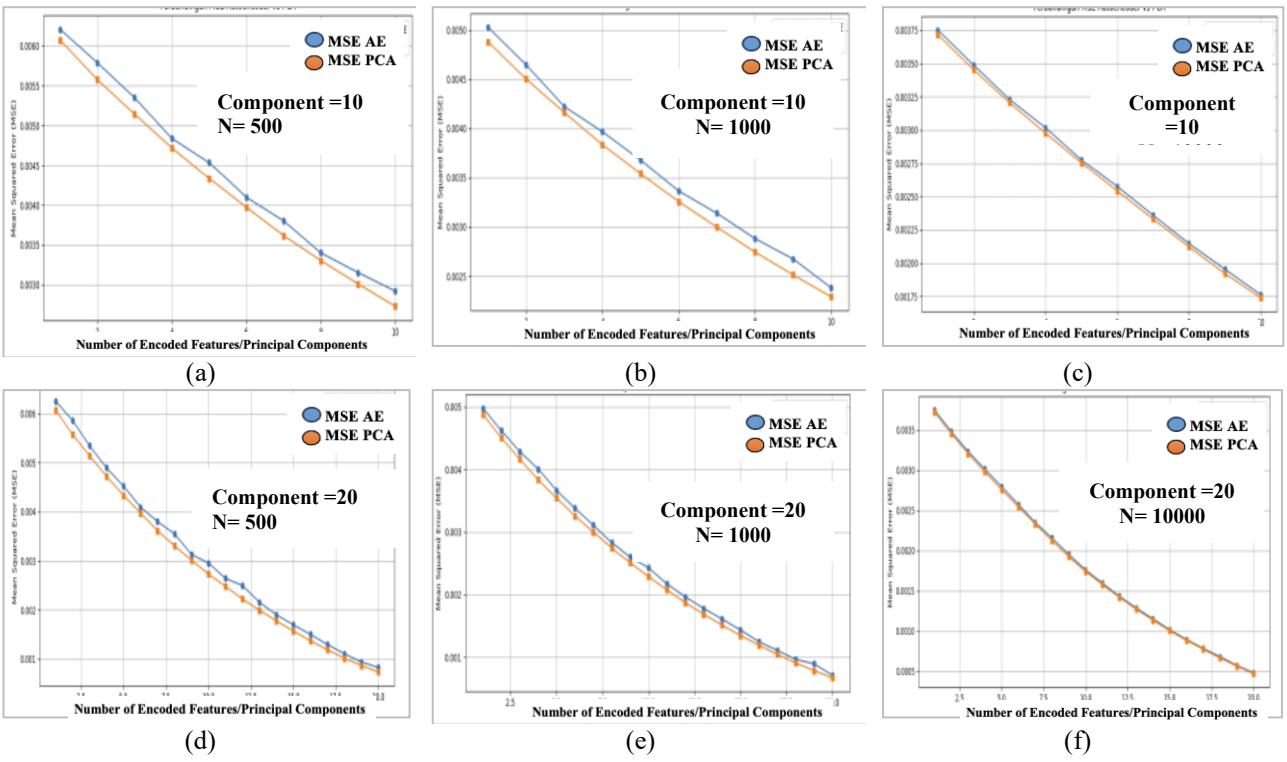
$$R^2 = 1 - \frac{\sum_{i=1}^{n}\sum_{j=1}^{d}(x_{ij} - \hat{x}_{ij})^2}{\sum_{i=1}^{n}\sum_{j=1}^{d}(x_{ij} - \bar{x}_j)^2}. \tag{11}$$

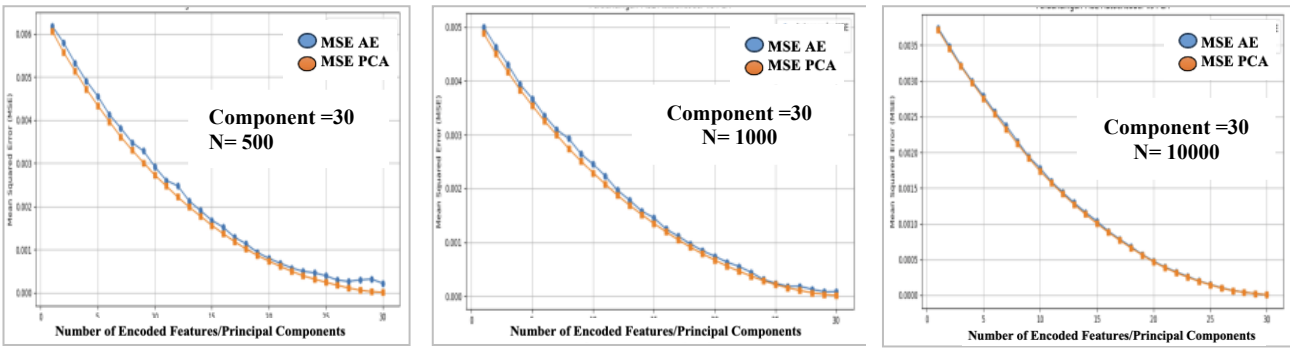
3. RESULTS AND DISCUSSION

Based on the analysis conducted using two different approaches, simulated and empirical, interesting findings emerged regarding the performance of the two evaluated methods.

3.1 Linear Relationship

Based on the simulation conducted on data with linear relationships, the results are presented in Fig. 2 below.





(g) (h) (i)

Figure 2. Comparison of PCA and AE on data with linear relationships for various numbers of components and sample sizes: (a) Component = 10, N = 500; (b) Component = 10, N = 1000; (c) Component = 10, N = 10000; (d) Component = 20, N = 500; (e) Component = 20, N = 1000; (f) Component = 20, N = 10000; (g) Component = 30, N = 500; (h) Component = 30, N = 1000; and (i) Component = 30, N = 10000.

Fig. 2 compares MSE values for the AE and PCA methods on data with linear correlations. The MSE decreases as the number of components increases, indicating that more information from the original data is retained. Overall, the MSE values for AE and PCA are similar, as evidenced by the overlapping MSE curves for both methods. This similarity is particularly apparent with a large number of observations, such as 10,000. However, the MSE values tend to diverge when the sample size is only 500. In this case, PCA performs slightly better than AE for smaller datasets. Nevertheless, the difference is not particularly significant when looking at the actual MSE values produced by both methods. PCA shows a slight advantage in producing lower errors for small datasets, while AE demonstrates greater stability on datasets with many observations.

The overall MSE values for AE and PCA are similar, particularly for data with linear correlations and large sample sizes. However, there are notable differences between the two methods. Despite its simplicity, PCA has an advantage when dealing with datasets with fewer observations due to its greater stability and faster computational speed. On the other hand, AE requires a more complex training process and is sensitive to network architecture parameters. This sensitivity can be a limitation for AE, especially if proper parameter tuning is not implemented. To clarify the comparison between these two methods, the output is presented in Table 3 below.

Table 3. Comparison of Average MSE And Standard Deviation of MSE

Method	Sample	Component 10		Component 20		Component 30	
		Average	Std. Deviation	Average	Std. Deviation	Average	Std. Deviation
PCA	500	0.0042	0.0011	0.0029	0.0016	0.0020	0.0018
	1000	0.0035	0.0008	0.0024	0.0013	0.0017	0.0015
	10000	0.0027	0.0006	0.0018	0.0010	0.0013	0.0011
AE	500	0.0044	0.0011	0.0031	0.0017	0.0022	0.0018
	1000	0.0036	0.0008	0.0025	0.0013	0.0018	0.0015
	10000	0.0027	0.0006	0.0018	0.0010	0.0013	0.0011

Based on Table 3, which presents the average MSE and the average standard deviation of MSE between the PCA and AE methods across various sample sizes and component numbers, it is evident that PCA and AE yield similar results. Although the average and standard deviation of the MSE for PCA are slightly lower than those for AE, the differences are not particularly significant. However, these minor differences can influence method selection depending on computational speed and performance stability, with PCA showing greater stability, especially on smaller datasets. Overall, both methods provide important insights: as the sample size and number of components increase, their performance becomes more stable on data structures with linear relationships.

According to the analysis presented, the performance of AE and PCA depends heavily on the structure of the relationships among variables. Generally, for data with linear relationships, the performance of AE and PCA appears comparable, which aligns with previous studies stating that AE tends to be on par with PCA in terms of reconstruction and dimensionality reduction [20], [38]. However, when the number of observations is limited, PCA tends to outperform AE due to its greater stability and computational efficiency. Conversely, AE requires a complex training process and is sensitive to parameter settings, such as the number of hidden neurons, the number of epochs, and the activation functions.

3.2 Nonlinear Relationship Data

When the data structure among variables is nonlinear, the differences between AE and PCA become more apparent. In this condition, PCA, being based on linear transformations, faces limitations in capturing complex patterns. This scenario is illustrated in Fig. 3 below.

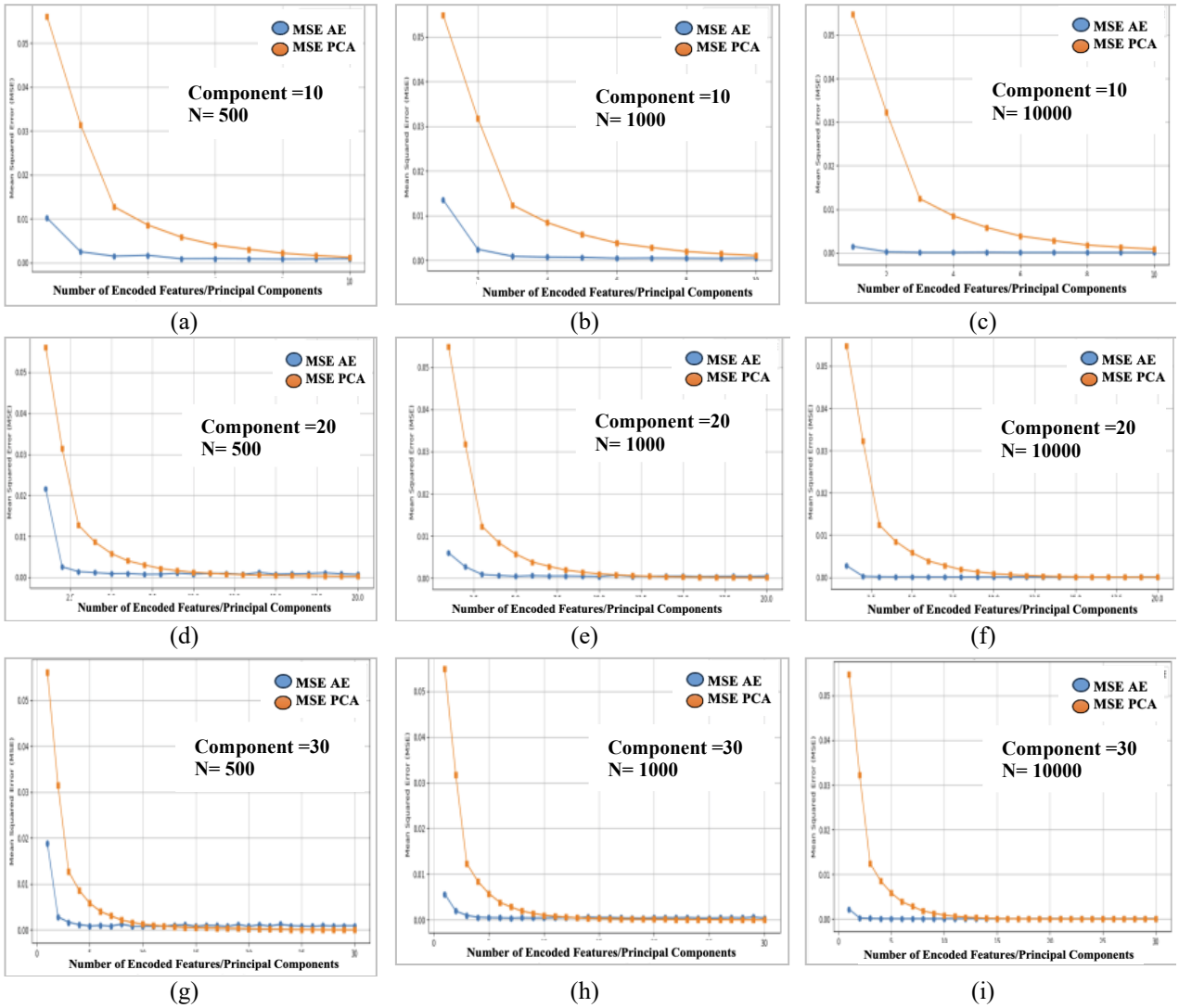


Figure 3. Comparison of PCA and AE on nonlinear data (sin, cosin, tangent, and exponential) for various numbers of components and sample sizes: (a) Component = 10, N = 500; (b) Component = 10, N = 1000; (c) Component = 10, N = 10000; (d) Component = 20, N = 500; (e) Component = 20, N = 1000; (f) Component = 20, N = 10000; (g) Component = 30, N = 500; (h) Component = 30, N = 1000; and (i) Component = 30, N = 10000.

In this condition, PCA based on linear transformation experiences significant limitations, especially in capturing complex patterns. Fig. 3 shows that PCA exhibits higher, more unstable MSE values across varying numbers of components and observations, indicating suboptimal performance on nonlinear data. On the other hand, AE appears to represent non-linear data patterns well. The MSE value in AE tends to be stable even in the fourth dimension, especially when the data is constructed using non-linear functions such as sine, cosine, tangent, and exponential. The superiority of AE becomes more apparent as the number of components increases, due to its ability to capture the complex structure of non-linear data. The difference is seen in Table 4, which presents the average MSE values and the average standard deviation of MSE from both methods.

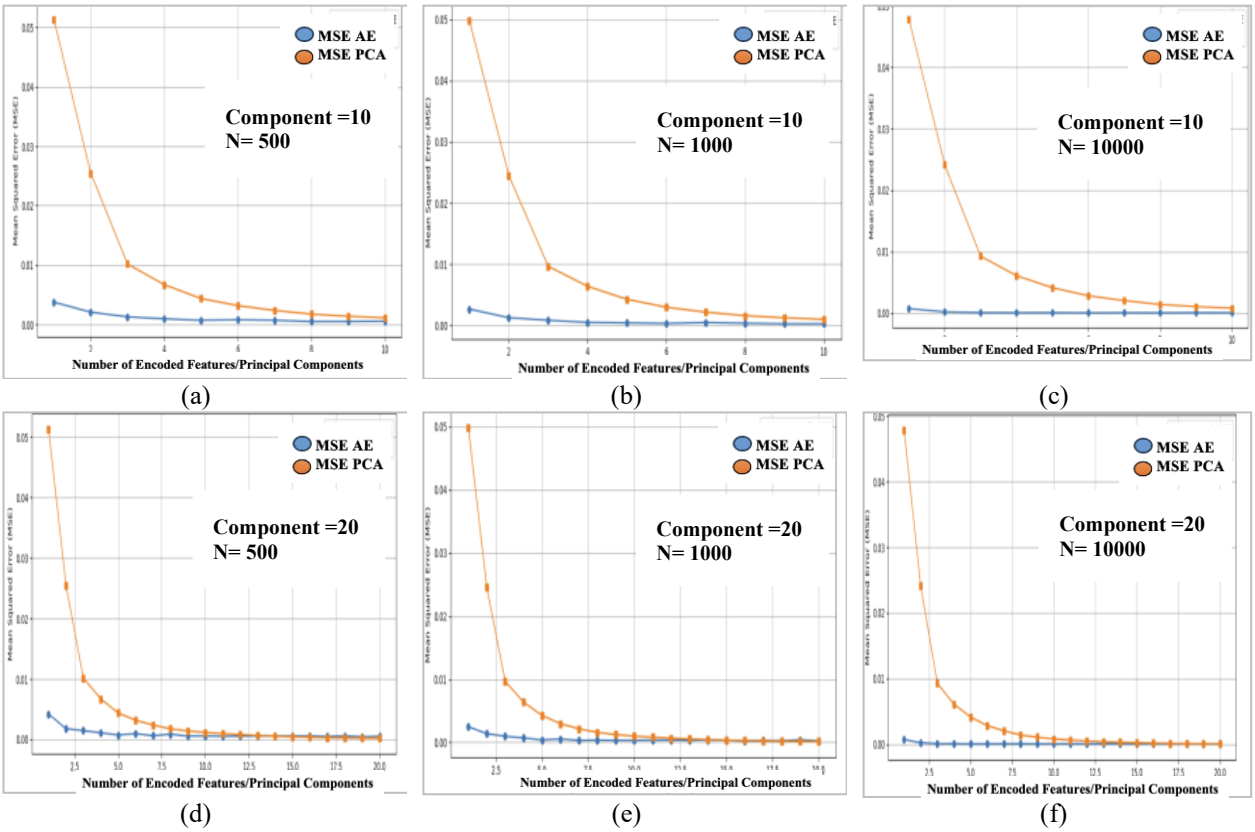
Table 4. Comparison of Average MSE And Average Standard Deviation of MSE For Nonlinear Relationship Data (Sinus, Cosine, Tangent, And Exponential)

Method	Sample	Component 10		Component 20		Component 30	
		Average	Std. Deviation	Average	Std. Deviation	Average	Std. Deviation
PCA	500	0.0100	0.0143	0.0066	0.0134	0.0044	0.0113

Method	Sample	Component 10		Component 20		Component 30	
		Average	Std. Deviation	Average	Std. Deviation	Average	Std. Deviation
AE	1000	0.0124	0.0166	0.0064	0.0132	0.0043	0.0112
	10000	0.0124	0.0167	0.0063	0.0133	0.0042	0.0113
	500	**0.0001**	**0.0002**	**0.0019**	**0.0036**	**0.0014**	**0.0027**
	1000	**0.0011**	**0.0013**	**0.0007**	**0.0008**	**0.0007**	**0.0007**
	10000	**0.0002**	**0.0004**	**0.0002**	**0.0006**	**0.0001**	**0.0002**

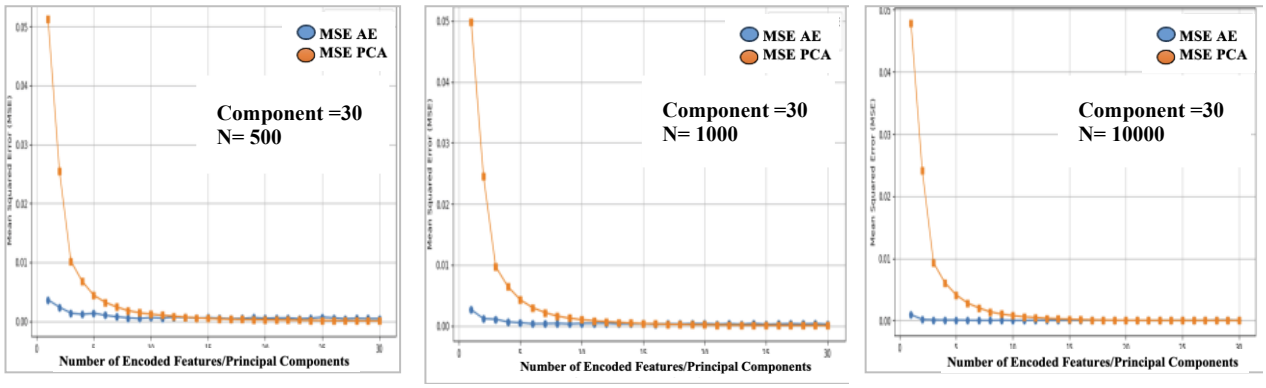
The output in Table 4 provides clear information regarding AE's ability to capture data patterns with nonlinear relationships. Overall, the average MSE and the average standard deviation of MSE from AE are significantly lower compared to PCA. This study finds that AE performs better in the presence of nonlinear relationships. The use of nonlinear activation functions in the model enables AE to capture complex patterns in the data and enhances the model's stability. Furthermore, as the number of samples increases, the average MSE consistently decreases, indicating improved performance.

However, the advantages of AE are not without drawbacks; the training process of AE requires tuning parameters such as the number of hidden neurons, the number of epochs, and the activation functions, which, if not optimally set, can lead to overfitting or underfitting. Additionally, the computation time of AE tends to be higher than that of PCA. On the other hand, although PCA appears less stable for non-linear data, it still offers advantages in terms of computational speed and ease of implementation, especially if only a rough estimate of the data structure is needed. Advanced simulations can incorporate other non-linear functions, such as logarithms, square roots, archin, and Arccos, to further test the robustness of AE. The study aims to determine the limits of AE's ability to handle increasingly complex patterns and to compare it with PCA as a simple baseline.



(a) (b) (c)

(d) (e) (f)



(g) (h) (i)

Figure 4. Comparison of PCA and AE on nonlinear data (sinus, Cosine, tangent, exponential, logarithm, square root, archin, and Arccos) for various numbers of components and sample sizes: (a) Component = 10, N = 500; (b) Component = 10, N = 1000; (c) Component = 10, N = 10000; (d) Component = 20, N = 500; (e) Component = 20, N = 1000; (f) Component = 20, N = 10000; (g) Component = 30, N = 500; (h) Component = 30, N = 1000; and (i) Component = 30, N = 10000.

The addition of variables with nonlinear relationships to the dataset further strengthens AE's consistent performance. AE can capture nonlinear data patterns, as reflected in significantly lower, more stable MSE values compared to PCA. Furthermore, as the number of variables increases, the MSE values for AE continue to decrease, whereas in PCA, increasing the number of nonlinear variables leads to greater inconsistencies. As in the previous results, AE successfully captures the presence of nonlinear variable groups in the data. The stabilization of the MSE values for AE occurs at the seventh dimension, while for PCA, it stabilizes at the twelfth dimension. At this point, the MSE values of the two methods become comparable. The level of MSE stability for the two methods is presented in Table 5 below.

Table 5. Comparison of Average MSE And Average Standard Deviation of MSE For Nonlinear Relationship Data (Sinus, Cosine, Tangent, Exponential, Logarithm, Square Root, Arcsin, And Arccos)

Method	Sample	Component 10		Component 20		Component 30	
		Average	Std. Deviation	Average	Std. Deviation	Average	Std. Deviation
PCA	500	0.0108	0.0152	0.0056	0.0119	0.0038	0.0101
	1000	0.0104	0.0148	0.0054	0.0116	0.0036	0.0098
	10000	0.0100	0.0143	0.00509	0.0112	0.0034	0.0094
AE	500	**0.0013**	**0.0010**	**0.0008**	**0.0007**	**0.0007**	**0.0006**
	1000	**0.0007**	**0.0006**	**0.0005**	**0.0006**	**0.0004**	**0.0005**
	10000	**0.0001**	**0.0002**	**0.0001**	**0.0002**	**0.0001**	**0.0001**

Overall, the results in Table 5 demonstrate that the AE method consistently produces a lower average MSE and a lower average standard deviation of MSE than PCA. The inclusion of four additional nonlinear transformation functions: logarithm, square root, arsinh, and arcosh, does not degrade AE's performance. On the contrary, AE has been shown to be effective at capturing complex nonlinear patterns. Furthermore, increasing the number of components and sample sizes tends to reduce MSE values in both methods; however, the reduction is significantly more pronounced in the AE method. Beyond testing on variables with nonlinear relationships, AE was also evaluated on datasets with missing values across variables, further validating its robustness under real-world data conditions.

The performance differences between the two methods become more pronounced when applied to data with nonlinear relationships. Under such conditions, PCA has limitations in capturing data patterns characterized by nonlinear relationships, as it only performs linear transformations. In contrast, AE, designed with nonlinear activation functions, is more flexible and can effectively represent complex data structures. Across varying numbers of observations and components, the MSE values for AE tend to be lower and more stable than those of PCA. When data are grouped by the number of functions, such as sine, tangent, cosine, exponential, logarithm, square root, arccos, and arcsin. AE successfully captures these patterns, as evidenced by stable MSE values after surpassing dimensions corresponding to the number of specified functions. AE's ability to handle data structures with nonlinear relationships is consistent with previous findings, though the training process of AE requires careful parameter tuning to avoid overfitting or underfitting [39], [40].

3.3 Data Noise (Missing Value and Outlier)

The introduction of noise into the data does not negatively affect the AE's performance in dimensionality reduction. The AE demonstrates greater stability than PCA, even under varying levels of noise, as shown in Fig. 5.

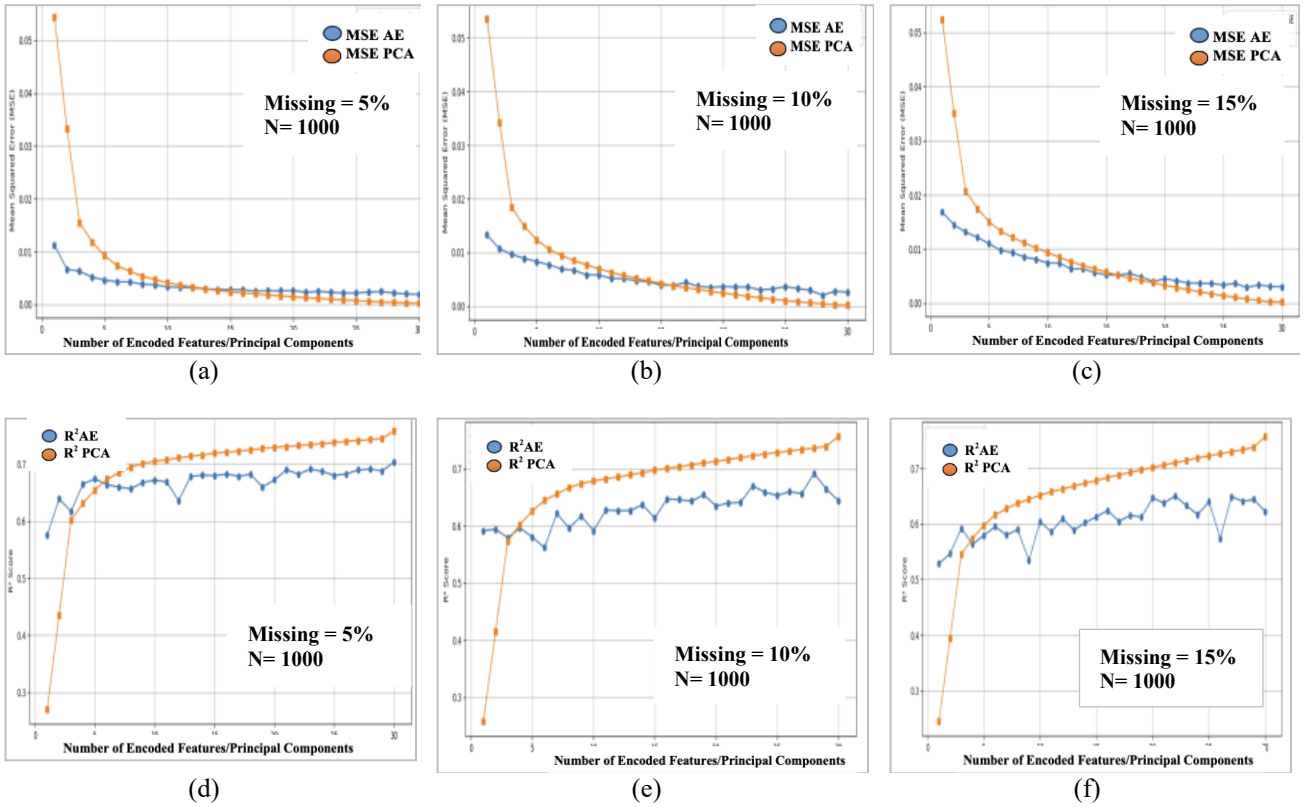
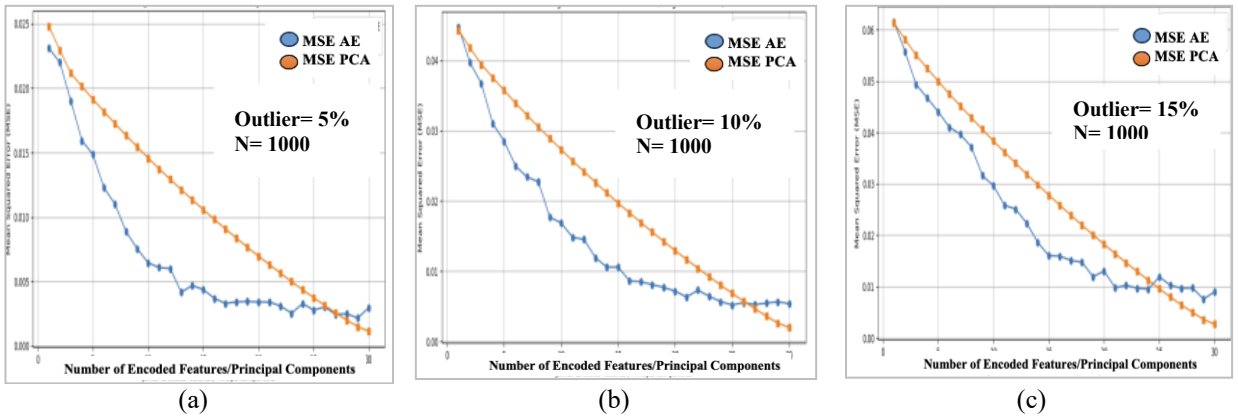
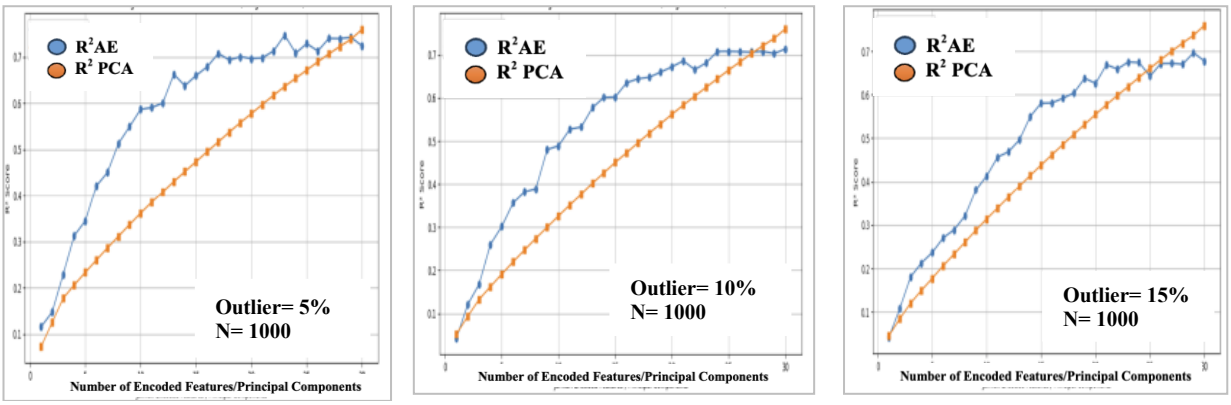


Figure 5. Comparison of PCA and AE on nonlinear data containing missing values for various numbers of missing: (a) Missing=5%, N = 1000; (b) Missing=10%, N = 1000; (c) Missing=15%, N = 1000; (d) R^2 AE and PCA with Missing=5%, N = 1000; (e) R^2AE and PCA with Missing=10%, N = 1000; (f) R^2 AE and PCA with Missing = 15%, N = 1000

Based on the visualization in Fig. 5, the AE method consistently outperforms PCA on datasets with missing values. The variations in the proportion of missing values do not affect AE's ability to produce low MSE values and consistently high R² values. In the first dimension, AE captures over 60% of the data's variance, while PCA captures approximately 20%. However, as the number of dimensions increases, PCA shows a significant improvement in its ability to capture data variance. In contrast, AE maintains a stable performance within the range of 55% to 70%, demonstrating its effectiveness, particularly with data that exhibit nonlinear patterns and missing values. Both methods show a decline in R^2 at the first dimension when the proportion of missing values per variable reaches 15%, though the decline is not substantial. We will apply both methods to data that contains outliers in each variable.





(d) (e) (f)

Figure 6. Comparison of PCA and AE on nonlinear data containing outlier values for various numbers of outliers: (a) Outlier=5%, N = 1000; (b) Outlier =10%, N = 1000; (c) Outlier =15%, N = 1000; (d) R^2 AE and PCA with Outlier =5%, N = 1000; (e) R^2AE and PCA with Outlier =10%, N = 1000; (f) R^2 AE and PCA with Outlier =15%, N = 1000

The results demonstrate that the AE consistently achieves lower MSE values than PCA, indicating its superior ability to reconstruct data despite the presence of outliers. PCA's linear framework and sensitivity to extreme deviations limit its effectiveness in handling outliers. AE's robustness is further supported by higher R^2 values relative to PCA, where R^2 quantifies the closeness of the reconstructed data to the original, with higher values reflecting better performance. This evidence confirms AE's resilience against outliers. Beyond the 25th principal component, the performance gap between AE and PCA narrows, suggesting both methods become comparably effective in representing the data. The average MSE and standard deviation of MSE on noisy datasets further elucidate the comparative performance.

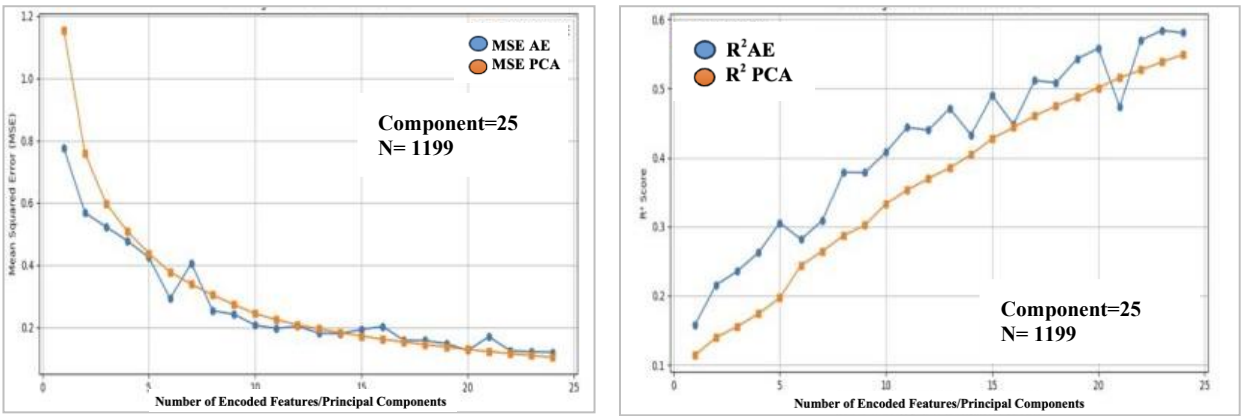
Table 6. Comparison of Average MSE And Average Standard Deviation of MSE For Nonlinear Relationship Data with Addition of Missing Values and Outliers

Method	Sample	Missing 5% Outlier 5%		Missing 10% Outlier 10%		Missing 15% Outlier 15%	
		Average	Std. Deviation	Average	Std. Deviation	Average	Std. Deviation
PCA	1000	0.0060	0.0110	0.0076	0.0109	0.0088	0.0109
		0.0109	0.0067	0.0202	0.0124	0.0284	0.0174
AE	1000	**0.0033**	**0.0015**	**0.0051**	**0.0026**	**0.0066**	**0.0036**
		0.0069	**0.0059**	**0.0150**	**0.0112**	**0.0237**	**0.0157**

Despite the addition of noise such as missing values and outliers at various levels (5%, 10%, and 15%), the AE method continues to demonstrate more consistent performance, evidenced by lower average MSE and average standard deviation of MSE compared to PCA. This evidence confirms that AE is more reliable in accurately reconstructing complex data than PCA. When data are subjected to disturbances such as outliers and missing values, AE exhibits superior stability. These findings reinforce the literature indicating AE's robustness in recognizing hidden patterns even when portions of the data are missing [44], [43]. AE consistently demonstrates lower MSE and higher R^2 values, whereas PCA is more vulnerable to data containing outliers or missing values. Subsequently, both methods were tested on empirical data with varying measurement scales, including nominal, interval, and ratio.

3.4 Empirical Approach

Further testing was carried out to assess the performance of both methods in representing complex empirical data across different measurement scales. Fig. 7 depicts the behavior of AE and PCA based on the values of Mean Squared Error (MSE) and R^2.



(a) (b)

Figure 7. Comparison of PCA with AE on Empirical Data: (a) Component 25, N=1199; (b) R^2AE

Scale variation in the data does not significantly impact the performance degradation of the AE compared to PCA. This advantage is evidenced by AE's lower MSE and higher R^2 values, as depicted in Fig. 7. The average and standard deviation of MSE for AE are 0.2412 and 0.1564, respectively, compared to 0.2553 and 0.2333 for PCA. These results substantiate AE's superior ability to capture latent representations of nonlinear data, particularly in complex datasets with significant scale variability. Although the average MSE values are comparable across methods, AE shows lower variability, as reflected in its lower standard deviation. AE's adaptability to diverse input types and measurement scales constitutes a primary advantage. In contrast, PCA's strengths include simplicity, computational efficiency, and effective handling of linear numerical data with well-defined structures. PCA remains pertinent for practical applications necessitating rapid interpretation and lightweight dimensionality reduction. Nonetheless, PCA's dependence on linearity assumptions and ideal data distributions limits its effectiveness when confronted with complex, noisy, or nonlinear data. While AE offers higher accuracy and greater flexibility, it requires more computational resources, and its encoded representations are more difficult to interpret directly. Therefore, the choice of method should be aligned with the data context and analytical needs: AE is preferable for in-depth exploration of complex data, whereas PCA is suitable for rapid modeling and clear interpretation of linear and well-structured data.

AE's ability to adapt to variations in scale and input types makes it a flexible tool for latent data representation. Although PCA remains advantageous for computational efficiency and ease of interpretation, it relies heavily on assumptions of ideal data distribution and linear relationships among variables. Therefore, method selection should be aligned with the data characteristics and analytical objectives. AE's superior performance when handling complex data with nonlinear inter-variable relationships compensates for its longer computational time [42].

4. CONCLUSION

The AE method is a dimensionality reduction technique that can serve as an alternative to PCA. AE can provide results comparable to PCA in data with linear relationships between variables. However, when the data exhibit strong linear relationships, PCA is still recommended due to its computational efficiency over AE. Nevertheless, PCA has limitations in capturing nonlinear data patterns. In this study, the observed data exhibited nonlinear relationships, rendering PCA less effective at preserving the underlying structure. This limitation becomes a challenge in analyzing complex data that does not meet the linearity assumption. Based on simulation studies and empirical research, AE has shown better performance than PCA in detecting nonlinear patterns. This advantage is evidenced by the lower MSE, indicating higher data reconstruction accuracy in the AE model. In this study, computational efficiency refers to the execution time required in the dimensionality reduction process, while accuracy is evaluated using the Mean Squared Error (MSE) obtained during reconstruction.

The main advantage of AE lies in its ability to map data into a latent space while preserving important nonlinear information. However, this method also presents several challenges, including higher computational requirements and dependence on training parameters, such as the number of epochs and the number of hidden layers. The determination of these parameters significantly affects the model's

performance: too few epochs can cause the model to fail to learn data patterns optimally, while too many epochs risk overfitting. Therefore, a trial-and-error approach is often necessary to achieve the optimal configuration. Furthermore, studies have demonstrated the robustness of AE in handling missing values, outliers, and varying measurement scales. The flexibility and stability of AE in handling diverse data conditions demonstrate its outstanding potential as an adaptive and reliable dimensionality reduction method across modern data analysis scenarios.

Author Contributions

Lukmanul Hakim: Conceptualization, Data Curation, Formal Analysis, Methodology, Software, Visualization, Writing, Original Draft. Asep Saefuddin: Conceptualization, Supervision, Validation, Writing, Review and Editing. Kusman Sadik: Methodology, Supervision, Validation, Writing, Review and Editing. Anwar Fitrianto: Formal Analysis, Methodology, Visualization, Writing, Review and Editing. Bagus Sartono: Methodology, Supervision, Visualization, Writing, Review and Editing. All authors discussed the results and contributed to the final manuscript.

Funding Statement

This research was funded by the Indonesian Education Scholarship through the Center for Higher Education Funding and Assessment, Ministry of Higher Education, Science, and Technology, Republic of Indonesia, and by the Endowment Fund for Education Agency, Ministry of Finance, Republic of Indonesia.

Acknowledgment

The author would like to express sincere gratitude to Universitas Insan Cita Indonesia for granting permission to pursue a doctoral program, and to the Indonesian Education Scholarship (for providing a full doctoral scholarship through the Endowment Fund for Education Agency (LPDP). Appreciation is also extended to colleagues and academic mentors for their guidance and valuable input throughout this research.

Declarations

The authors declare no conflicts of interest to report study.

Declaration of Generative AI and AI-assisted technologies

The authors declare that no generative AI or AI-assisted technologies were used in the preparation of this manuscript, including for writing, editing, data analysis, or the creation of tables and figures.

REFERENCES

[1] H. K. Yaseen and A. M. Obaid, "BIG DATA: DEFINITION, ARCHITECTURE & APPLICATIONS," *JOIV Int. J. Informatics Vis.*, vol. 4, no. 1, pp. 45–51, Feb. 2020, doi: https://doi.org/10.30630/joiv.4.1.292

[2] M. Ashraf *et al.*, "A SURVEY ON DIMENSIONALITY REDUCTION TECHNIQUES FOR TIME-SERIES DATA," *IEEE Access*, vol. 11, pp. 42909–42923, 2023, doi: https://doi.org/10.1109/ACCESS.2023.3269693

[3] R. Ramachandran, G. Ravichandran, and A. Raveendran, "EVALUATION OF DIMENSIONALITY REDUCTION TECHNIQUES FOR BIG DATA," in *2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC)*, IEEE, Mar. 2020, pp. 226–231, doi: https://doi.org/10.1109/ICCMC48092.2020.ICCMC-00043

[4] C. JI *et al.*, "BIG DATA PROCESSING: BIG CHALLENGES AND OPPORTUNITIES," *J. Interconnect. Networks*, vol. 13, no. 03n04, p. 1250009, Sep. 2012, doi: https://doi.org/10.1142/S0219265912500090

[5] L. H. Nguyen and S. Holmes, "TEN QUICK TIPS FOR EFFECTIVE DIMENSIONALITY REDUCTION," *PLOS Comput. Biol.*, vol. 15, no. 6, p. e1006907, Jun. 2019, doi: https://doi.org/10.1371/journal.pcbi.1006907

[6] W. Zhang, H. Cheng, S. Zhan, M. Luo, F. Wang, and Z. Huang, "DIMENSIONALITY REDUCTION AND MACHINE LEARNING BASED MODEL OF SOFTWARE COST ESTIMATION," *Front. Phys.*, vol. 12, Mar. 2024, doi: https://doi.org/10.3389/fphy.2024.1324719

[7] S. Nanga *et al.*, "REVIEW OF DIMENSION REDUCTION METHODS," *J. Data Anal. Inf. Process.*, vol. 09, no. 03, pp. 189–231, 2021, doi: https://doi.org/10.4236/jdaip.2021.93013

[8] N. Ahmad and A. B. Nassif, "DIMENSIONALITY REDUCTION: CHALLENGES AND SOLUTIONS," *ITM Web Conf.*, vol. 43, p. 01017, Mar. 2022, doi: https://doi.org/10.1051/itmconf/20224301017

[9] M. DPatil and S. S. Sane, "DIMENSION REDUCTION: A REVIEW," *Int. J. Comput. Appl.*, vol. 92, no. 16, pp. 23–29, Apr. 2014, doi: https://doi.org/10.5120/16094-5390

[10] N. Abd-Alsabour, "ON THE ROLE OF DIMENSIONALITY REDUCTION," *J. Comput.*, pp. 571–579, 2018, doi: https://doi.org/10.17706/jcp.13.5.571-579
[11] D. M. Abdulqader, A. M. Abdulazeez, and D. Q. Zeebaree, "MACHINE LEARNING SUPERVISED ALGORITHMS OF GENE SELECTION: A REVIEW," *Technol. Reports Kansai Univ*, vol. 62, no. 03, pp. 233–244, 2020.
[12] B. M. Salih Hasan and A. M. Abdulazeez, "A REVIEW OF PRINCIPAL COMPONENT ANALYSIS ALGORITHM FOR DIMENSIONALITY REDUCTION," *J. Soft Comput. Data Min.*, vol. 02, no. 01, Apr. 2021, doi: https://doi.org/10.30880/jscdm.2021.02.01.003
[13] S. Velliangiri, S. Alagumuthukrishnan, and S. I. Thankumar joseph, "A REVIEW OF DIMENSIONALITY REDUCTION TECHNIQUES FOR EFFICIENT COMPUTATION," *Procedia Comput. Sci.*, vol. 165, pp. 104–111, 2019, doi: https://doi.org/10.1016/j.procs.2020.01.079
[14] H. K. Palo, S. Sahoo, and A. K. Subudhi, "DIMENSIONALITY REDUCTION TECHNIQUES: PRINCIPLES, BENEFITS, AND LIMITATIONS," in *Data Analytics in Bioinformatics*, Wiley, 2021, pp. 77–107, doi: https://doi.org/10.1002/9781119785620.ch4
[15] J. T. Vogelstein *et al.*, "SUPERVISED DIMENSIONALITY REDUCTION FOR BIG DATA," *Nat. Commun.*, vol. 12, no. 1, p. 2872, May 2021, doi: https://doi.org/10.1038/s41467-021-23102-2
[16] K. Pearson, "LIII. *ON LINES AND PLANES OF CLOSEST FIT TO SYSTEMS OF POINTS IN SPACE*," *London, Edinburgh, Dublin Philos. Mag. J. Sci.*, vol. 2, no. 11, pp. 559–572, Nov. 1901, doi: https://doi.org/10.1080/14786440109462720
[17] H. Hotelling, "ANALYSIS OF A COMPLEX OF STATISTICAL VARIABLES INTO PRINCIPAL COMPONENTS.," *J. Educ. Psychol.*, vol. 24, no. 6, pp. 417–441, Sep. 1933, doi: https://doi.org/10.1037/h0071325
[18] J. Liu, Lai Xu, A. Caprihana, and V. D. Calhoun, "EXTRACTING PRINCIPLE COMPONENTS FOR DISCRIMINANT ANALYSIS OF FMRI IMAGES," in *2008 IEEE International Conference on Acoustics, Speech and Signal Processing*, IEEE, Mar. 2008, pp. 449–452, doi: https://doi.org/10.1109/ICASSP.2008.4517643
[19] N. Salem and S. Hussein, "DATA DIMENSIONAL REDUCTION AND PRINCIPAL COMPONENTS ANALYSIS," *Procedia Comput. Sci.*, vol. 163, pp. 292–299, 2019, doi: https://doi.org/10.1016/j.procs.2019.12.111
[20] Monica Casella, Pasquale Dolce, Michela Ponticorvo, and Davide Marocco, "AUTOENCODERS AS AN ALTERNATIVE APPROACH TO PRINCIPAL COMPONENT ANALYSIS FOR DIMENSIONALITY REDUCTION. AN APPLICATION ON SIMULATED DATA FROM PSYCHOMETRIC MODELS," *Proccedings Third Symp. Psychol. Technol.*, 2021. doi: https://doi.org/10.1109/MetroXRAINE54828.2022.9967686
[21] P. Switzer, "EXTENSIONS OF LINEAR DISCRIMINANT ANALYSIS FOR STATISTICAL CLASSIFICATION OF REMOTELY SENSED SATELLITE IMAGERY," *J. Int. Assoc. Math. Geol.*, vol. 12, no. 4, pp. 367–376, Aug. 1980, doi: https://doi.org/10.1007/BF01029421.
[22] K. V. Ravi Kanth, D. Agrawal, A. El Abbadi, and A. Singh, "DIMENSIONALITY REDUCTION FOR SIMILARITY SEARCHING IN DYNAMIC DATABASES," *Comput. Vis. Image Underst.*, vol. 75, no. 1–2, pp. 59–72, Jul. 1999, doi: https://doi.org/10.1006/cviu.1999.0762.
[23] M. Al Digeil *et al.*, "PCA-ENHANCED AUTOENCODERS FOR NONLINEAR DIMENSIONALITY REDUCTION IN LOW DATA REGIMES," *Proc. Can. Conf. Artif. Intell.*, Jun. 2023, doi: https://doi.org/10.21428/594757db.05a13011.
[24] L. Qu and Y. Pei, "A COMPREHENSIVE REVIEW ON DISCRIMINANT ANALYSIS FOR ADDRESSING CHALLENGES OF CLASS-LEVEL LIMITATIONS, SMALL SAMPLE SIZE, AND ROBUSTNESS," *Processes*, vol. 12, no. 7, p. 1382, Jul. 2024, doi: https://doi.org/10.3390/pr12071382.
[25] W. Wang, Y. Huang, Y. Wang, and L. Wang, "GENERALIZED AUTOENCODER: A NEURAL NETWORK FRAMEWORK FOR DIMENSIONALITY REDUCTION," in *2014 IEEE Conference on Computer Vision and Pattern Recognition Workshops*, IEEE, Jun. 2014, pp. 496–503, doi: https://doi.org/10.1109/CVPRW.2014.79.
[26] C. Hu, X. Hou, and Y. Lu, "IMPROVING THE ARCHITECTURE OF AN AUTOENCODER FOR DIMENSION REDUCTION," in *2014 IEEE 11th Intl Conf on Ubiquitous Intelligence and Computing and 2014 IEEE 11th Intl Conf on Autonomic and Trusted Computing and 2014 IEEE 14th Intl Conf on Scalable Computing and Communications and Its Associated Workshops*, IEEE, Dec. 2014, pp. 855–858, doi: https://doi.org/10.1109/UIC-ATC-ScalCom.2014.50.
[27] Y. Wang, H. Yao, S. Zhao, and Y. Zheng, "DIMENSIONALITY REDUCTION STRATEGY BASED ON AUTO-ENCODER," in *Proceedings of the 7th International Conference on Internet Multimedia Computing and Service*, New York, NY, USA: ACM, Aug. 2015, pp. 1–4, doi:https://doi.org/10.1145/2808492.2808555.
[28] J. Wang, H. He, and D. V. Prokhorov, "A FOLDED NEURAL NETWORK AUTOENCODER FOR DIMENSIONALITY REDUCTION," *Procedia Comput. Sci.*, vol. 13, pp. 120–127, 2012, doi: https://doi.org/10.1016/j.procs.2012.09.120.
[29] I.T. Jolliffe, *Principal Component Analysis*, Second Edition. New York: Springer-Verlag, , 1986. doi: https://doi.org/10.1007/978-1-4757-1904-8.
[30] H. Abdi and L. J. Williams, "PRINCIPAL COMPONENT ANALYSIS," *WIREs Comput. Stat.*, vol. 2, no. 4, pp. 433–459, Jul. 2010, doi: https://doi.org/10.1002/wics.101.
[31] 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, Jul. 2019, doi: https://doi.org/10.36294/jurti.v3i1.686.
[32] R. A. Johnson and D. W. Wichern, "APPLIED MULTIVARIATE STATISTICAL ANALYSIS.," *Biometrics*, vol. 54, no. 3, p. 1203, Sep. 1998, doi: https://doi.org/10.2307/2533879.
[33] H. Li, M. Trocan, M. Sawan, and D. Galayko, "SERIAL DECODERS-BASED AUTO-ENCODERS FOR IMAGE RECONSTRUCTION," *Appl. Sci.*, vol. 12, no. 16, p. 8256, Aug. 2022, doi: https://doi.org/10.3390/app12168256.
[34] A. Mehrabinezhad, M. Teshnehlab, and A. Sharifi, "AUTOENCODER-PCA-BASED ONLINE SUPERVISED FEATURE EXTRACTIONSELECTION APPROACH," *J. Artif. Intell. Data Min. (JAIDM),* vol. 11, pp. 525–534, 2023.
[35] S. Chen and W. Guo, "AUTO-ENCODERS IN DEEP LEARNING—A REVIEW WITH NEW PERSPECTIVES," *Mathematics*, vol. 11, no. 8, p. 1777, Apr. 2023, doi: https://doi.org/10.3390/math11081777.
[36] G. Dong, G. Liao, H. Liu, and G. Kuang, "A REVIEW OF THE AUTOENCODER AND ITS VARIANTS: A COMPARATIVE PERSPECTIVE FROM TARGET RECOGNITION IN SYNTHETIC-APERTURE RADAR IMAGES," *IEEE Geosci. Remote Sens. Mag.*, vol. 6, no. 3, pp. 44–68, Sep. 2018, doi: https://doi.org/10.1109/MGRS.2018.2853555
[37] F. Laakom, J. Raitoharju, A. Iosifidis, and M. Gabbouj, "Reducing redundancy in the bottleneck representation of

autoencoders," *Pattern Recognit. Lett.*, vol. 178, pp. 202–208, Feb. 2024, doi: 10.1016/j.patrec.2024.01.013. doi: https://doi.org/10.1016/j.patrec.2024.01.013

[38] E. Palut, "FROM PRINCIPAL SUBSPACES TO PRINCIPAL COMPONENTS WITH LINEAR AUTOENCODERS," *Mach. Learn.*, 2018.

[39] M. P. Gore, A. A. Shinde, and A. Choudhury, "APPLYING PRINCIPAL COMPONENT ANALYSIS AND AUTOENCODERS FOR DIMENSIONALITY REDUCTION IN DATA STREAM," *Int. J. Innov. Res. Eng. Manag.*, vol. 11, no. 5, pp. 121–126, Oct. 2024, doi: https://doi.org/10.55524/ijirem.2024.11.5.17.

[40] D. Cacciarelli and M. Kulahci, "HIDDEN DIMENSIONS OF THE DATA: PCA VS AUTOENCODERS," *Qual. Eng.*, vol. 35, no. 4, pp. 741–750, Oct. 2023, doi: https://doi.org/10.1080/08982112.2023.2231064.

[41] R. Cardoso Pereira, M. Seoane Santos, P. Pereira Rodrigues, and P. Henriques Abreu, "REVIEWING AUTOENCODERS FOR MISSING DATA IMPUTATION: TECHNICAL TRENDS, APPLICATIONS AND OUTCOMES," *J. Artif. Intell. Res.*, vol. 69, pp. 1255–1285, Dec. 2020, doi: https://doi.org/10.1613/jair.1.12312

[42] A. Abhaya and B. K. Patra, "AN EFFICIENT METHOD FOR AUTOENCODER BASED OUTLIER DETECTION," *Expert Syst. Appl.*, vol. 213, p. 118904, Mar. 2023, doi: https://doi.org/10.1016/j.eswa.2022.118904

[43] N. Olabi and M.V Zurich, "DIMENSIONALITY REDUCTION FOR DISTRIBUTED MACHINE LEARNING," University of Zurich, 2025.

[44] T. O. Hodson, T. M. Over, and S. S. Foks, "MEAN SQUARED ERROR, DECONSTRUCTED," *J. Adv. Model. Earth Syst.*, vol. 13, no. 12, Dec. 2021, doi: https://doi.org/10.1029/2021MS002681

[45] A. G. Priya Varshini, K. Anitha Kumari, D. Janani, and S. Soundariya, "COMPARATIVE ANALYSIS OF MACHINE LEARNING AND DEEP LEARNING ALGORITHMS FOR SOFTWARE EFFORT ESTIMATION," *J. Phys. Conf. Ser.*, vol. 1767, no. 1, p. 012019, Feb. 2021, doi: https://doi.org/10.1088/1742-6596/1767/1/012019