Journal of Applied Data Sciences
Vol. 5, No. 3, September 2024, pp. 1210-1222

Logistic Regression Analysis of Factors that Influence User Experience in Student Medical Report Applications

Tenia Wahyuningrum[1,*], Novian Adi Prasetyo[2], Gita Fadila Fitriana[3], Dimas Fanny Hebrasianto Permadi[4], Rr. Setyawati[5], Joewandewa Yuliansyah[6], Khoem Sambath[7]

[1,2,4,6]*Department of Informatics, Institut Teknologi Telkom Purwokerto, Banyumas, Indonesia*

[3]*Department of Software Engineering, Institut Teknologi Telkom Purwokerto, Banyumas, Indonesia*

[5]*Department of Psychology, Universitas Muhammadiyah Purwokerto, Banyumas, Indonesia*

[7]*Department of Computer Science, National Polytechnic Institute of Cambodia, Phnom Penh, Cambodia*

(Received: May 23, 2024; Revised: June 26, 2024; Accepted: July 19, 2024; Available online: August 8, 2024)

Abstract

Monitoring student health efficiently requires collaboration between schools and government health services. Traditional methods often need more agility and user-friendliness, leading to delays and inaccuracies. This research aims to verify a fast and agile student medical report that we have previously developed using the Modified Agile User Experience (UX) method, with a focus on simplicity, usability, and accessibility. The system's evaluation employs non-functional testing methods to identify factors influencing user satisfaction within the scope of the user experience. We measure task-level and overall user satisfaction using the Single Ease Questions (SEQ) questionnaire as the response variable. This study also investigates test-level satisfaction as predictor variables using Usability Metric for User Experience (UMUX) and UMUX-Lite questionnaire as predictor variables, as well as each student's Interest in learning and learning motivation concerning test-level satisfaction. Binary Logistic Regression (BLR) analysis determined the relationship between test-level and task-level satisfaction, revealing significant correlations between these variables. Based on the results, the Interest to Learn variable is the most important factor that influences task-level satisfaction, but with a small probability value (42.9%). To ensure these accurate results, we changed the scale on SEQ from Easy and Hard to seven scales with normalized values. We compared the results using 4 algorithms: Logistic Regression, Random Forest, Support Vector Machine (SVM), and Gradient Boosting as the most effective model. For a test size of 0.2 and a random state of 40, Logistic Regression achieved an accuracy of 0.80 and a Receiver Operating Characteristics (ROC) and Area Under Curve (AUC) score of 0.83. Random Forest also had an accuracy of 0.80 but a slightly lower ROC AUC score of 0.77. SVM also performed well, with accuracies of 0.83 and ROC AUC scores of 0.77. Gradient Boosting showed the lowest performance with an accuracy of 0.77 and a ROC AUC score of 0.73. These results indicate that Logistic Regression is the most robust model for predicting user satisfaction. Significant data correlations between SEQ, UMUX, and UMUX-Lite guide the development of user-centered applications, enhancing the effectiveness of educational tools by ensuring higher user satisfaction. Future research should consider more extensive, more diverse samples and additional factors influencing user experience to refine these models and their applications.

Keywords: Agile UX Method, Usability Metric for User Experience, Single Ease Question, Logistic Regression

1. Introduction

Monitoring student health requires collaboration between schools and government health services, necessitating special treatment and continuous, regular efforts to obtain an overview of student health status [1]. In Indonesia, the health screening process for school-aged children is a routine activity conducted every semester involving schools and public health centers [2]. Providing a health reporting system for elementary school students, teachers, and health center doctors can serve as a primary source for recording cases, enabling medical officers to observe and identify health issues more efficiently [3]. Additionally, health information from schools in a particular area can facilitate information exchange and communication between different health services [4]. Furthermore, the physical and mental health of students can be detected by those closest to them, such as parents and teachers [5], [6]. Mental health plays a crucial role in student learning success, especially after the COVID-19 pandemic, which has impacted all aspects of human

***Corresponding author: Tenia Wahyuningrum (tenia@ittelkom-pwt.ac.id)**
DOI: https://doi.org/10.47738/jads.v5i3.285

life [7]. To develop a student health reporting system, despite limited resources, costs, and a small team, we require simple usability and accessibility, leading us to choose the Modified Agile UX method for development [8].

The modified Agile UX is onto the design sprint process at each iteration, facilitating continuous improvement [9]. Three fundamental principles demonstrate how agile methods enhance UX design: team member interaction, user feedback and analysis, and incremental project completion. Agile UX modifications can accelerate processing time by focusing on critical functions and producing a minimum viable product. We accelerated the sprint process from the original five stages [10], [11], [12], [13] to three stages, which are the Ideate, Create, and Iterate stages in developing the student medical report application.

In previous research, we successfully developed a student medical report using the Modified Agile UX method, with a satisfaction score using the UMUX-Lite questionnaire 74.9. However, we have yet to test the factors that influence system user satisfaction, both at the test level and task level [9]. In this research, we carried out iterative stages by evaluating and testing the system on users using non-functional testing methods. This testing is usually related to software quality [14]. This non-functional testing is to find out how far the user's experience with the application is. This test also examines the factors that influence user satisfaction measures within the scope of user experience. We measure the influence of task level satisfaction on test level satisfaction using the SEQ questionnaire [15] and UMUX [16], which represents the level of task satisfaction. SEQ is usually used to see how difficult a given task is based on the user's perspective.

Meanwhile, UMUX as test-level satisfaction, collects feedback regarding overall user satisfaction with the overall application testing experience. The advantage of the UMUX measuring tool compared to the System Usability Scale (SUS) is that UMUX only focuses on 4 aspects, while SUS focuses on 10 assessment aspects [17], [18], [19]. The SUS method assesses perceived usability and learnability, while UMUX targets usability in terms of effectiveness, efficiency, and satisfaction. UMUX targets usability similar to that obtained by SUS, regulated according to usability according to ISO series 9241-11[16]. So, it can be said that UMUX is concise enough to function as a usability factor and broader user experience metric. UMUX is a short qualitative assessment using an ordinal scale, or it can be concluded that UMUX is a mini version of SUS. UMUX-Lite, developed by [20], [21] is a short version containing only positive statements. Factor analysis conducted by [20] shows that UMUX has a two-dimensional structure with an alignment of positive and negative items. This condition then led to the selection of two items for UMUX-Lite to create an ultra-short metric for user-perceived usability. We also investigated including the predictor variables of each student's Interest in learning and learning motivation regarding satisfaction [22], [23].

To determine the relationship between test-level satisfaction and task-level satisfaction, we used BLR analysis, which is an analysis method that uses categorical or continuous variables. It provides interpretable coefficients indicating the impact of predictors on the likelihood of the outcome. In these studies, BLR can determine essential factors in decision-making or future predictions regarding a case based on dichotomous or binary data. So, this research aims to develop a fast and agile system for building a student health reporting system, employing the Agile UX method to ensure simple usability and accessibility. By evaluating and testing the system using non-functional testing methods, we aim to identify the factors that influence user satisfaction within the scope of user experience.

Furthermore, BLR analysis will be used to determine the relationship between test-level satisfaction and task-level satisfaction, providing insights into key factors influencing user experience. The insights gained from this study are intended to guide the development of more user-centered applications, ultimately enhancing the effectiveness of educational tools. To confirm the research results, we compared them with other methods, such as the Logistic Regression, Random Forest, SVM, and Gradient Boosting models.

2. The Proposed Method

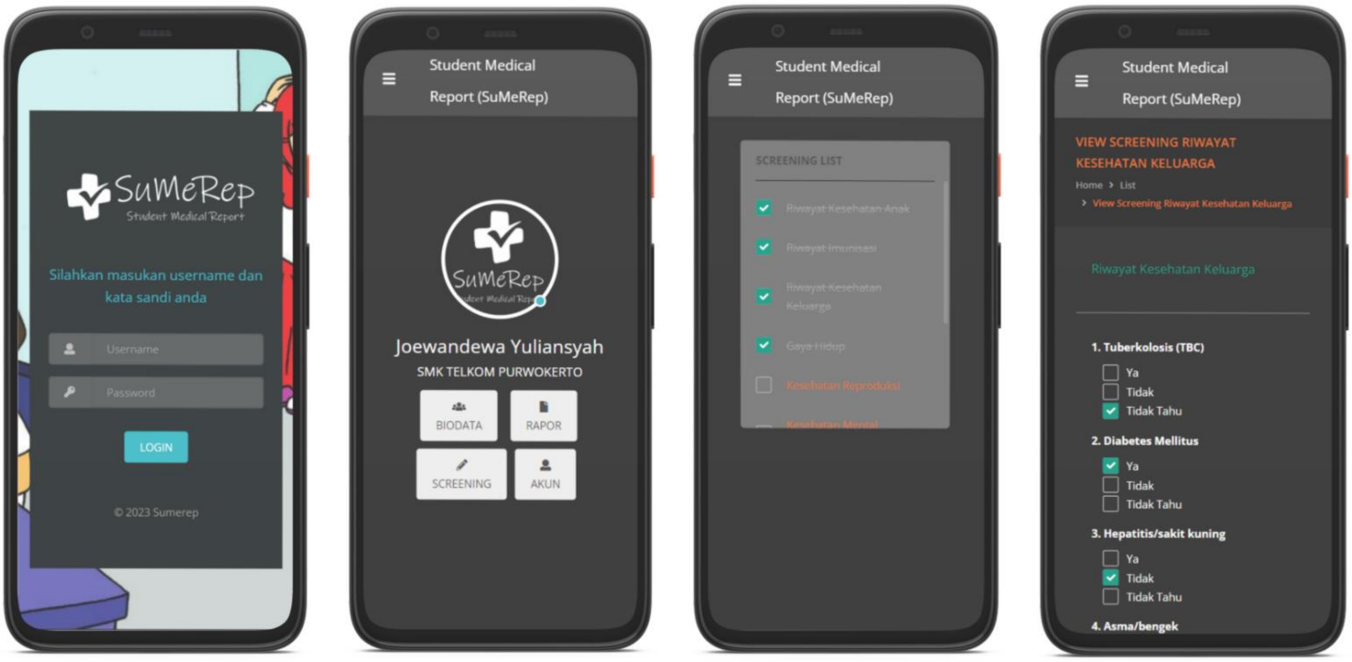
2.1. Data Source

The data used in this research is secondary data from the results of measuring the user experience of student users in operating the student medical report application. We have successfully developed an application using the Modified Agile-UX method [9] with a user interface (UI) display, as described in figure 1. In this study, 149 students from junior high school Telkom Purwokerto in 8th grade were involved in the measurements, with details of 109 male students and

ISSN 2723-6471

40 female students. The students' age range is between 14 and 16 years old, and they are under teacher guidance. Previously, students were asked to carry out two tasks: 1) fill out and edit the biodata form, and 2) fill out and edit the physical and mental health screening form independently (Interest in learning, motivation to learn, knowledge of reproductive health, child and family health history). The UI uses Indonesian as the language of instruction in the system.

Figure 1 shows the student medical report UI, which consists of a login page, menu page, report page, and screening page. The UI communicates with users using Indonesian. On the login page, students need to enter their username and password according to what they have previously registered (Figure. 1. (a)). The menu page has four main features: health report, self-screening form, biodata form, and user account (Figure. 1. (b)). Users can select one of the features and return to the main menu or the burger menu by clicking back. In Figure 1 (c), the students can see their health reports during school. Students also get feedback from health workers through recommendations; if the illness they suffer is severe enough, students can see a doctor's referral to the nearest hospital for follow-up. Figure 1 (d) shows one of the independent screening forms that students can fill out: the psychological well-being form. From the results of student input, health workers and teachers can provide recommendations regarding their physical and mental health.



(a). login page (b). menu page (c). report page (d). screening page

Figure 1. Student UI

2.2. Variables

The target variable used for this research is the response variable (Y), which is used for SEQ Score. The feature variables are UMUX (X_1), UMUX-Lite (X_2), Interest to Learn (X_3), and Motivation to Learn (X_4). The normalization method is used as expressed in equations (1) and (2). SEQ scale is divided into 7 scales. The scales are between 1 to 7 from the lowest point until the highest point. In every point, the name is Strongly Disagree, Disagree, Somewhat Disagree, Neither Disagree nor Agree, Somewhat Agree, Agree, Strongly Agree. This scale is normalized into 0-100 scales based on equation (1). So, the scales become respectively into 0, 16.67, 33.33, 50, 66.67, 83.33, and 100. For the purposes of calculating BLR, we categorize binary values 0 and 1; the Y variable into "Hard or 0" for values 0.00, 16.67, 33.33, and 50.00 and "Easy or 1" for values 66.67, 83.33, and 100.00.

$$\mathrm{Score_{norm}} = \frac{(\mathrm{Score_{raw}} - 1)}{6} \times 100 \tag{1}$$

The variables for Interest to Learn (X_3) and Motivation to Learn (X_4) are scales using the Likert Scale for every question. Every question uses a score from 1-5, and the number of every Variable is 12. The questions are about how many students are interested in learning and how many students are motivated about learning. So, the lowest and the highest scores for every Variable were about 12 and 60, respectively. This score is converted into 3 scales with a minimum score of 0-28 for Low, then 29-44 for Medium, and 45-60 for High. So, the normalized scale using equation (2). For example, a raw score of 3 on a Likert scale of 1-5 will be normalized to 50.

$$\text{Norm} = 100 \times \frac{\text{Raw}_{\text{value}} - \text{Min}_{\text{value}}}{\text{Max}_{\text{value}} - \text{Min}_{\text{value}}} \tag{2}$$

Meanwhile, the variables UMUX and UMUX-Lite Score are not normalized because this score uses a scale of 0-100. So, the variables are respectively represented as (X_1) and (X_2). The detailed variable research is shown in table 1.

Table 1. Research variables

Variable	Description	Categories	Scale
Y	Task Level Satisfaction (by SEQ score)	0: Hard 1: Easy	Ordinal
X_1	Test Level Satisfaction (by UMUX score)	0-100	Ratio
X_2	Test Level Satisfaction (by UMUX-Lite score)	0-100	Ratio
X_3	Interest to Learn	0-100	Ratio
X_4	Motivation to Learn	0-100	Ratio

3. Method

In his study, the regression method establishes the connection between the response variable (Y) and one or more predictor variables (X). For more detail, we used BLR to illustrate how dichotomous variables are related, specifically variables with only two categories (on a nominal or ordinal) scale and a set of continuous or categorical predictor variables.

3.1. Data Correlation

Correlation is a statistical measure that describes the degree to which two variables move in relation to each other. It is a common way to understand the relationship between variables, which can be positive, negative, or zero. The correlation matrix revealed the relationships between the different variables, showing how changes in one Variable might be associated with changes in another. When one variable increases, the other variable also increases. Also, a correlation coefficient greater than 0 is called Positive Correlation. When one variable increases, the other variable decreases. Also, a correlation coefficient of less than 0 is called a negative correlation. The last one is there is no relationship when the coefficient is around 0. The correlation matrix provides insights into the linear relationship between variables. The correlation coefficient (corr) ranges from -1 to 1, as seen in equation (3). For the Variable x_i is the individual sample value of variable x and y_i is the individual sample value of variable y. The Variable $\bar{x}$ is the mean of variable x and the variable $\bar{y}$ is the mean of variable y. Those variables include correlation about variable SEQ Score (Y), UMUX (X_1), UMUX-Lite (X_2), Interest to Learn (X_3), and Motivation to Learn (X_4).

$$\text{corr} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}} \tag{3}$$

3.2. Binary Logistic Regression

We performed a regression analysis to predict SEQ_Score based on the normalized values of Interest to Learn, Motivation to Learn, UMUX score, and UMUX-Lite score. The BLR equation used in the form of an interpretation of the probability function $\pi(x) = (Y|x)$ is expressed in the form of the following equation (4).

$$\pi(x) = \frac{e^{\beta_0+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_p x_p}}{1 + e^{\beta_0+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_p x_p}} \tag{4}$$

Then, a logit transformation is carried out to simplify equation (5) in logit form as follows

$$g(x) = \ln\left[\frac{\pi(x)}{1-\pi(x)}\right] = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p \tag{5}$$

Equation (5) is a logit model, where the function is a linear function of the parameters.

3.3. Model Suitability Test

Model suitability testing is carried out to determine whether there is a difference between the observation results and the possible model prediction results. The model suitability test hypothesis expressed the null hypothesis (H_0), which is that the model is fit (there is no significant difference between the observed results and the possible model predictions). Alternative hypothesis (H_1) is the model does not fit (there is a significant difference between the observed results and the possible model predictions).

Using equation (6), it can be seen that the Hosmer Leme show test statistic follows a Chi-Square distribution with degrees of freedom g-2. If a significance level of α is determined, then a decision to reject H_0 is obtained if the value of $\hat{C} > \chi^2_{(a,g-2)}$.

$$\hat{C} = \sum_{l=1}^{g} \frac{(o_l - n_l'\bar{\pi}_l)^2}{n_l'\bar{\pi}_l(1-\bar{\pi}_l)} \tag{6}$$

3.4. Parameter Significance Test

Parameter estimation testing is a test used to test the significance of the β coefficient of the model. This test can use partial or simultaneous tests.

3.4.1. Simultaneous testing

The hypothesis used in the simultaneous test is as follows: the null hypothesis (H_0) is $\beta 1= \beta 2=\ldots= \beta p=0$ (The predictor variable does not have a significant influence on the response variable). Alternative hypothesis (H_1) is $\beta j \neq 0$ where j=1,2,…, p (There is at least one predictor variable that has a significant influence on the response variable).

Based on equation (7), the G test statistic follows the Chi-Square distribution. If the significance level is determined to be α and the degrees of freedom are v, then it is decided to reject H0 if the value $G > \chi^2_{(a,p)}$.

$$G = -2\ln \frac{\left[\frac{n_1}{n}\right]^{n_1}\left[\frac{n_0}{n}\right]^{n_0}}{\prod_{j=1}^{n} \hat{\pi}_j^{y_j}[1-\hat{\pi}_j]^{1-y_j}} \tag{7}$$

3.4.2. Partial testing

The hypothesis used in the partial test is as follows: the null hypothesis (H_0) is $\beta j= 0$ (The predictor variables partially do not have a significant influence on the response variable). Alternative hypothesis (H_1) is $\beta j \neq 0$ where j=1,2,…, p (There is at least one predictor variable that has a significant influence on the response variable partially).

Based on equation (8), the W^2 test statistic follows the Chi-Square distribution. If the significance level is determined to be α and the degrees of freedom are v, then it is decided to reject H_0 if the value $W^2 > \chi^2_{(a,v)}$.

$$W^2 > \frac{\hat{\beta}_j^2}{[SE(\hat{\beta}_j)]^2} \tag{8}$$

3.5. Model Interpretation

The odds ratio value is used to make it easier to interpret the model, namely the tendency for the response variable to have a specific value if given x=1 and compared to x=0.

$$OR = \frac{\pi(1)/[1-\pi(1)]}{\pi(0)/[1-\pi(0)]} \quad (9)$$

Noted that $\pi(1) = \frac{\exp(\beta_0+\beta_j)}{1+\exp(\beta_0+\beta_j)}$ and $\pi(0) = \frac{\exp(\beta_0)}{1+\exp(\beta_0)}$ where j=1, 2, ..., p. Based on equation (9), the decision that there is no relationship between the predictor variables is made if the odds ratio equals 1. Suppose the odds ratio value is less than 1. In that case, there is a negative relationship between the predictor variable and the response variable every time the value of the independent variable (x) changes, and vice versa.

4. Result and Discussion

4.1. Descriptive Statistic

Descriptive statistics is a part of statistical science that summarizes, presents and describes data in an easy-to-read form so as to provide complete information. Based on table 2, the mean value obtained for each variable SEQ, UMUX, UMUX Lite, Interest to learn, and motivation to learn is 53.13 (easy); 63.09 (high); 62.31 (high); 71.49 (OK), and 72.82 (OK). The minimum and maximum values in the SEQ score show that some users are not satisfied with the task given at all (value 0), but some are very satisfied (value 100). However, the other variables have a range of values that are not too far apart in terms of minimum and maximum values. Meanwhile, the skewness value of almost all variables is close to 0; this condition shows that the data is normally distributed. The standard deviation shows a small value from the mean, which means that the data distribution is normal.

Table 2. Descriptive Statistics

Descriptive Statistics	SEQ score (Y)	UMUX score (X_1)	UMUX-Lite score (X_2)	Interest to Learn (X_3)	Motivation to Learn (X_4)
Mean	53.13	63.09	62.31	71.49	72.82
Median	50.00	66.67	66.23	70.00	70.00
Min	0.00	20.00	22.00	55.00	55.00
Max	100.00	100.00	87.00	100.00	100.00
Standard Deviation	26.31	19.20	16.21	8.26	8.44
Sample Variance	692.05	368.58	262.82	68.21	71.20
Skewness	-0.28	-0.26	-0.56	1.05	0.69

4.2. Model Suitability Test

The Hosmer and Lemeshow Test is a Goodness of Fit test (GoF), which determines whether the model formed is fit or not. The model is fit if there is no significant difference between the model and the observed values. Based on Table 3, the sig value is 0.233>0.05, and the Hosmer and Lemeshow Chi-Square value calculated is 10.483< Chi-Square table 15.507 (df is 8). It can be concluded that H0 is accepted (the model is fit), meaning that the BLR model is suitable for use for further analysis because there is no real difference between the predicted probabilities and the observed classifications.

4.3. Parameter Significance Test

4.3.1. Coefficient of determination test

In the next analysis, the Cox and Snell R Square and Nagelkerke R Square values were used to determine the ability of the predictor variable to respond to the response variable. The Nagelkerke R Square value in table 3 is 0.548, which shows that the ability of the independent variable to explain the dependent Variable (Test Level Satisfaction) is 54.8%, and the rest is explained by other variables not studied. In BLR analysis, the Nagelkerke R Square value is 0.548; this value indicates that the model has a good fit but could be better. This value can be used to compare with other models to see whether one model is better at explaining the data than the other model.

Table 3. Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	99.933a	0.372	0.548

a. Estimation terminated at iteration number 7 because parameter estimates changed by less than 001.

4.3.2. Simultaneous Test

The Omnibus Test of model coefficients in BLR is a statistical test used to evaluate the entire logistic regression model. This test assesses whether the model as a whole is significant in predicting the dependent Variable. The result calculation with the degree of freedom is 5, generating the Chi-Square is 69.271, and the Sig value is 0.000 (less than 0.05), which means H_0 is rejected and H_1 is accepted. It can be concluded that at least one predictor variable in the model has a significant relationship with task-level satisfaction as the response variable. Omnibus test of model coefficient is shown in table 4.

Table 4. Omnibus Test of Model Coefficient

		Chi-square	df	Sig.
Step 1	Step	69.271	5	0.000
	Block	69.271	5	0.000
	Model	69.271	5	0.000

4.3.3. Partially Test

Next, a partial test was carried out to find out which predictor variables had a significant effect on the test level satisfaction. In this test, a variable is declared to have a significant effect if the significance value is less than α of 0.05 is shown in table 5.

Table 5. Variables in the Equation

Variables		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1a	UMUX score	0.086	0.086	1.010	1	0.315	1.090
	UMUX-Lite score	-0.161	0.089	3.236	1	0.072	0.851
	Interest to Learn	-0.079	0.037	4.627	1	0.031	0.924
	Motivation to Learn	0.031	0.038	0.680	1	0.410	1.032
	UMUX by UMUX-Lite	0.001	0.001	1.361	1	0.243	1.001
	Constant	-0.203	5.727	0.001	1	0.972	0.817

a. Variable(s) entered on step 1: UMUX, UMUX-Lite, Interest to Learn, Motivation to Learn, UMUX * UMUXLite. Significance α=5%.

According to table 5, the variable that significantly affects test-level satisfaction is task-level satisfaction, represented by the Interest to Learn. It can be seen that the significance value approaching 0.05. The variables of UMUX Score, UMUX-Lite Score, and Motivation to Learn do not significantly influence test-level satisfaction.

4.4.2. Model Interpretation

The BLR model formed based on the results of parameter testing is as follows.

$$g(x) = \ln\left[\frac{\pi(x)}{1-\pi(x)}\right] = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 = -0.203 - 0.079x_3 \quad (10)$$

The odds ratio value makes it easier to interpret the BLR model, where it is known from the results of the partial parameter significance test that the most influencing Variable is Interest to Learn. The odds ratio or Exp (B) value of the Interest to learn Variable is 0.924 (almost 1). This indicates that users with a high level of Interest in learning will increase their task satisfaction level by 1-fold. The probability or predicted BLR model with the following coefficients is 42.9%; this probability value is low. Therefore, it is necessary to use other methods to explore important factors in task-level satisfaction, such as Logistic Regression.

$$\text{probability} = \frac{\exp((-0.203)+(-0.079\times 1))}{1+\exp((-0.203)+(-0.079\times 1))} = 0.429 \quad (11)$$

4.5. Classification Test

In classification performance, shown in table 6, it is known that the number of students who had an Easy perception when doing assignments was 111 people, 10 of whom were predicted to have a Hard experience with a prediction correctness level of 91.0%. Meanwhile, the number of students who did not perform well was 38 people, 10 of whom were predicted to have Easy experiences, with a prediction correctness level of 73.7%. So, the percentage accuracy of the model that can be predicted correctly is 86.6% (good classification).

Table 6. Classification Table

Observed			Predicted		
			SEQ		Percentage Correct
			Easy	**Hard**	
Step 1	SEQ	Easy	101	10	91.0
		Hard	10	28	73.7
	Overall Percentage				**86.6**

a. The cut value is ,500

4.6. Comparison Classification Test

To enrich the findings further, we have carried out a comparison of classification tests. As seen in the model interpretation, the probability is low (42,9%). We tried to dig deeper, using alternative methods to see the most significant factors influencing task level satisfaction, as well as those that have the highest accuracy in classification. In this study, four machine learning models were evaluated to predict whether SEQ_Score is high (>50). The models compared are Logistic Regression, Random Forest, Support Vector Machine (SVM), and Gradient Boosting. The data we use is as in Table 1. Still, we change the SEQ variable into ratio size data (0-100) using normalization techniques (Eq. 1). Logistic Regression in previous research is widely used in research in the health sector [24], in the field of economics [25], and in the field of government [26]. Logistic Regression was conducted to predict whether the SEQ score is greater than 50. The logistic Regression provided coefficients, intercept, confusion matrix, classification report, ROC and AUC score, indicating the model's performance in binary classification. While the initial comparison of machine learning models (Logistic Regression, Random Forest, SVM, and Gradient Boosting) was based on accuracy, further analysis highlighted several reasons for selecting Logistic Regression as the most effective method.

Interpretability is logistic Regression provides transparent and interpretable coefficients that indicate the direction and magnitude of the effect of each predictor variable, which is crucial for understanding and explaining the model. Performance is although other models like Random Forest and Gradient Boosting performed well, Logistic Regression showed comparable accuracy and ROC AUC scores, making it a reliable choice. Simplicity is logistic Regression is computationally less intensive and more accessible to implement, requiring fewer hyperparameter tuning efforts compared to more complex models like SVM and Gradient Boosting. And robustness is the model demonstrated robustness across different test sizes, maintaining consistent performance, which is essential for real-world applications. Here, we provide a detailed explanation of each model's performance based on accuracy, ROC AUC score, and other metrics is shown in table 7.

Table 7. Summary Metric of Test Sizes and Random States

Model	Accuracy	ROC AUC Score	Precision (Class 0)	Recall (Class 0)	F1-Score (Class 0)	Precision (Class 1)	Recall (Class 1)
Logistic Regression	0.80	0.83	0.75	0.86	0.80	0.86	0.75
Random Forest	0.80	0.77	0.72	0.93	0.81	0.92	0.69
SVM	0.83	0.77	0.76	0.93	0.84	0.92	0.75
Gradient Boosting	0.77	0.73	0.71	0.86	0.77	0.85	0.69

The performance of four different models (Logistic Regression, Random Forest, SVM, and Gradient Boosting) was evaluated under various test sizes and random states. The evaluation metrics included accuracy, precision, recall, F1-score, and ROC AUC score. These metrics offer a comprehensive view of each model's ability to classify instances correctly, balance between precision and recall, and distinguish between classes. The performance of four different machine learning models—Logistic Regression, Random Forest, SVM, and Gradient Boosting—was evaluated using test sizes 0.2 and random states 40. These metrics provide a comprehensive view of each model's ability to classify instances correctly, balance between precision and recall, and distinguish between classes. The results can be seen in table 7, and the comparison graph can be seen in figure 2.

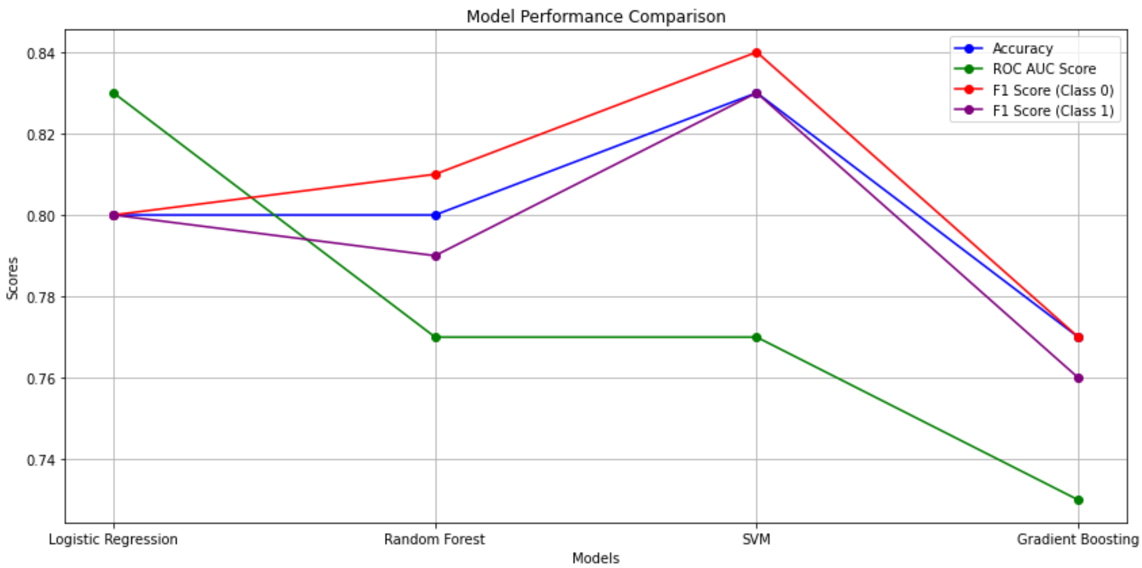


Figure 2. Model Performance Comparison

In the classification models, Logistic Regression achieved an accuracy of 0.80 and a ROC AUC score of 0.83, with a balanced precision and recall for both classes. Random Forest also had an accuracy of 0.80 but a slightly lower ROC AUC score of 0.77, showing high precision and recall for class 0 (0.72 and 0.93) and lower for class 1 (0.92 and 0.69). SVM outperformed slightly with an accuracy of 0.83 and a ROC AUC score of 0.77, with balanced precision and recall. Gradient Boosting showed the lowest performance with an accuracy of 0.77 and a ROC AUC score of 0.73, with slightly lower precision and recall for class 1. Random Forest also performed well but showed variability between classes, while Gradient Boosting lagged. These findings suggest focusing on logistic Regression and SVM for future model developments, given their robust and balanced performance.

4.7. Data Correlation

Based on the result of the performance comparison to see in detail the data correlation between variables using Logistic Regression with the results in figure 3.

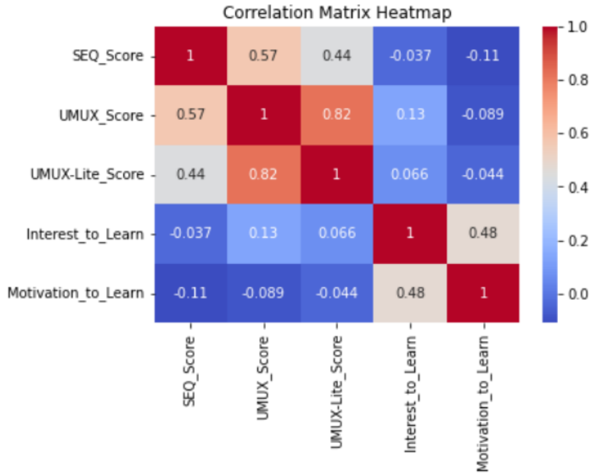


Figure 3. Matrix Data Correlation

Correlation analysis reveals significant relationships between SEQ Score, UMUX Score, and UMUX-Lite Score. These findings indicate that these metrics consistently measure similar aspects of user satisfaction and usability. The correlation between the variables: SEQ Score (Y), UMUX Score (X_1), UMUX-Lite Score (X2), Interest to Learn (X3), and Motivation to Learn (X4). Based on the result, SEQ Score and UMUX Score correlate 0.57, indicating a moderate positive relationship, suggesting that as SEQ scores increase, UMUX Score also tends to grow. SEQ Score and UMUX-Lite Score for the correlation of 0.44 show a low positive relationship. It can also be seen that the UMUX Score and UMUX-Lite Score have the strongest positive correlation of 0.82, indicating that these two measures are closely related and often increase together. The last, Interest to Learn and Motivation to Learn, for the correlation of 0.48, indicates a moderate positive relationship between Interest to Learn and Motivation to Learn. Other correlations between these variables and the scores are relatively low, indicating weaker or negligible relationships. For example, learning motivation has a weak negative correlation with -0.11, suggesting a minor inverse relationship between SEQ scores and student's learning motivation. Lastly, learning interest has a very weak negative correlation with -0.037, implying almost no relationship between SEQ scores and student interest in learning.

4.8. Logistic Regression

Logistic Regression was chosen due to its high ROC AUC score (0.83), indicating its effectiveness in balancing precision and recall. Although Random Forest achieved the highest accuracy (0.80), Logistic Regression's interpretability and ability to provide probabilistic predictions made it more suitable for this study. Additionally, the logistic model's coefficients offer insights into the relationship between UX scores and student satisfaction.

The regression analysis reveals that UMUX has the most significant positive impact on SEQ with a coefficient of 0.98, whereas UMUX_Lite negatively impacts SEQ with a coefficient of -0.42. The intercept of the regression model is 11.98. The mean squared error (MSE) is 354.33, indicating the average squared difference between observed and predicted values, and the R-squared (R^2) value is 0.48, meaning that the features explain 48% of the variance in SEQ. In Logistic Regression, the model achieved an accuracy of 0.73 and a ROC AUC score of 0.82. The confusion matrix shows 13 true negatives, 4 false positives, 4 false negatives, and 9 true positives. The classification report indicates balanced performance with precision and recall at 0.73 for both classes. This result can be seen in figure 4.

ISSN 2723-6471


```
Regression Analysis for predicting Score_SEQ:
Coefficients: [ 0.97854848 -0.01294097 -0.41859815  0.13937569]
Intercept: 11.979443168348034
Mean squared error: 354.33
Coefficient of determination (R^2): 0.48
Logistic Regression:
Accuracy: 0.80
ROC AUC Score: 0.83
Confusion Matrix:
[[12  2]
 [ 4 12]]
Classification Report:
              precision    recall  f1-score   support
           0       0.75      0.86      0.80        14
           1       0.86      0.75      0.80        16
    accuracy                           0.80        30
   macro avg       0.80      0.80      0.80        30
weighted avg       0.81      0.80      0.80        30
```

Figure 4. Logistic Regression Results

5. Conclusion

This study successfully demonstrated that SEQ, UMUX, and UMUX-Lite scores can effectively predict user satisfaction with a student health reporting system. Factors influencing SEQ differ in results from BLR and Logistic Regression. This difference is due to different models, interactions, and interpretations; factors that influence the response variable can show different results in binary and multinomial Logistic Regression. Four machine learning models were evaluated to predict whether the Score_SEQ was high (>50), including Logistic Regression, Random Forest, SVM, and Gradient Boosting. Logistic Regression achieved an accuracy of 0.80 and a ROC AUC score of 0.83, with balanced precision and recall for both classes. Random Forest also had an accuracy of 0.80 but a slightly lower ROC AUC score of 0.77, showing high precision and recall for class 0 (0.72 and 0.93) and lower for class 1 (0.92 and 0.69). SVM outperformed slightly with an accuracy of 0.83 and a ROC AUC score of 0.77, with balanced precision and recall. Gradient Boosting showed the lowest performance with an accuracy of 0.77 and a ROC AUC score of 0.73, with slightly lower precision and recall for class 1. These metrics provide a comprehensive view of each model's ability to classify instances correctly, balance precision and recall, and distinguish between classes.

However, the study has several limitations that must be acknowledged. The sample size and scope of the study may limit the generalizability of the findings. Additionally, the data was collected exclusively from users of the student medical report application, which might introduce bias as the sample may not represent the broader student population. The models used, including Logistic Regression, may have limitations in capturing complex, non-linear relationships between variables, which could affect the accuracy of the predictions. Future research should address these limitations by including more significant and diverse samples, exploring additional factors influencing user experience, and implementing advanced machine learning models like deep learning techniques to improve prediction accuracy. An in-depth investigation into the weak correlations between satisfaction scores and learning motivation or interest metrics is also necessary, possibly requiring additional features or alternative metrics.

Furthermore, qualitative methods such as interviews could complement quantitative analysis and provide deeper insights into user experiences. Longitudinal studies to observe changes in user satisfaction over time and usability testing in real-world educational settings are crucial to identifying potential issues that may not be apparent in controlled environments. By addressing these areas, future studies can enhance the understanding of user experience and contribute to the development of more effective and user-centered educational tools.

6. Declaration

6.1. Author Contributions

Conceptualization: T.W., N.A.P., G.F.F., D.F.H.P., R.S., J.Y., and K.S.; Methodology: T.W., N.A.P. and D.F.H.P.; Software: N.A.P., J.Y., D.F.H.P.; Validation: T.W., N.A.P., G.F.F., D.F.H.P., R.S., and K.S.; Formal Analysis: T.W., N.A.P., G.F.F., D.F.H.P., R.S., J.Y., and K.S.; Investigation: G.F.F. and R.S.; Resources: J.Y.; Data Curation: K.S.; Writing Original Draft Preparation: T.W., N.A.P., G.F.F., D.F.H.P., R.S., and K.S.; Writing Review and Editing: T.W., N.A.P., G.F.F., D.F.H.P., R.S., J.Y., and K.S.; Visualization: J.Y.; All authors have read and agreed to the published version of the manuscript.

6.2. Data Availability Statement

The data presented in this study are available on request from the corresponding author.

6.3. Funding

This study was provided funding by LPPM from the Institut Teknologi Telkom Purwokerto for Tenia Wahyuningrum (IT Tel3029/ LPPM-000/Ka. LPPM/ III/2023).

6.4. Institutional Review Board Statement

Not applicable.

6.5. Informed Consent Statement

Not applicable.

6.6. Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] M. Soenarnatalina, Y. Sulistyorini, Mahmudah, and D. Indriani, "The importance of student health record books in school to realize children's health," *Opcion*, vol. 35, no. Special Issue 22, pp. 2899–2921, 2019.

[2] T. Ripursari and B. M. Suhita, "School Children's Health Screening at SDN Bangsal 1 with Puskesmas Pesantren 1 Kediri City," *J. Community Engagem. Heal.*, vol. 6, no. 1, March, pp. 66–72, 2023, doi: 10.30994/jceh.v6i1.419.

[3] C. Ortega-Loubon, C. Culquichicón, and R. Correa, "The Importance of Writing and Publishing Case Reports During Medical Training," *Cureus*, vol. 9, no. 12, pp. 10–12, 2017, doi: 10.7759/cureus.1964.

[4] J. A. Almulhem, "Medical students' experience with accessing medical records in Saudi Arabia: a descriptive study," *BMC Med. Educ.*, vol. 21, no. 1, pp. 1–10, 2021, doi: 10.1186/s12909-021-02715-7.

[5] N. Agustin, F. Roisatul, and M. Karimun, "Developing guidelines for early detection of child and adolescent mental health problems," *Journal of Community Services and Engagements*, vol. 4, no. 1, pp. 152–161, 2023.

[6] T. Wahyuningrum, G. F. Fitriana, A. C. Wardhana, M. A. Sidiq, and D. Wahyuningsih, "Developing Suicide Risk Idea Identification for Teenager (SERIINA) Mobile Apps Prototype using Extended Rapid Application Development," *in 2021 9th International Conference on Information and Communication Technology-Proceedings, ICoICT 2021*, vol. 9, no. 1, pp. 92–97, 2021. doi: 10.1109/ICoICT52021.2021.9527508.

[7] T. Wahyuningrum, N. A. Prasetyo, A. R. Bahtiar, L. Latifah, I. D. Ramadhani, and D. Yunitawati, "Google Trends Data about Mental Health during COVID-19 Pandemic Using Time Series Regression," *in 2022 5th International Seminar on Research of Information Technology and Intelligent Systems-Proceedings, ISRITI 2022*, vol. 5, no. 1, pp. 125–129, 2022. doi: 10.1109/ISRITI56927.2022.10052965.

[8] S. R. Nadikattu, "Integrating User Experience (UX) Development with Agile Software Development Practices," Blekinge Institute of Technology, Thesis, Sweden, 2016.

[9] T. Wahyuningrum, N. A. Prasetyo, G. F. Fitriana, D. F. H. Permadi, I. Puspitasari, and M. Al Fatoni, "Modified Agile User Experience for Developing Student Medical Report," *in The 7th International Conference of SNIKOM 2023-Proceedings, Binjai, Indonesia*, vol. 7, no. 1, , pp. 1–7, 2023.

[10] R. W. Dias Pedro, A. Machado-Lima, and F. L. S. Nunes, "Towards an approach using grammars for automatic classification of masses in mammograms," *Comput. Intell.*, vol. 2020, Special Issue, pp. 1-30, 2020, doi: 10.1111/coin.12320.

[11] J. Knapp, J. Zeratsky, and B. Kowitz, Sprint How To Solve Big Problems and Test New Ideas In Just Five Days. Simon and Schuster, 2016.

[12] H. Sumual, J. Reimon Batmetan, and M. Kambey, "Design Sprint Methods for Developing Mobile Learning Application,"

KnE Soc. Sci., vol. 3, no. 12, pp. 394, 2019, doi: 10.18502/kss.v3i12.4106.

[13] J. Thomas and C. Shin, “Implementing design sprints in the education of industrial designers,” *Des. Princ. Pract.*, vol. 10, no. 1, pp. 59–73, 2016, doi: 10.18848/1833-1874/CGP/v10i01/59-73.

[14] L. Buglione, “Software product quality: Some thoughts about its evolution and perspectives,” *in 20th IMEKO TC4 Symposium on Measurements of Electrical Quantities: Research on Electrical and Electronic Measurement for the Economic Upturn, Together with 18th TC4 International Workshop on ADC and DCA Modeling and Testing, IWADC 2014, Gruppo Utenti Function Point Italia, Italian Software Metrics Association, Italy*, vol. 20, no. 1, pp. 737–742, 2014.

[15] J. Sauro and J. S. Dumas, “Comparison of three one-question, post-task usability questionnaires,” *in Conference on Human Factors in Computing Systems - Proceedings*, vol. 2009, no. 1, pp. 1599–1608, 2009. doi: 10.1145/1518701.1518946.

[16] K. Finstad, “The usability metric for user experience,” *Interact. Comput.*, vol. 22, no. 5, pp. 323–327, 2010, doi: 10.1016/j.intcom.2010.04.004.

[17] J. R. Lewis, B. S. Utesch, and D. E. Maher, “Investigating the Correspondence Between UMUX-LITE and SUS Scores,” *in Conference on HCI International- Lecture Notes In Computer Science, Los Angeles, CA, USA*, vol. 9186, no. 1, pp. 204–211, 2015, doi: 10.1007/978-3-319-20886-2.

[18] J. Brooke, “SUS : A Retrospective,” *J. Usability Stud.*, vol. 8, no. 2, pp. 29–40, 2013, doi: 10.1074/jbc.R115.675280.

[19] J. R. Lewis and J. Sauro, “The Factor Structure of the System Usability Scale,” *in International Conference on Human Centered Design-Proceedings, Berlin, German*, vol. 5619, no. 1, pp. 94–103, 2009.

[20] J. R. Lewis, B. S. Utesch, and D. E. Maher, “UMUX-LITE - When there’s no time for the SUS,” *in Conference on Human Factors in Computing Systems - Proceedings, Paris, France*, vol. 2013, no. 1, pp. 2099–2102, 2013. doi: 10.1145/2470654.2481287.

[21] J. R. Lewis, “Measuring Perceived Usability : The CSUQ , SUS , and UMUX,” *Int. J. Human–Computer Interact.*, vol. 34, no. 12, pp. 1–9, 2018, doi: 10.1080/10447318.2017.1418805.

[22] N. S. M. Satar, “Does e-learning usability attributes correlate with learning motivation ?,” *in 21st AAOU Annual Conference-Proceedings, Kuala Lumpur, Malaysia*, vol. 21, no. 1, pp. 1–14.

[23] P. Zaharias and A. Poylymenakou, “Developing a usability evaluation method for e-learning applications: Beyond functional usability,” *Int. J. Hum. Comput. Interact.*, vol. 25, no. 1, pp. 75–98, 2009, doi: 10.1080/10447310802546716.

[24] Z. A. A. Al-Bairmani and A. A. Ismael, “Using Logistic Regression Model to Study the Most Important Factors Which Affects Diabetes for the Elderly in the City of Hilla / 2019,” *J. Phys. Conf. Ser.*, vol. 1818, no. 1, pp. 1-10, 2021. doi: 10.1088/1742-6596/1818/1/012016.

[25] P. A. Sunarya, U. Rahardja, S. C. Chen, Y. M. Li, and M. Hardini, “Deciphering Digital Social Dynamics: A Comparative Study of Logistic Regression and Random Forest in Predicting E-Commerce Customer Behavior,” *J. Appl. Data Sci.*, vol. 5, no. 1, pp. 100–113, 2024, doi: 10.47738/jads.v5i1.155.

[26] S. N. Nurdiansah and L. Khikmah, “Binary Logistic Regression Analysis of Variables that Influence Poverty in Central Java,” *J. Intell. Comput. Heal. Informatics*, vol. 1, no. 1, pp. 6–8, 2020.