Proposed Two-Parameter Estimator for Estimating Linear Regression Model and Comparing It with Some Other Estimators

NoorAlzahraa Naeem Abd Ali[1,*], Shrooq Abdul Redha Al Sabah[1]

[1]Department of statistics , Faculty of Administration & Economics, University of Kerbala, Kerbala , Iraq

Emails: nooralzahraa.n@s.uokerbala.edu.iq; shorouq.a@uokerbala.edu.iq

Abstract

In this paper, a new two-parameter estimator was proposed to estimate the parameters of the linear regression model that has the ability to face the problem of Multicollinearity based on the previous information about the parameters to be estimated and this estimator was compared with the two-parameter estimator of the linear regression model of Kaciranlar and the two-parameter estimator of the linear regression model (Lokman et al. [1]) using the mean square error criterion (MSE) for each model by conducting Monte-Carlo simulation to study the behavior of the proposed estimator. It was concluded that the proposed method is better than the rest of the estimation methods because it achieved the lowest comparison criteria, and in the case of high Multicollinearity between the explanatory variables, the proposed method was very effective in solving this problem. Data representing (100) observations of the number of women with Irritable Bowel Syndrome (IBS) for the years (2020-2023) from the Karbala Holy Health Department were used, which represents the dependent variable (y) and a group of variables affecting the incidence of the disease, with nineteen variables. It was concluded that irritable bowel syndrome among women is decreasing, as the predictive values according to the proposed method are appropriate for the estimated values during the next five years.

Keywords: Linear Regression Model; mean square error criterion (MSE); Irritable Bowel Syndrome (IBS)

1. Introduction

Regression analysis is a fundamental tool for understanding and modeling relationships between variables and is a cornerstone of statistical methodology, providing a powerful framework for exploring relationships and making informed decisions in the face of uncertainty. It is increasingly important across the scientific community across a variety of fields, including economics, sociology, psychology, biology, medicine, industry, agriculture, and more. At its core, regression analysis seeks to uncover underlying patterns within data, allowing us to make predictions, test hypotheses, and derive meaningful insights. The essence of regression analysis lies in its ability to explore the extent to which one variable (the dependent variable) depends on one or more other variables (the independent variables). By fitting a regression model to the observed data, we aim to measure the strength and nature of these relationships. One of the key features of regression analysis is its flexibility. It accommodates a wide range of data types and can handle both continuous and categorical variables. Regression models can take many forms, from simple linear models to complex nonlinear models, providing a versatile toolkit for data analysis. However, regression analysis is not without challenges and problems, as the problem of Multicollinearity can arise when there is a strong correlation between the explanatory variables, which in turn leads to an increase in the variance of the regression coefficient estimates, which makes the OLS method unreliable. The problem of Multicollinearity: Several methods have been proposed, including Ridge Regression, which in turn had a great diversity in avoiding the problem of Multicollinearity. Ordinary least squares (OLS) estimators have a set of assumptions that must be met according to the Gauss–Markov theorem, namely that the samples representing the population are drawn randomly, that the estimated parameters are the best linear unbiased estimate (BLUE) and have the least variance among all other estimators, that there is no linear relationship between the independent variables in the model, that random errors are not related to the independent variables, and that the variance of the errors is constant. However,

DOI: https://doi.org/10.54216/PMTCS.040201

Received: January 24, 2024 Revised: May 14, 2024 Accepted: August 08, 2024

these assumptions are rarely met in practical life, and the variables are linked together by a relationship called the multicollinearity problem, which causes a violation of the assumptions of the Gauss–Markov theorem, and as a result the estimators are not good, unstable, and do not represent the reality represented by these phenomena. Therefore, it is necessary to find an estimator who has the ability to confront and deal with this problem to reach accurate estimates of the parameters of the model that represents these phenomena, and among these phenomena is irritable bowel syndrome in women. From a statistical point of view, studying this phenomenon is considered complex due to the Multicollinearity between the various factors that affect the disease, such as age, weight, insulin levels, blood pressure levels, etc. Therefore, it is necessary to study the effect of Multicollinearity on the analysis of the factors affecting irritable bowel syndrome and to extract accurate results about these factors. Therefore, this research came to propose a new two-parameter estimator to estimate the parameters of the linear regression model that has the ability to confront the problem of Multicollinearity based on previous information about the parameters to be estimated and compare this estimator with the two-parameter estimator for the linear regression model of Kaciranlar and the two-parameter estimator for the linear regression model (Lokman et al. [1]) using the mean square error criterion MSE for each model. In (2010) Yang & Chang [2] proposed a new two-parameter estimator that includes ordinary least squares (OLS) estimator, regression graph (RR) estimator, and Liu estimator as special cases. The necessary and sufficient conditions for the superiority of the new estimator over OLS, RR, Liu estimators, and the two-parameter estimator proposed by Ozkale and Kaciranlar [3] in the mean square error matrix (MSEM) were derived. Moreover, the bias coefficient estimates were obtained. It was concluded that the proposed estimator actually achieved higher performance than the other estimators [3]. In (2019), Lukman et al. [4] proposed a new modified bi-parameter estimator based on prior information of the parameter vector to circumvent the multicollinearity problem which included the special cases of ordinary least squares estimator (OLSE), edge regression estimator, Liu estimator, modified edge estimator, and modified Liu estimator. The new estimator was found to outperform OLSE, RRE, LE, MRE, MLE, and the bi-parameter estimator proposed by Özkale and Kaçıranlar [5] using the mean square error matrix criterion by conducting simulation experiments to evaluate the performance of the new estimator and compare it with other estimators [4]. In the same year, Olasunkanmi [6] proposed a new modified Liu Ridge estimator called Liu-Dawoud-Kibria instead of OLS in estimating the parameters of the general linear model. The results of theoretical comparison and simulation study showed that the proposed estimator outperforms others in some conditions, using the mean square error criterion [6]. In (2023), Abdulrasheed et al. [7] developed a new two-parameter estimator to address the risk of multicollinearity of the linear regression model. Some necessary and sufficient conditions were obtained for the dominance of the proposed estimator over the ordinary least squares (OLS) estimator, the letter regression estimator, the Liu estimator, the KL estimator, and some two-parameter estimators in the mean square error matrix [7]. In (2014) Dorugade [8] proposed an efficient two-parameter alternative estimator and its properties were examined. It was compared with the ordinary least squares (OLS) and ordinary regression estimators (ORR) using the mean squares error criterion, and it was found that the proposed estimator performs more efficiently than the OLS and ORR estimators through Monte Carlo simulation experiments [8], In (2024) Abdelwahab et al. [9] introduced a new two-parameter Liu estimator to address the problem of multicollinearity in the Poisson regression model (PRM). The Poisson maximum likelihood estimator (PMLE) traditionally estimates PRM. When the explanatory variables are correlated, leading to multicollinearity, the variance or standard error of the PMLE is inflated. To address this problem, several alternative estimators have been introduced, including the Poisson ridge regression estimator (PRRE), the Liu estimator (PLE), and the modified Liu estimator (PALE) [9].

2. Gauss–Markov theorem

We can express the linear regression model as follows [10]:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots . + \beta_K X_{ip} + u_i \quad ; i=1,2,\ldots.., n \tag{1}$$

Where $\beta_0, \beta_1, \beta_2 \ldots\ldots \beta_p$ are regression Coefficients, p: Number of explanatory variables

u_i random error. X_{ip} are explanatory variables [11].

It can be written in matrix form as follows:

$$\underline{Y} = X\underline{\beta} + \underline{U} \tag{2}$$

Where Y is vector nx1 is assumed to be normally distributed with zero mean and constant variance

X: matrix of independent variables nxp, β vector of coefficients px1, U vector of errors nx1

The parameters of model (2) are estimated using the OLS method according to the following formula:

$$\hat{\beta} = (X'X)^{-1}(X'Y) \tag{3}$$

DOI: https://doi.org/10.54216/PMTCS.040201
Received: January 24, 2024 Revised: May 14, 2024 Accepted: August 08, 2024

The estimated regression equation that gives n values of the mean response ($\hat{y}$) is written as follows:

$$\hat{y} = X\hat{\beta} \quad (4)$$

$\hat{\beta}$ best linear unbiased estimator (BLUE) of the model parameters β, under ordinary least squares assumptions [12]. The word best is defined as having minimal variance and being unbiased, and the regularity conditions are the OLS assumptions. There is no guarantee that a nonlinear method, for example, will not be better for our data by some other measure, but if we want to use an unbiased linear model and if the OLS assumptions apply to our data, then we should use only OLS. The optimum estimator in this way is sometimes referred to as "BLUE", for best unbiased linear estimator [12].

3. Multicollinearity Problem

(Ranger Frisch, 1934) developed the concept of multicollinearity, or multicollinearity, and described it as the existence of a complete or incomplete linear relationship between all or some of the explanatory variables in the regression model, which leads to a violation of one of the assumptions of ordinary least squares (OLS) in the absence of a relationship between the values of the observations of the explanatory variables in the regression model to be estimated, with which the effect of the variables cannot be separated from each other, which leads to inaccurate estimates of the parameters of the regression model [13]. The term multilinearity consists of the word (Multi) which means multiple, the word (Co) which means common, overlapping or linked, and the word (Linearity) which means linear [14]. (Al-Sabbah et al., 2019, 493). Multicollinearity can be defined through the concept of orthogonality, i.e. when there is a perfect, linear relationship between two or more explanatory variables, i.e. the information matrix (X'X) is of full rank and all the eigenvalues are equal to one. This indicates that orthogonality exists, which makes it impossible to find the inverse of the information matrix $(X'X)^{-1}$, because the determinant of this matrix will be equal to zero [15], [16].

4. Bayesian Ridge Regression

The traditional regression method has been used in many studies, as this method has proven its effectiveness in dealing with the problem of multicollinearity, but its estimates are biased. However, in the regular regression method, the prior information will be entered into the regular regression model to get rid of the bias problem in estimating the parameters. As we have previously shown, the regular regression method was developed by [17]. This method is summarized by adding the constant K to the matrix (X'X) before taking its inverse. Swindle [18] suggested introducing the prior information into the letter regression formula, so the updated letter regression formula will be as in Equation (40-2) where we notice that the constant K was introduced with the vector b to the vector X'Y which represents the prior information. We also notice that in the case of K=O we will get the ordinary least squares (OLS) estimate. (Al-Kafishi and the Swimmer, 2020, 22). Through the concept of Bayes' theorem, which is as follows:

$$\text{Posterior dist.} \propto \text{Prior Prob.} \times \text{Likelihood}$$

If we have the vector b which represents the initial information vector, and as is known from the assumptions of the linear model, the error u_i is normally distributed with mean 0, and variance $\sigma^2 I_n$, that is, $u_i \sim N(0, \sigma^2 I_n)$, then the least squares estimates are normally distributed with mean β and variance $\sigma^2(X'X)^{-1}$ [19]. Let us assume that the vector b containing the previous information has a normal distribution with mean β, and the variance-covariance matrix V is $b \sim N(B, V)$. Assuming that V is a full rank matrix and represents the variance-covariance matrix, then the convex estimator: Assuming that V is a full rank matrix and represents the variance-covariance matrix, then the convex estimator [20].

$$\beta(C, b) = C\hat{B}_{Ols} + (I - C)b \quad (5)$$

$\beta(C, b)$ is the convex estimate and it is a convex real-valued function. One of its properties is that if there are any two points in the domain of this function, such as x and y, then if the straight line that connects any two points on its graph lies above the graph of the function, then,

$$B(C, b) = cf(x) + (1 - c)f(y) \quad (6)$$

Then $cf(x)$=$f(x)$, I is a one-dimensional matrix of order P×P, C is a matrix of dimension P×P.

We note from formula (6) that matrix C is an unknown matrix. To find this matrix, we will find the average sum of squares of the error for the convex estimate, which takes the following form:

$$MSE(\beta(C, b)) = E((\beta(C, b) - \beta)'((\beta(C, b) - \beta) \quad (7)$$

Substituting $\hat{B}_{Ols}$in equation (6) yields: (Al-Kafishi and Al-Sabbah, 2020, 34)

DOI: https://doi.org/10.54216/PMTCS.040201
Received: January 24, 2024 Revised: May 14, 2024 Accepted: August 08, 2024

Also, substituting the value of Y yields:

$\beta(C,b) = C(X'X)^{-1}X'(X\beta + U) + (I - C)b$

$\beta(C,b) = C(X'X)^{-1}X'X\beta + C(X'X)^{-1}X'U + (I - C)b$

$\beta(C,b) = C\beta + C(X'X)^{-1}X'U + (I - C)b$

$(\beta(C,b) - \beta) = C(X'X)^{-1}X'U + (I - C)b$ (8)

By squaring the relationship (8) as follows:

$(\beta(C,b) - \beta)'(\beta(C,J) - \beta) = (C(X'X)^{-1}X'U + (I - C)b)'\,(C(X'X)^{-1}X'U + (I - C)b)$

$=U'X(X'X)^{-1}C' + b'(I - C)')(C(X'X)^{-1}X'U + (I - C)b)$

$=C(X'X)^{-1}X'UU'X(X'X)^{-1}C' + U'X(X'X)^{-1}C'(I - C)b + b'(I - C)'C(X'X)^{-1}X'U + b'(I - C)'(I - C)b$ (9)

Taking the expectation of both sides and assuming that $U \sim N(0, \sigma^2{}_{In})$ yields:

$= \sigma^2\, C'(X'X)^{-1}C + 0 + 0 + V(I - C)'(I - C)$

$= \sigma^2\, C'(X'X)^{-1}C + V(I - C'I - CI + CC')$ (10)

$MSE(\beta(C,b)) = tr(\sigma^2\, C'(X'X)^{-1}C + CVC' - CV - VC' + V)$ (11)

$$\frac{\partial MSE(\beta(C,b))}{\partial C} = 2C\sigma^2\,(X'X)^{-1} + 2CV - V - V$$

$$2C(\sigma^2\,(X'X)^{-1} + V) - 2V = 0$$

$$C = \frac{V}{\sigma^2\,(X'X)^{-1} + V} \quad (12)$$

The matrix C was estimated, considering that J is normally distributed with rate □ and variance V, and that $\hat{\beta}_{ols}$ is normally distributed with rate β and variance Σ, i.e. $\hat{\beta}_{ols} \sim N(\beta, \boldsymbol{\Sigma})$, $J \sim (\beta, V)$, So the matrix C will be as follows [21]:

$$C = \frac{V}{V + \Sigma} \quad (13)$$

Therefore, the convex estimate in formula (13) is written in the following form:

$$\beta(\boldsymbol{C}, \beta_{\boldsymbol{0}}) = \boldsymbol{C}\hat{\beta}_{\boldsymbol{Ols}} + (\boldsymbol{I} - \boldsymbol{C})\beta_{\boldsymbol{0}} \quad (14)$$

The value of C that was estimated in the previous steps will be as follows:

$$\boldsymbol{C} = \frac{V_0\hat{\beta}_{Ols} + \beta_0\sigma^2\,(\boldsymbol{X'X})^{-1}}{V_0 + \sigma^2\,(\boldsymbol{X'X})^{-1}} \quad (15)$$

The estimated of the Bayesian ridge regression is as follows:

$$\hat{B}_{BaysRR} = (X'X + \sigma^2 V_0^{-1})^{-1}\,(X'Y + \sigma^2 V_0^{-1}\beta_0) \quad (16)$$

5. Two parameters Regression Model estimator for Kaciranlar estimator

The main idea of the two-parameter (Kaciranlar) estimator is to derive a two-parameter graph estimator that combines the advantages of the graph regression estimator and the advantages of the shrink estimator as follows:

Since the distance between $\hat{\beta}_{OLS}$ and β increases in the case of multicollinearity problem, our goal will be to minimize the distance between the true parameter and the estimated parameter, but minimizing the distance will be to a certain level of mean squares of the residuals, as we can get many estimators that give the same sum of squares of the residuals, but we choose the estimator that has the smallest square distance. It may then be better to minimize the distance between the estimator and a certain point. This point can be chosen based on the previous information (Hu & Xinfeng, 2010, 12). Suppose the prior information is given in the form of a non-random vector b as in Swindel (1976) but choosing the point as the specified b values as the prior information leads to the loss of important properties of the original model. When there is multicollinearity, the OLS method produces estimated regression coefficients that are very large in absolute value and the $\hat{\beta}_{OLS}$ vector is longer than the β vector. In other words $\|\beta\| > \|\hat{\beta}_{OLS}\|$ [22]. The RR and shrinkage estimators have a shorter length than the OLS estimator, i.e.

DOI: https://doi.org/10.54216/PMTCS.040201
Received: January 24, 2024 Revised: May 14, 2024 Accepted: August 08, 2024

$||\hat{\beta}_{OLS}|| < ||\hat{\beta}_{RR}||,\ \hat{\beta}_{SH}(d)||$. Which can also be considered as a remedy when $||\hat{\beta}_{OLS}||$ is too long due to collinearity. Therefore, a preferred choice for this point is $\hat{\beta}_{SH}(d)$ because it confines the estimator to a suitable parameter space. It may be thought that instead of $\hat{\beta}_{SH}(d)$different estimators can be used to obtain the same parameter space but $\hat{\beta}_{SH}(d)$is a linear function of d so the calculations are much simpler than any other estimators are. Use the procedure of obtaining the regression estimator to arrive at a new estimator that faces the multicollinearity problem as a solution to the following minimization problem [23]:

$$\therefore \hat{\beta}_{SHRR}(k,d) = \text{Min}\ [(Y-XB)'\,(Y-XB)\ ] \qquad (17)$$

Subject to: $(\beta - \hat{\beta}_{SH}(d))'(\beta - \hat{\beta}_{SH}(d)) \leq c$

Considering the shrinkage estimator as a constraint imposed on the parameters to be estimated,

$$\therefore \hat{\beta}_{SHRR}(k,d) = \text{Min}\ (Y-XB)'\,(Y-XB) + k(:\ (\beta - \hat{\beta}_{SH}(d))'(\beta - \hat{\beta}_{SH}(d)) - c) \qquad (18)$$

Since c is a constant that has no effect on the solution and $k \geq 0$ is the bias parameter, equation (18) can be written as follows [24]:

$$\therefore \hat{\beta}_{SHRR}(k,d) = \text{Min}(Y-XB)'\,(Y-XB) + k:\ \ (\beta - \hat{\beta}_{SH}(d))'(\beta - \hat{\beta}_{SH}(d)) \qquad (19)$$

The estimated values of the model parameters are obtained by the modified letter regression method by minimizing the sum of squares of the following penalty errors [25]:

$$(Y-XB)'\,(Y-XB) + k:\ \ (\beta - \hat{\beta}_{SH}(d))'(\beta - \hat{\beta}_{SH}(d))) \qquad (20)$$

By deriving the equation (20) with respect to β and setting the derivative equal to zero, we get:

$$\therefore \hat{\beta}_{SHRR}(k,d) = \left(X'X + kI_p\right)^{-1}\ \ (X'Y - k\,\hat{\beta}_{SH}(d))$$

$$= \left(X'X + kI_p\right)^{-1}\ \ (X'X + kd)\,\hat{\beta}_{OLS} = T_{kd}\hat{\beta} \qquad (21)$$

6. Two parameters regression model estimator Lokman et al. [1]:

(Lukman et al.) proposed a new estimator for the two-parameter linear regression model based on the modified Liu estimator $\hat{\beta}_{MLE}(d)$ and the modified letter regression estimator $\hat{\beta}_{MRR}$which are a convex combination of prior information b and the ordinary least squares estimator $\hat{\beta}_{OLS}$. He modified Liu estimator $\hat{\beta}_{MLE}(d)$ is:

$$\hat{\beta}_{MLE}(d) = \left(\acute{X}X + I\right)^{-1}\left[\left(\acute{X}Y + dI\right)\hat{\beta}_{OLS} + (1-d)b\right] \qquad (22)$$

Let:

$$T_K = (\acute{X}X + kI)^{-1}\acute{X}X = I - k(\acute{X}X + kI)^{-1} \qquad (23)$$

The estimated slope of the modified letter becomes:

$$\hat{\beta}_{MRR}(k,b) = \left(X'X + kI_p\right)^{-1}\ (X'Y - kb)$$

$$= \left(X'X + kI_p\right)^{-1}\ X'Y +\ K(\acute{X}X + KI)^{-1}b$$

$$= \left(X'X + kI_p\right)^{-1}\ X'X\hat{\beta}_{OLS} +\ K(\acute{X}X + KI)^{-1}b$$

$$= T_k\hat{\beta}_{OLS} + (I - T_k)b \qquad (24)$$

Let:

$$T_d = \left(\acute{X}X + I\right)^{-1}\left(\acute{X}X + dI\right) \qquad (25)$$

From equation (25) we get:

$$T_{kd} = \ \left(X'X + kI_p\right)^{-1}(X'X + kd)$$

$$=\ I - k(1-d)\left(X'X + kI_p\right)^{-1} \qquad (26)$$

Then :

$$\hat{\beta}_{LOKRR}(k,d,b) = T_{kd}\hat{\beta}_{ols} - (I - T_{Kd})b$$

$$= \acute{X}X + KI)^{-1}(\acute{X}X + kdI)\hat{\beta}_{OLS}$$

DOI: https://doi.org/10.54216/PMTCS.040201
Received: January 24, 2024 Revised: May 14, 2024 Accepted: August 08, 2024

$$+\left(I-(\acute{X}X+kI)^{-1}(\acute{X}X+kId)\right)b$$

$$=(\acute{X}X+kI)^{-1}(\acute{X}X+kdI)\hat{\beta}_{ols}$$

$$+\left(k(I-d)(\acute{X}X+kI)^{-1}\right)b$$

$$=(\acute{X}X+kI)^{-1}\left[(\acute{X}X+kdI)\hat{\beta}_{ols}+k(I-d)b\right] \quad (27)$$

7. Suggested two parameters regression model estimator [26]:

A new two-parameter estimator based on the ridge regression perspective is proposed using the Bayesian ridge regression estimator procedure as a constraint on the parameters to be estimated to arrive at a new estimator that addresses the Multicollinearity problem and is unbiased at the same time as a solution to the following minimization problem:

$$\hat{\beta}_{SURR}(k,b)=\text{Min}\ [(Y-XB)'\ (Y-XB)\ ] \quad (28)$$

Subject to: $(\beta-\ \hat{\beta}_{RRBayes})'(\beta-\ \hat{\beta}_{RRBayes})\leq c$

Considering the shrinkage estimator as a constraint imposed on the parameters to be estimated,

$$\hat{\beta}_{SURR}(k,b)=\text{Min}\ (Y-XB)'\ (Y-XB)+k\ (\beta-\ \hat{\beta}_{RRBayes})'(\beta-\ \hat{\beta}_{RRBayes})-c) \quad (29)$$

Since c is a constant that has no effect on the solution and $k\geq 0$ is the bias parameter, equation (29) can be written as follows:

$$\therefore\ \hat{\beta}_{SHRR}(k,d)=\text{Min}(Y-XB)'\ (Y-XB)+k\ (\beta-\ \hat{\beta}_{RRBayes})'(\beta-\ \hat{\beta}_{RRBayes}) \quad (30)$$

The estimated values of the model parameters are obtained by the modified letter regression method by minimizing the sum of squares of the following penalty errors:

$$(Y-XB)'\ (Y-XB)+k\ (\beta-\ \hat{\beta}_{Bays RR})'(\beta-\ \hat{\beta}_{Bays RR}) \quad (31)$$

By deriving the equation (31) with respect to β and setting the derivative equal to zero, we get:

$$\hat{\beta}_{SURR}(k,b)=\left(X'X+kI_p\right)^{-1}\quad (X'Y-k\,\hat{\beta}_{BaysRR}) \quad (32)$$

8. Simulation study

Simulations are computer experiments that involve generating data by taking pseudo-random samples from known probability distributions. They are an important tool for statistical research, especially for evaluating new methods and comparing alternative methods. Simulations are often used in the pages of statistics in medicine, but they suggest that some statisticians lack the understanding needed to confidently conduct a simulation study, while others are overconfident and therefore fail to think carefully about the design and report the results poorly. A sound understanding of simulation studies would enable the former to conduct and critically evaluate published simulation studies themselves, while the latter would enable simulation studies to be conducted more carefully and reported transparently. Simulation studies are experimental experiments, and therefore statisticians must apply their knowledge of experimental design and analysis to their conduct. As we will see, shortcomings in design, analysis, and reporting lead to uncritical use and interpretation of simulation studies. In this context, a better understanding of the rationale, design, implementation, analysis, and reporting of simulation studies is essential to improve understanding and interpretation of results. (Tim P et al., 2019:1). In the simulation experiments, the assumed regression models were set as follows:

Table 1: Default parameters of the studied models

Model	ρ	n	p
	0.3	50, 100	10
	0.9	50, 100	

Where p Number of explanatory variables in the model, n Sample size, ρ Correlation coefficient to determine the degree of association of explanatory variables in model.

DOI: https://doi.org/10.54216/PMTCS.040201
Received: January 24, 2024 Revised: May 14, 2024 Accepted: August 08, 2024

The data generation: First: Generating explanatory variables with a number p from the normal distribution according to the instruction X = randn(n, p). Second: Adding Multicollinearity to the variables by generating the covariance matrix according to the following formula Sigma = rho * ones(p) + (1 - rho) * eye(p), Three correlation coefficients were chosen rho=0.1, 0.9, Then finding the amount of Multicollinearity generated between the explanatory variables L using Cholesky decomposition according to the following instruction L = chol(Sigma, 'lower') , Cholesky decomposition is a mathematical algorithm used to analyze positively deterministic symmetric matrices. This decomposition is an important tool in linear algebra and many practical applications, especially in solving systems of linear equations and calculating inverses, and in statistical models and other places [27].

The idea of Cholesky decomposition is to transform a positively definite symmetric matrix A into the product of two matrices: a lower triangular matrix L and an upper triangular matrix LT, which is the reciprocal transformation (transposition) of the matrix L. The mathematical formula for this decomposition is A=LLT (Hao et al., 2023). Bayesian Ridge Regression estimated by By choosing the hyperbolic parameter of the prior distribution b = 2 and the value of σ^2=0.9; then calculating the mean of the prior distribution and its covariance as follows:

A = X' * X + k * eye (p); % Design matrix + regularization term (33)

V_0 = inv (A + b * eye (p)); % Covariance matrix of weights (34)

β_0 = Sigma * (X' * y); % Posterior mean of weights (35)

The regularization parameter k was chosen according to the empirical method assuming an initial value of k = 0.1. Two parameters Regression Model estimator for Kaciranlar estimator estimated and Two parameters regression model estimator Lokman et al. [1] by using and Suggested two parameters regression model estimator by using $d = (1 + \text{Rho})^{-1}$ At each value of the correlation coefficient Rho by choosing the parameter b = 2. The results as following:

First experiment: The regression model was applied with the number of variables p=10 and sample size n=50, 100, 200 and values rho=0.1, 0.9 according to the following model:

$Y_i = \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \beta_4 X_{i4} + \beta_5 X_{i5} + +\beta_6 X_{i6} + \beta_7 X_{i7} + \beta_8 X_{i8} + \beta_9 X_{i9} + \beta_{10} X_{i10} + u_i$; i=1,2,….., n (36)

Table 2: Estimates of regression model coefficients for estimation methods at n=50, p=10, rho=0.3

Estimate	Bay RR	KACRR	LOKRR	SURR
$\hat{\beta}_1$	1.960	1.721	1.660	3.052
$\hat{\beta}_2$	2.177	2.131	2.103	1.557
$\hat{\beta}_3$	-0.340	-0.036	0.031	-2.027
$\hat{\beta}_4$	0.215	0.225	0.233	0.889
$\hat{\beta}_5$	3.220	2.909	2.818	3.937
$\hat{\beta}_6$	1.481	1.377	1.344	1.495
$\hat{\beta}_7$	1.404	1.436	1.432	0.657
$\hat{\beta}_8$	3.357	2.846	2.711	4.812
$\hat{\beta}_9$	5.416	4.638	4.437	8.415
$\hat{\beta}_{10}$	1.068	0.869	0.820	1.808

Table 3: Comparison criteria for each method at n=50, p=10, and rho=0.3

Criteria	MSE	AIC	BIC	HNQ	Best
'Bayesian'	12.47597	288.08406	307.20429	295.36516	Suggested
'Kaciranlar'	21.59494	315.51681	334.63704	322.79791	
'Lokman'	24.54630	321.92190	341.04213	329.20300	
'Proposed'	0.24979	92.53670	111.65693	99.81779	

DOI: https://doi.org/10.54216/PMTCS.040201
Received: January 24, 2024 Revised: May 14, 2024 Accepted: August 08, 2024

Table 4: Estimates of regression model coefficients for estimation methods at n=100, p=10, rho=0.3

Estimate	Bay RR	KACRR	LOKRR	SURR
$\hat{\beta}_1$	2.418	2.236	2.185	2.934
$\hat{\beta}_2$	1.417	1.394	1.386	1.432
$\hat{\beta}_3$	-0.675	-0.309	-0.218	-1.867
$\hat{\beta}_4$	1.376	1.445	1.456	0.993
$\hat{\beta}_5$	3.290	3.012	2.934	4.024
$\hat{\beta}_6$	1.514	1.476	1.461	1.508
$\hat{\beta}_7$	0.878	0.923	0.931	0.665
$\hat{\beta}_8$	3.651	3.272	3.168	4.667
$\hat{\beta}_9$	6.639	5.905	5.700	8.510
$\hat{\beta}_{10}$	1.463	1.330	1.293	1.789

Table 5: Comparison criteria for each method at n=100, p=10, and rho=0.3

Criteria	MSE	AIC	BIC	HNQ	Best
'Bayesian'	4.04108	443.43898	469.49069	453.98258	Suggested
'Kaciranlar'	7.63514	507.06385	533.11555	517.60744	
'Lokman'	8.87099	522.06635	548.11805	532.60994	
'Proposed'	0.20323	144.44795	170.49965	154.99154	

Table 6: Estimates of regression model coefficients for estimation methods at n=100, p=10, rho=0.9

Estimate	Bay RR	KACRR	LOKRR	SURR
$\hat{\beta}_1$	1.111	0.965	0.930	3.013
$\hat{\beta}_2$	3.246	3.038	2.971	1.259
$\hat{\beta}_3$	2.694	2.875	2.891	-2.082
$\hat{\beta}_4$	3.295	3.293	3.267	0.828
$\hat{\beta}_5$	3.338	3.171	3.122	4.012
$\hat{\beta}_6$	2.919	3.028	3.035	1.659
$\hat{\beta}_7$	2.248	2.219	2.197	1.104
$\hat{\beta}_8$	5.126	4.744	4.633	4.858
$\hat{\beta}_9$	6.225	5.486	5.289	8.720
$\hat{\beta}_{10}$	1.964	2.012	2.017	1.269

Table 7: Comparison criteria for each method at n=100, p=10, and rho=0.9

Criteria	MSE	AIC	BIC	HNQ	Best
'Bayesian'	5.77021	249.52926	268.64949	256.81035	Suggested
'Kaciranlar'	8.49603	268.87379	287.99402	276.15488	
'Lokman'	9.44610	274.17397	293.29420	281.45506	
'Proposed'	0.29272	100.46763	119.58786	107.74873	

DOI: https://doi.org/10.54216/PMTCS.040201
Received: January 24, 2024 Revised: May 14, 2024 Accepted: August 08, 2024

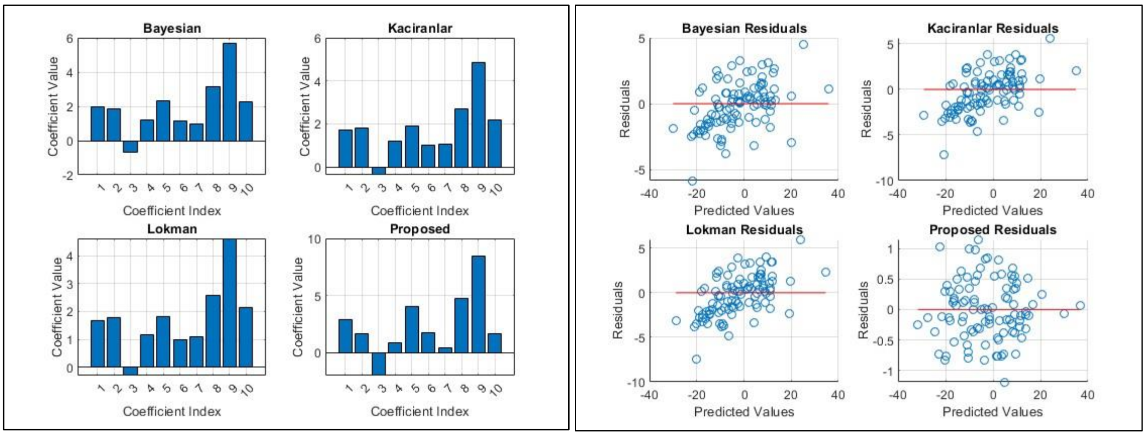

Figure 1. (a) Estimated regression model coefficients for each method at n=50, p=10, and rho=0.3; (b) Comparison criteria for each method at n=50, p=10, and rho=0.3

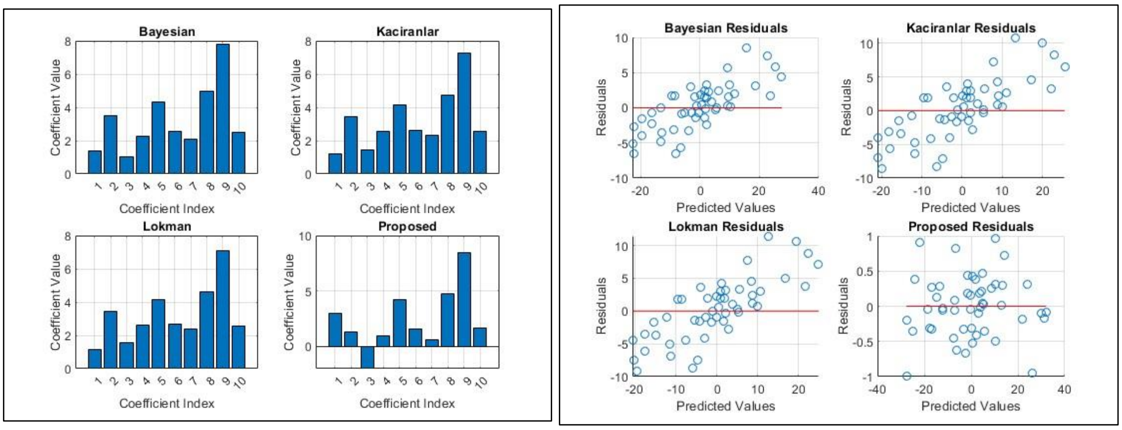

Figure 2. (a) Estimated regression model coefficients for each method at n=100, p=10, and rho=0.9; (b) Comparison criteria for each method at n=100, p=10, and rho=0.9

9. Data set applied

Data representing (100) observations of the number of women with Irritable Bowel Syndrome (IBS) for the years (2020-2023) were obtained from the Karbala Health Department - Hospital () which represents the dependent variable (y) and a group of variables affecting the incidence of the disease with nineteen explanatory variables shown in Appendix (A) which are: X1: The woman's age, X2: The woman's weight, X3: The number of hours of exercise, X4: The number of hours of sleep, X5: The woman's profession (housewife - employee - university professor), X6: Taking contraceptives (yes - no), X7: The woman's profession (housewife - employee - university professor), X8: Infection with thyroiditis, X9: The woman's educational attainment, X10: Smoking (smokers - non-smokers), X11: The percentage of insulin in the blood, X12: The level of blood pressure, X13: The percentage of cortisol (micrograms per (dL), X14: Age of woman at marriage, X15: Educational attainment of husband, X16: Age of husband, X17: Occupation of husband, X18: Duration of marriage, X19: Duration of breastfeeding, The data on the number of women with irritable bowel syndrome were tested using the Jarque-Bera test to test the normality of the data. The Jarque-Bera test (JB Test) is a statistical test used to assess whether a set of data follows a normal distribution. This test is based on comparing the skewness and kurtosis of the data with those of a normal distribution. The formula for the test is:

$$\mathrm{JB} = \frac{n}{6}\left(S^2 + \frac{(K-3)^2}{4}\right) \quad (37)$$

To test the following statistical hypothesis:

H0: The data have Normal distribution

H1: The data do not have Normal distribution

Table (6) shows the test results:

DOI: https://doi.org/10.54216/PMTCS.040201
Received: January 24, 2024 Revised: May 14, 2024 Accepted: August 08, 2024

Table 8: Results of the normal distribution test for real data

Variable	Jarque-Bera Statistic	Sig.	Decision
y	0.09010	0.12422	Accept.
X_1	0.98828	0.13432	Accept.
X_2	0.58954	0.33111	Accept.
X_3	0.46786	0.09067	Accept.
X_4	0.78976	0.78382	Accept.
X_5	0.88919	0.56742	Accept.
X_6	0.89191	0.89982	Accept.
X_7	1.89978	0.99993	Accept.
X_8	1.90891	0.77821	Accept.
X_9	0.98901	0.78811	Accept.
X_{10}	1.45633	0.88990	Accept.
X_{11}	2.89981	0.56626	Accept.
X_{12}	2.98919	0.45689	Accept.
X_{13}	0.90091	0.99010	Accept.
X_{14}	1.89919	0.16433	Accept.
X_{15}	0.89891	0.56632	Accept.
X_{16}	1.19010	0.78383	Accept.
X_{17}	2.89919	0.89975	Accept.
X_{18}	1.23443	0.88992	Accept.
X_{19}	1.38933	0.99467	Accept.

We notice from Table (8) that the Sig. values for all variables are greater than the significance level (0.01), which indicates the acceptance of the null hypothesis that states that the data follow the normal distribution

The variance inflation factor were used, as well as the discriminant values as the conditional criterion, and the conditional number Table (9) shows the variance inflation factors, the discriminant values, the conditional criterion, the Farrar & Glauber test statistic, the correlation coefficient, and the coefficient of determination for real data.

Table 9: Variance inflation coefficients, discriminant values, conditional criterion, Farrar & Glauber test, correlation coefficient and coefficient of determination for real data

Variable	VIF	Eigen Value	Condition Index CI
X_1	4.09676	0.538	109.75242
X_2	2.73850	0.363	44.67121
X_3	1.55240	0.207	25.98279
X_4	1.33986	0.125	19.46838
X_5	2.69007	0.108	13.50498
X_6	3.16177	0.088	9.47698
X_7	2.13037	0.069	8.71072
X_8	5.14657	0.056	6.39584
X_9	4.01513	0.043	5.34815
X_{10}	2.28241	0.027	4.91664
X_{11}	1.28455	0.019	4.49395
X_{12}	1.22556	0.016	3.75007
X_{13}	1.12206	0.014	3.63755
X_{14}	3.69400	0.011	3.37055
X_{15}	4.56490	0.007	3.11288
X_{16}	13.13043	0.006	2.29150
X_{17}	1.65055	0.004	1.67166
X_{18}	16.33889	0.003	1.30361
X_{19}	1.73958	0.001	1.00000
Farrar & Glauber test statistic χ_0^2			929.56841
R			0.63
$\mathbf{R^2}$			0.40

DOI: https://doi.org/10.54216/PMTCS.040201
Received: January 24, 2024 Revised: May 14, 2024 Accepted: August 08, 2024

It is clear from Table (7) that the variance inflation coefficients for the explanatory variables are greater than (1) and some of them exceed (5), which is the criterion that was set for the purpose of saying that the data has a multicollinearity problem, and that the characteristic values of the matrix (X'X) are close to zero for many explanatory variables, and that the conditional criterion (CI) is greater than (30) for the variables (X1, X2). Accordingly, the data that was studied suffers from the problem of multicollinearity, and this is also proven by the value of the Farrar & Glauber test statistic χ_0^2, which is (929.56841), which is greater than the table value of the test, which is (216.9378) at a degree of freedom (171). Figures (3) and (4) show the variance inflation factor, characteristic values, and conditional criterion, respectively.

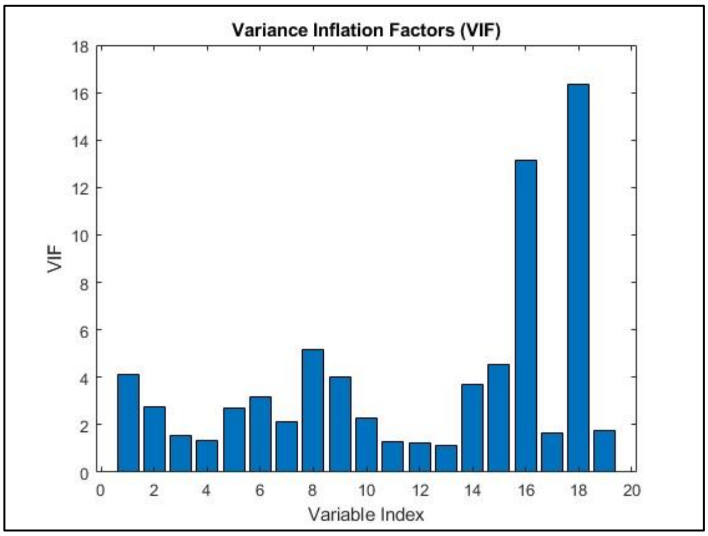

Figure 3. Variance inflation factor for real data.

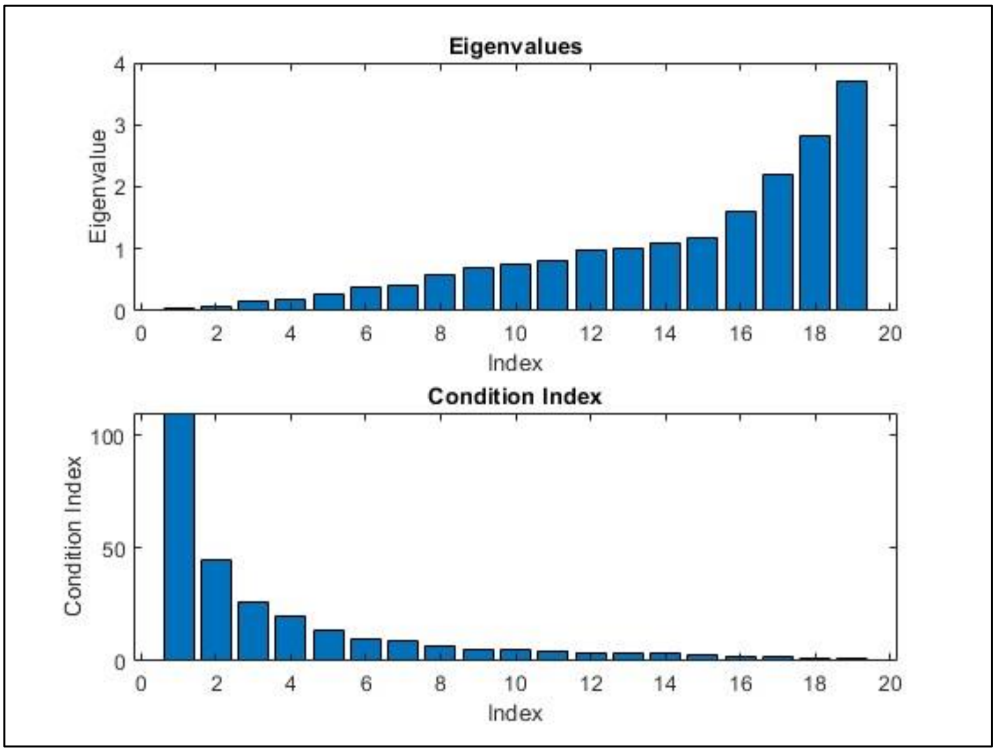

Figure 4. Distinctive values and the standard for the conditional for real data.

Analysis of real data using estimation methods that were tested in the experimental aspect of the message, which is data on the number of women with irritable bowel syndrome, which represents the dependent variable and a group of explanatory variables, the number of which is (19) variables. The following tables show the results of the analysis.

DOI: https://doi.org/10.54216/PMTCS.040201
Received: January 24, 2024 Revised: May 14, 2024 Accepted: August 08, 2024

Table 10: Analysis of real data by methods

Estimate	Bay RR	KACRR	LOKRR	SURR
$\hat{\beta}_1$	0.042	0.045	0.046	-0.016
$\hat{\beta}_2$	-0.015	-0.006	-0.004	-0.058
$\hat{\beta}_3$	-0.127	-0.118	-0.115	-0.149
$\hat{\beta}_4$	-0.186	-0.141	-0.129	-0.311
$\hat{\beta}_5$	-0.042	-0.118	-0.134	0.926
$\hat{\beta}_6$	0.691	0.613	0.590	0.741
$\hat{\beta}_7$	-1.165	-0.799	-0.712	-2.930
$\hat{\beta}_8$	-0.091	-0.088	-0.088	-0.103
$\hat{\beta}_9$	1.862	1.624	1.561	2.822
$\hat{\beta}_{10}$	0.694	0.613	0.593	1.409
$\hat{\beta}_{11}$	0.000	0.000	0.000	-0.003
$\hat{\beta}_{12}$	-0.012	-0.010	-0.010	-0.024
$\hat{\beta}_{13}$	0.030	0.027	0.026	0.030
$\hat{\beta}_{14}$	-0.111	-0.107	-0.105	0.073
$\hat{\beta}_{15}$	-0.047	-0.075	-0.081	0.101
$\hat{\beta}_{16}$	0.066	0.063	0.062	-0.068
$\hat{\beta}_{17}$	-0.622	-0.549	-0.530	-0.861
$\hat{\beta}_{18}$	0.124	0.104	0.099	0.369
$\hat{\beta}_{19}$	0.139	0.121	0.116	0.169

Table 11: Mean square error for each method for real data

Criteria	MSE	Best
'Bayesian'	1.82481	Suggested
'Lui'	2.09053	
'Kaciranlar'	2.55577	
'Lokman'	2.75964	
'Proposed'	0.00004	

DOI: https://doi.org/10.54216/PMTCS.040201
Received: January 24, 2024 Revised: May 14, 2024 Accepted: August 08, 2024

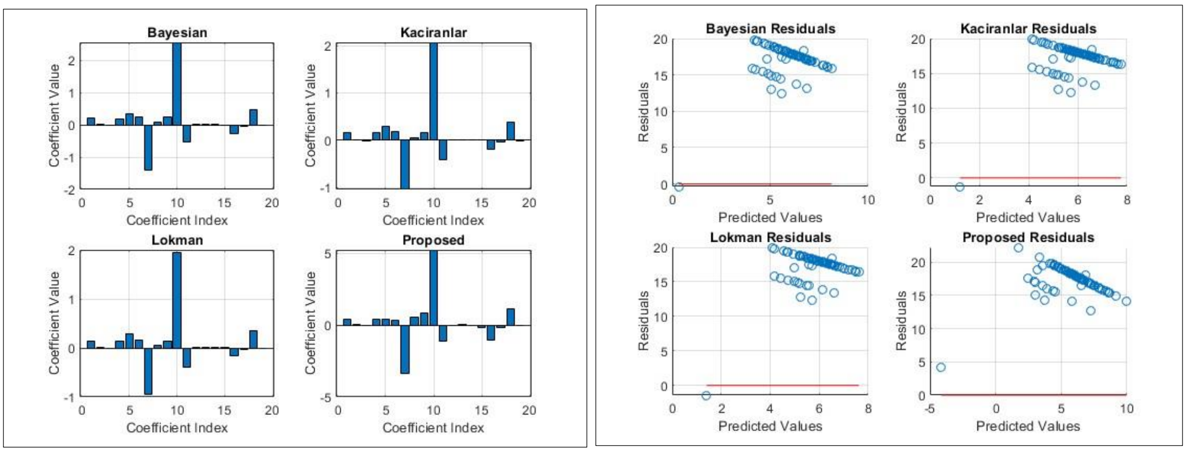

Figure 5. (a) estimated regression model coefficients for each method for real data; (b) Residuals for each method for real data

Therefore, the regression equation estimated according to the best method, which is the proposed method, is written as follows:

$$Y_i = -0.016X_{i1} - 0.058X_{i2} - 0.149X_{i3} - 0.311X_{i4} + 0.926X_{i5} + +0.741X_{i6} - 2.930X_{i7} - 0.103X_{i8} + 2.822X_{i9} + 1.409X_{i10} + -0.003X_{i11} - 0.024X_{i12} + 0.030X_{i13} + 0.073X_{i14} + 0.101X_{i15} - 0.068X_{i16} - 0.861X_{i17} + 0.369X_{i18} + 0.169_i \quad (37)$$

Then, the predictive values for the next five years were found, and the results are shown in the following table:

Table 12: Predictive values of each method for real data

'Bayesian'	'Kaciranlar'	'Lokman'	'Proposed'
0.83612	0.59592	0.53946	1.58012
1.24452	0.94255	0.86983	1. 19537
2.62870	2.15813	2.03551	1.97087
1.15652	1.08224	-1.05083	-2.04806
-0.94486	-0.67064	-0.60487	-4.60679

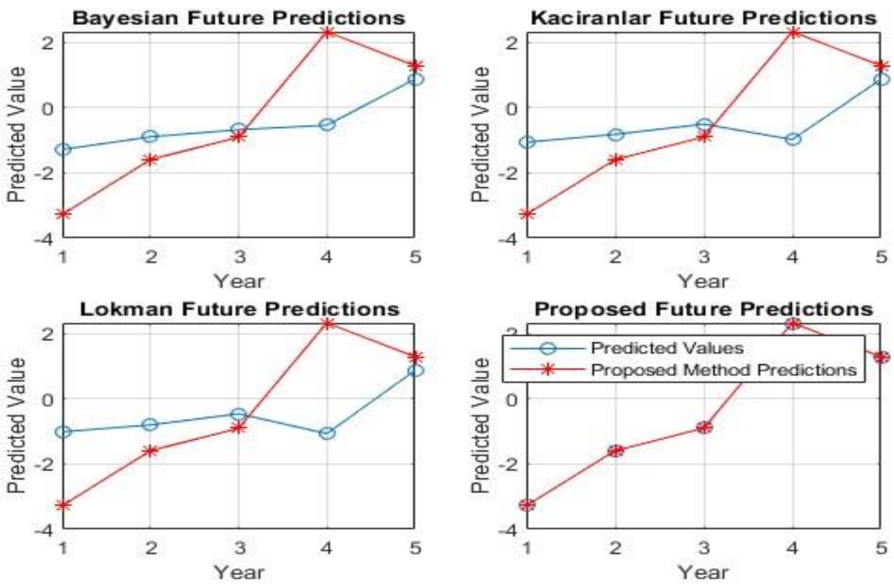

Figure 6. Predictive values for the next five years for each method for real data.

DOI: https://doi.org/10.54216/PMTCS.040201
Received: January 24, 2024 Revised: May 14, 2024 Accepted: August 08, 2024

We note from Table (12) and Figure (6) that the predictive values according to the proposed method are appropriate for the estimated values for the years and were decreasing. This indicates that irritable bowel syndrome in women is decreasing during the next five years.

10. Conclusions

- The proposed method is better than the rest of the estimation methods because it achieved the lowest comparison criteria, but this method was close to the RR method in the event that the correlation between the explanatory variables was very low.
- Irritable bowel syndrome decreased in women because the predictive values according to the proposed method were suitable for the estimated values for the years and was decreasing during the next five years.

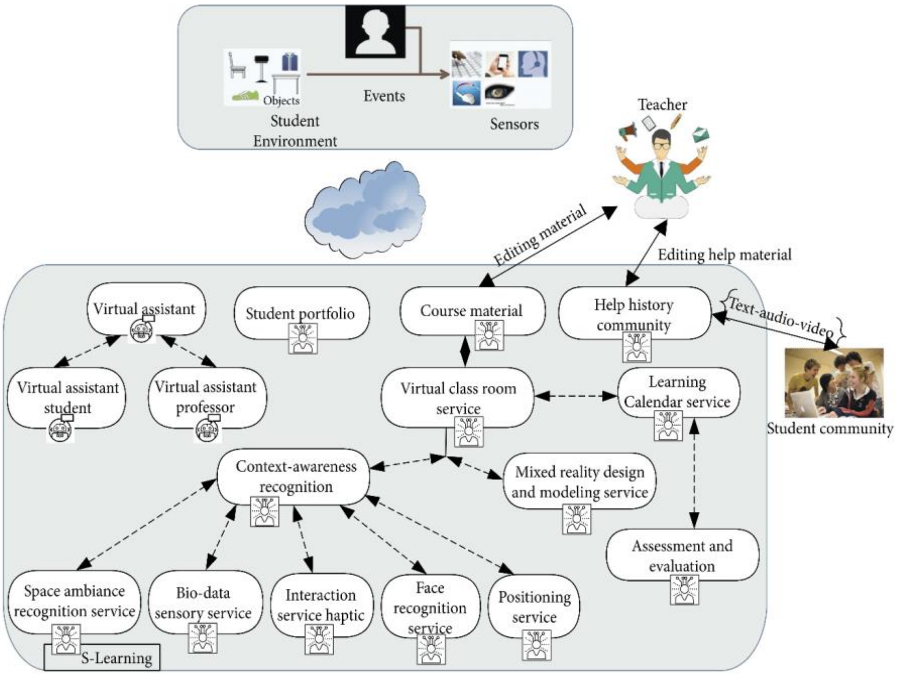

Figure 7. IoT IES services

11. Research Foundations

Research on the application of AI in virtual learning is a fascinating and rapidly evolving field. This research aims to investigate how AI helps improve learning experiences and learning (educational) outcomes as well as learning personalization.

Understanding the current landscape: The first step in the research on AI in virtual learning is to understand the current landscape, including literature review, academic papers, and industry reports to gain insights into multiple AI technologies in virtual learning, e.g., adaptive learning platforms, intelligent educational systems, and educational chatbots (ECs).

Identifying research questions: Specific research questions must be identified if the status quo of AI in virtual learning is to be well understood. Research questions should primarily be focused on the potential benefits of AI in virtual learning, AI challenges and limitations, and the effect of AI on student learning and involvement (participation).

Research method design: This allows collecting data and answering research questions, possibly including experiments, surveys, case studies, or dataset analysis, to explore the process of applying AI in virtual learning and its impact on student outcomes [6].

Since AI is active in dynamic education, it is essential to stay up-to-date with the latest developments, emerging technologies, and best procedures for meaningful relevant research.

DOI: https://doi.org/10.54216/PMTCS.040201
Received: January 24, 2024 Revised: May 14, 2024 Accepted: August 08, 2024

A. Internet of Things

The Internet of Things algorithm is a vast system of connected devices and tools that improve communication between systems and the cloud as well as between themselves, also due to the availability of economical computer chips and telecommunications; there are now thousands of devices connected to the Internet. There is. For this purpose, everyday devices, such as televisions and mobile phones, can use sensors to analyze and review information and prepare suitable and accurate solutions for people [7].

IoT integrates everyday "things" with the Internet. Computer engineers have been adding processors and sensors to everyday objects since the 1990s. Nevertheless, progress was slow at first due to big and bulky chips. Low-power computer components called RFID were originally used to locate expensive equipment. With the reduction in the size of computing systems, these parts have become smaller, stronger and smarter over time. The price and value of combining computational understanding in small parts has been greatly reduced. For example, connecting with Alexa services can be done on an MCU with a size of less than 1MB. These smart devices have the ability to transfer information quickly to the Internet. All these devices and related technologies are collectively known as the Internet of Things [8].

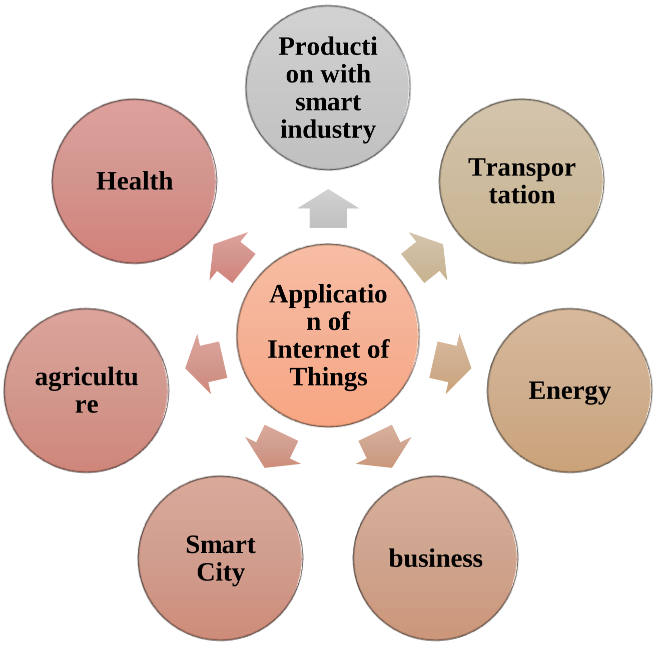

Figure 8. Application of Internet of Things

- **History of Internet of Things**

The idea of adding smart chipsets to objects was explored in the 1980s and 1990s, for example with Satfade's vending machine at Carnegie Mellon University that was connected to the Internet. This was the first system that was connected to the Internet, which was able to announce the availability of its drinks and that it gives accurate information about the hotness and coldness of the drinks. Also, Mark Wieser's 1991 research on online computing, "21st Century Computing", as well as academic sites such as UbiComp set the future of the Internet of Things [9].

- **IoT components (figure 2)**

The Internet of Things (IoT) complex is a very broad system of new technologies connected together to form an improved system for a specific purpose. The Internet of Things is a system that connects "things" by communicating with "intelligent" "things". The facilities help the devices to communicate closely with each other in the IoT system complex. IoT must be built with a managed and planned mindset to improve efficiency. However, some systems are more complex than others, six foundational ingredients are required to make them work [10].

DOI: https://doi.org/10.54216/PMTCS.040201
Received: January 24, 2024 Revised: May 14, 2024 Accepted: August 08, 2024

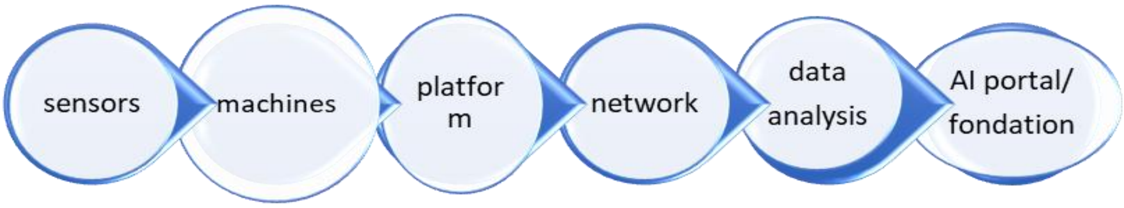

Figure 9. Foundational ingredients of IoT [11]

12. Results

The word "intelligence" implies the possibility of reasoning. Besides, researchers disagree on whether AI can reach the potential for reasoning. The main AI textbooks define it as the study of "intelligent agents": a system that perceives its environment and takes actions to maximize its chance of achieving its goals. Some popular accounts use the term to describe machines that imitate "cognitive" functions associated with the human mind, e.g., learning to solve problems. However, the main AI researchers have rejected the above definition [12].

The artificial intelligence system has the tasks of machine learning from human thinking, including training, reasoning and working methods of problems, natural language understanding, and perception, by machines, particularly computer systems. AI possesses the following capabilities: data analysis, pattern recognition, and decision-making [13].

A. AI applications in virtual learning

- **Personalized learning**

AI can evaluate student learning speed and style and offer customized instructional materials and adaptive feedback accordingly. This allows students to learn at a certain pace tailored to their learning style. Thus, they optimize their learning experiences.

- **Intelligent educational systems**

AI powers virtual tutors to interact with students, answer their questions, offer real-time feedback, and adapt lesson plans to student progress. Such systems are capable of simulating human learning and providing personalized guidance.

- **Intelligent content creation and management**

AI-powered algorithms can analyze large amounts of educational content and recommend personalized resources to learners according to their preferences and learning history. This can help create and tailor educational content to suit student needs and preferences.

- **Administrative automation**

AI can automate administrative tasks, e.g., grading, scheduling, and attendance tracking, thereby saving educators much time to further focus on teaching and interacting with students.

- **Virtual labs and simulations**

AI is used in creating virtual labs and simulating certain subjects, including science, engineering, and medicine. Such simulations provide students with practical learning experiences and allow them to conduct experiments and practice their skills in a virtual environment.

- **NLP for language learning**

AI-based NLP can facilitate the language learning process through real-time translation, feedback on pronunciation, and conversation practice. It can also help students enhance their language skills more interactively and comprehensively.

- **Adaptive assessment and feedback**

AI is used to create adaptive assessments based on student performance and offer personalized feedback to assist students in identifying areas for improvement and tracking their achievements over time.

DOI: https://doi.org/10.54216/PMTCS.040201
Received: January 24, 2024 Revised: May 14, 2024 Accepted: August 08, 2024

AI applications in virtual learning can revolutionize our learning and education procedures. Educators can share personalized, adaptive learning experiences through AI technologies, while students can benefit from engaging and interactive educational content. As it continues to develop, AI is expected to contribute significantly to shaping the future of virtual learning [11].

B. E-learning (virtual learning)

For example, most virtual trainings are used to teach the types of learning that can be done through virtual networks. Due to this, there are several challenges [14]. More than 40% (the number is only set to keep growing) of various companies and organizations are presently utilizing e learning as a regular part of their training initiatives. The evolution of the company owes much too e-learning benefits. E-learning is applied due to some of its most important features [15]:

It has flexibility: using face-to-face training requires all students to be in one place, and this creates many limitations. Virtual education allows more convenience and freedom. More freedom of content is possible and the possibility of accessing content from all places and locations is possible by downloading and uploading [12].

It is economical: non-virtual education requires space, transportation costs, accommodation and many other costs. These problems will disappear with virtual training. Moreover, these things are extremely economical. Medium-sized companies such as IBM conducted research to find a solution for on-site learning [16].

Convenient reporting: Virtual training provides the possibility to manage a lot of unmanaged information through a platform, and all information can be reported in real time. And he did all the activities of the students, including attendance, absence, academic evaluation, etc [17].

Large scope: it is difficult to use face-to-face training in this work scale and in some cases, it is not applicable. It may be difficult to start and use it, and its implementation may have a lot of economic pressure. Students can receive and evaluate a virtual training course without moving and in a personal place [14].

The data obtained from holding elite conferences regarding the development of AI in virtual education According to the subjects they had, using the future triple method including factors related to traction the future, empowering factors and inhibiting factors are as described in Table No. 13.

Table 13: Analysing the point of view of Ebergan and Merhzat [14].

Themes	**Factors**
Future traction	Change in the nature of jobs
	Having vision and planning for ai
Empowerment factors	Appropriate financing
	Having the necessary technology infrastructure
inhibiting factors	Lack of understanding of the applications and benefits of ai
	Not using an interdisciplinary approach in AI development

C. Research Background

Corona disease has affected the lives of most students in schools and countries. Also, all universities and schools were required to provide remote education based on international guidelines in order to prevent the spread of the disease. The Internet has become a widely used and excellent network for the development of distance education classes in universities and schools. In addition, the system that is continuously designed to support the improvement of virtual learning called motivational systems and virtual education, during the training process, the amount of monitoring and follow-up of students' performance is reviewed. The purpose of this article is to study an economic, effective and ideal system that is designed to monitor the information accessed and participation of students in virtual education environments. In addition, an initial research was done in the civil engineering department of the university. Based on the results of the research, e-LMSs act as an effective system for professors in educational discussion and analysis of students' performance, which especially helps to improve the efficiency and academic progress of students [18].

In another research, he specified the design process of a virtual laboratory for the purpose of training systems built through the Internet of Things. The goals of this remote research were to improve sensors, interfaces, and real world programming in engineering education, as well as dealing with industrial systems and automotive applications. With the growth of the Internet of Things system, the remote presence of the training recipient in physical facilities through cameras and graphic interfaces has been implemented. Through improved systems,

DOI: https://doi.org/10.54216/PMTCS.040201
Received: January 24, 2024 Revised: May 14, 2024 Accepted: August 08, 2024

students can exploit programming through virtual access to laboratory facilities. This system allows students to practice teaching and learning information on tablets and small computers in various places. Therefore, students can first practice control systems used in industries such as automobiles and then implement them on existing industrial systems. A number of laboratory papers in the student center were improved by adopting a pedagogical thinking based on experiential learning theory to complete the site. Expanded systems are a system where remote access is much more economical that is incorporated into expensive systems [19].

13. Discussion

The application of AI in virtual learning has offered a promising and revolutionary opportunity for the future of learning. Via personalized learning experiences, adaptive tutoring, and intelligent feedback systems (IFSs), AI can revolutionize the interaction between students and educational content. In addition, virtual education systems that use artificial intelligence have attracted good flexibility to encourage students to use virtual education systems and finally, it has brought an excellent virtual education system for the academic progress of students.

Due to the increasing progress of new technologies, professors and universities should exploit artificial intelligence as a useful tool to improve old educational methods. In addition, the combination of artificial intelligence with professors' experiences can improve university professors with new educational methods and important and valuable points regarding the academic performance of students and new educational methods. In addition, artificial intelligence can analyze and improve the ability to develop new learning methods based on students' weaknesses. According to the conclusions obtained in this research, new educational systems based on artificial intelligence can create many changes by carefully considering the challenges in virtual education and learning. Protecting student data and ensuring the responsible adoption of AI technologies are among the critical considerations in developing and deploying AI-powered educational tools. Consequently, AI has shown great potential in virtual learning and played an undeniable role in shaping the future of learning. A more effective, equitable education system can be created by harnessing the AI power to personalize learning experiences, empower educators, and create inclusive learning environments.

A. Critical challenges of IoT in e-learning

The Internet of Things technology makes the teaching process much easier and makes the classroom smarter. The exceptional improvement of global computing and Internet of Things technologies, cloud computing and analytics support educational values and good research quality. Therefore, the Internet of Things will change the society and promote the improvement of the new digital system. The amount of virtual classes as well as educational programs is increasing significantly, which has led to more virtual acceleration in educational institutions, especially universities. (Table 13):

- All universities use the hybrid cloud system to host the Internet of Things system as their organizational architecture. Students who study in the university use devices such as tablets or mobile phones. It opens a new way to increase the effectiveness of organizational architecture, educational system, teaching and educational environment. The cloud space provides us with comprehensive and integrated information technology services. Enterprise architectures used by higher education institutions rely on hybrid cloud on high-powered private clouds. Due to the significant growth of dynamic media training, there is a great demand for dynamic training.

- Using modern systems, such as blackboard, produces a significant amount of audio and video information. Classes and workshops using the Internet of Things, equipped with the latest and most up-to-date recording and broadcasting systems, allow students to have full access to important information upon request.

In addition, the Internet of Things provides exceptional challenges for the use of digital courses with increased quality. This virtual guideline highlights diverse student ethics, work honesty, and information fraud in academic communities. In addition, data application processes based on the Internet of Things, various tools and systems for professors and scientific communities have been improved for the advancement of study work and handling of these issues in higher education institutions.

- Different modern systems based on the Internet of Things are dramatically used to evaluate the performance of students with the connection and combination of channel training programs in smartphones. This very useful software can support students in teaching resources and working on researches. Professors use some other application software to show basic concepts, simulations and teaching.

- The use of new Internet of Things technologies in the university education sector includes various challenges, including security. Given that some work is being done on the security and infrastructure of the Internet of Things, there is currently no defined method to identify the business risks associated with a data breach. Based on this, the virtual education department needs to develop several standards to ensure security in these applications supported by the Internet of Things, security and the use of information privacy has become valuable.

DOI: https://doi.org/10.54216/PMTCS.040201
Received: January 24, 2024 Revised: May 14, 2024 Accepted: August 08, 2024

B. Challenges of AI-powered virtual learning and solutions

Table 14: Some challenges and solutions

N.	Challenge	Description	Solution
1	Access and equality	Most of the students do not have the same access to the systems or internet network required for virtual learning and teaching.	Investing in virtual education system: Universities can invest in virtual education development infrastructure to facilitate unrestricted access to modern e-learning technology for all students. Development of mobile phone software: to develop educational software based on artificial intelligence compatible with all types of mobile phones to meet the needs of students who have limited access to smart devices.
2	Privacy and data security	The collection and storage of large amounts of student data for AI-powered personalization poses concerns about maintaining privacy and data security.	Clarification of artificial intelligence information: Governments can apply a transparent policy regarding the collection and use of student information and ensure that this information is collected only for the purposes of virtual education. Network security: in order to protect students' information from unauthorized access, implement advanced encryption algorithms to control access to information.
3	Bias in AI-powered algorithms	Error: AI-based systems may mistakenly choose orientations, which suddenly lead to unequal actions for students based on multiple factors such as gender, or socio-economic status.	Diverse education systems: ensuring the use of educational models based on artificial intelligence by using the latest educational information to make the educational methods uniform and purposeful. The rule of artificial intelligence: artificial intelligence should try to reduce the errors in modern algorithms in intelligent education that are used for educational purposes.
4	Tutor training and support	Many educators may not have undergone the required training to effectively integrate AI tools into virtual classrooms.	Professional development programs (courses): Offer continuing professional development programs to train educators on the application of AI-powered educational tools and platforms. **Peer support networks (PSNs) **: Strengthen PSNs, where educators can share the best procedures and strategies for the integration of AI into VLEs.
5	Overreliance on technology	Overreliance on AI technology may result in reduced human interaction and developed social skills among students.	Equal treatment: Achieving equal treatment in smart e learning. Share AI tools to achieve better and more collaborative learning experiences. Interaction and combination of social-emotional learning: the components of these interactions in the modern educational system cause the development and academic progress of students.
6	Evaluation (Assessment) and responsiveness	Effective evaluation of student progress and responsiveness in VLEs with AI-based personalized learning experiences can be challenging.	Clear evaluation criteria: Define clear evaluation criteria to assess student progress and learning outcomes in AI-powered virtual learning. Continuous monitoring: Implement some continuous monitoring and feedback systems to ensure responding to students and educators' needs.

DOI: https://doi.org/10.54216/PMTCS.040201
Received: January 24, 2024 Revised: May 14, 2024 Accepted: August 08, 2024

14. Conclusion

By addressing the aforementioned challenges by adopting thoughtful strategies and preemptive measures, the integration of AI into VLEs (table 13) can result in more inclusive, personalized, and effective learning experiences for students, while supporting educators in their instructional roles.

The e-learning and distance education processes are complex with certain parameters that require using different indicators and parameters on the side of the educator (trainer) and learner (trainee) for a more detailed and extensive analysis of the features and criteria affecting distance education. Governments, higher education institutions, and various organizations have focused heavily on the above issue. To increase user satisfaction with their services, many of them are attempting to adopt certain solutions to bring e-learning quality closer to face-to-face training and to fill the existing gaps in this field. Hence, the integration of cognitive IoT in e learning and distance education systems will lead to increased student understanding and concentration. Besides, the professors will be able to present the subjects with more facilities and concentration.

The studies conducted show that the speed of expansion and penetration of AI in higher education is affected by the influence of different factors. In the research conducted by (International Data Corporation), the most important drivers of the use of artificial intelligence are:

- Improving the efficiency of education
- Better integration of students in the process
- Higher competitiveness
- Accelerating the innovation process
- More funding

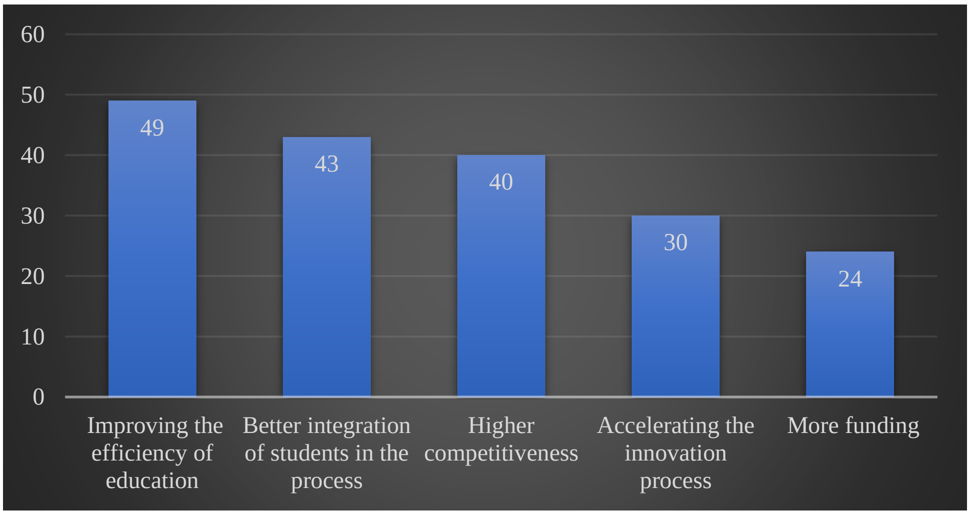

Figure 10. Drivers of using AI in education (%).

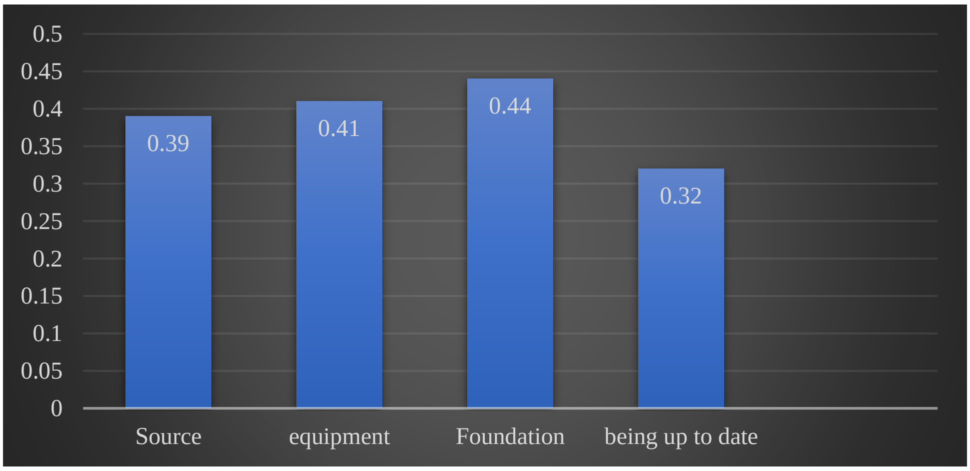

Figure 11. The results of the t-test: sub-components of the infrastructure (path coefficient β)

DOI: https://doi.org/10.54216/PMTCS.040201
Received: January 24, 2024 Revised: May 14, 2024 Accepted: August 08, 2024

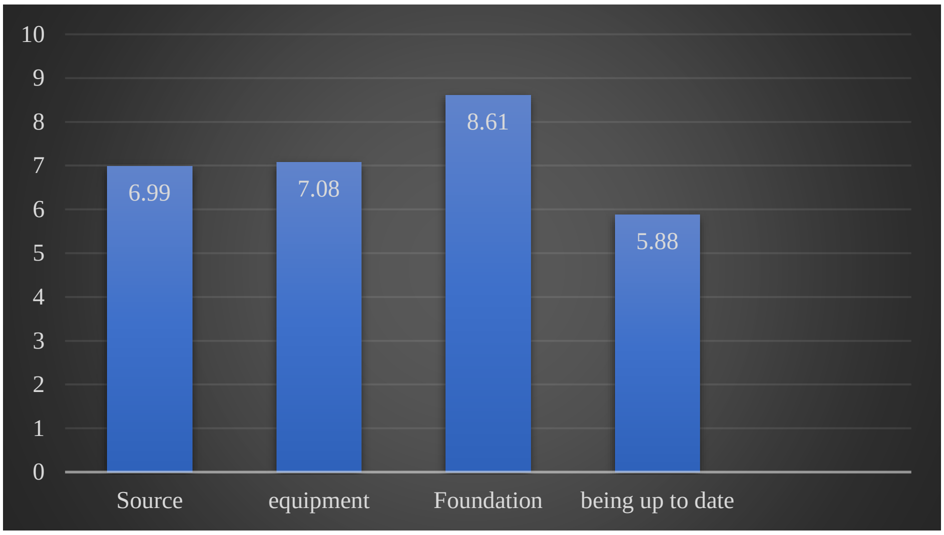

Figure 12. The results of the t-test: sub-components of the infrastructure (t-test)

Based on this, the effect of all identified sub-components related to the sub-structure category of the item Confirmed.

15. Recommendations for Future Research

- For future research, the above problem can be combined with other criteria and parameters to boost the quality and student understanding of the curriculum. For example, several emotional factors, e.g., news and educational policies, can be taken into account during the evaluation of the quality of education using IoT.
- Adopting a solution that takes into account different personal and environmental aspects with different applications in various institutions and low startup costs.

Funding: “This research received no external funding”

Conflicts of Interest: “The authors declare no conflict of interest.”

References

[1] Dunai L, Martins J, Umetani K, Lucia O, Ibrahim Y, Appuhamillage GK. E-Learning in Industrial Electronics during Covid-19. 2021 22nd IEEE International Conference on Industrial Technology (ICIT) 2021. https://doi.org/10.1109/icit46573.2021.9453467.

[2] T. B. L, P. S. S. An Analysis on E-Learning and Its Recommendations. Advances in Data Mining and Database Management 2021:166–87. https://doi.org/10.4018/978-1-7998-8061-5.ch009.

[3] Habes M, Ali S, Khalid A, Haykal HA, Elareshi M, Khan T, et al. E-Learning Acceptance During the Covid-19 Outbreak: A Cross-sectional Study. Lecture Notes in Networks and Systems 2021:65–77. https://doi.org/10.1007/978-3-030-77246-8_7.

[4] Basha RFK, Venusamy K. A Detailed Study Of Effective Online Technical Teaching Aids For Augmented Learning in Covid’ 19 Pandemic. 2020 Sixth International Conference on E-Learning (Econf) 2020. https://doi.org/10.1109/econf51404.2020.9385485.

[5] Dung HTT, Hai TQ. Changes in students’ experiences and perceptions towards e-learning at Hoa Sen University during Covid-19 pandemic. Hoang, TTD, & Tran, QH (2021) Changes in Students’ Experiences and Perceptions towards E-Learning at Hoa Sen University during Covid-19 Pandemic AsiaCALL Online Journal 2021;13:22–39.

[6] Shinghal K, Saxena A, Saxena N, Misra R. IOT Based Modified Hybrid Blended Learning Model for Education. 2020 International Conference on Advances in Computing, Communication & Materials (ICACCM) 2020. https://doi.org/10.1109/icaccm50413.2020.9213049.

[7] Popchev I, Orozova D, Stoyanov S. IoT and Big Data Analytics in E-Learning. 2019 Big Data, Knowledge and Control Systems Engineering (BdKCSE) 2019. https://doi.org/10.1109/bdkcse48644.2019.9010666.

[8] Mohammed CM, Askar S. Machine learning for IoT healthcare applications: a review. International Journal of Science and Business 2021;5:42–51.

DOI: https://doi.org/10.54216/PMTCS.040201
Received: January 24, 2024 Revised: May 14, 2024 Accepted: August 08, 2024

[9] Ilieva G, Yankova T. IoT in Distance Learning during the COVID-19 Pandemic. TEM Journal 2020:1669–74. https://doi.org/10.18421/tem94-45.

[10] Lamri M, Akrouf S, Boubetra A, Merabet A, Selmani L, Boubetra D. From local teaching to distant teaching through IoT interoperability. 2014 International Conference on Interactive Mobile Communication Technologies and Learning (IMCL2014) 2014. https://doi.org/10.1109/imctl.2014.7011115.

[11] Rukmana AA, Mulyanti B. Internet of Things (IoT): Web learning for smart school system. IOP Conference Series: Materials Science and Engineering 2020;830:32042. https://doi.org/10.1088/1757-899x/830/3/032042.

[12] Lea R, Blackstock M. City Hub: A Cloud-Based IoT Platform for Smart Cities. 2014 IEEE 6th International Conference on Cloud Computing Technology and Science 2014. https://doi.org/10.1109/cloudcom.2014.65.

[13] Shamshoddin S, Khader J, Gani S. Predicting consumer preferences in electronic market based on IoT and Social Networks using deep learning based collaborative filtering techniques. Electronic Commerce Research 2019;20:241–58. https://doi.org/10.1007/s10660-019-09377-0.

[14] Rahmani AM, Ali Naqvi R, Hussain Malik M, Malik TS, Sadrishojaei M, Hosseinzadeh M, et al. E-Learning Development Based on Internet of Things and Blockchain Technology during COVID-19 Pandemic. Mathematics 2021;9:3151. https://doi.org/10.3390/math9243151.

[15] Kim J, Yun J, Choi S-C, Seed DN, Lu G, Bauer M, et al. Standard-based IoT platforms interworking: implementation, experiences, and lessons learned. IEEE Communications Magazine 2016;54:48–54. https://doi.org/10.1109/mcom.2016.7514163.

[16] Bayani M, Leiton K, Loaiza M. Internet of things (IoT) advantages on e-learning in the smart cities. International Journal of Development Research 2017;7:17747–53.

[17] SİRİPONGDEE K, PIMDEE P, TUNTIWONGWANICH S. A blended learning model with IoT-based technology: effectively used when the COVID-19 pandemic? Journal for the Education of Gifted Young Scientists 2020;8:905–17. https://doi.org/10.17478/jegys.698869.

[18] Jasim NA, Salim AlRikabi HT, Farhan MS. Internet of Things (IoT) application in the assessment of learning process. IOP Conference Series: Materials Science and Engineering 2021;1184:12002. https://doi.org/10.1088/1757-899x/1184/1/012002.

[19] Pandey D, Singh N, Singh V, Khan MW. Paradigms of Smart Education with IoT Approach. EAI/Springer Innovations in Communication and Computing 2021:223–33. https://doi.org/10.1007/978-3-030-77528-5_11.

DOI: https://doi.org/10.54216/PMTCS.040201

Received: January 24, 2024 Revised: May 14, 2024 Accepted: August 08, 2024