Published *online* on the journal's web page : http://innovatics.unsil.ac.id

Innovation in Research of Informatics (INNOVATICS)

| ISSN (Online) 2656-8993 |

Performance and Effectiveness Evaluation of the National Digital Samsat as a Public E-Government Service Using the PIECES Framework

Rahma Fitria[1], Amanda Syakhila[2], Desvina Yulisda[3], Azham Hussain[4], Arafat Febriandirza[5]

[1,2,3]*Department of Information Systems, Universitas Malikussaleh, Bukit Indah Campus, Lhokseumawe, 24355, Indonesia*
[4]*School of Computing, Universiti Utara Malaysia, Sintok, 06010, Malaysia*
[5]*Research Center for Data and Information Science, National Research and Innovation Agency (BRIN), Bandung, 40135, Indonesia*

[1]*rahmafitria@unimal.ac.id,* [2]*syakhilaamanda@gmail.com,* [3]*desvina.yulisda@unimal.ac.id,* [4]*azham.h@uum.edu.my,* [5]*arafat.febriandirza@brin.go.id*

ARTICLE INFORMATION

Article History:
Received: June 4, 2025
Last Revision: September 9, 2025
Published *Online*: September 30, 2025

KEYWORDS

System Evaluation,
SIGNAL Application,
User Satisfaction,
PIECES Framework,
Apptim

CORRESPONDENCE

Phone: 085368686691
E-mail: *rahmafitria@unimal.ac.id*

ABSTRACT

The SIGNAL application is a digital service from the Korlantas Polri that makes it easier for people to take care of STNK validation online. However, some users still experience problems such as the verification process that is not always successful, slow delivery of physical documents, less rapid customer service response, and slow verification or login process. Therefore, this study aims to evaluate the performance of the SIGNAL application with the PIECES framework approach which includes six aspects: Performance, Information, Economic, Control, Efficiency, and Service. The data collection method was carried out through distributing questionnaires to 300 respondents who use the SIGNAL application. In addition, technical performance testing was also carried out using Apptim tools to measure application technical metrics. The results showed that overall, users were satisfied with the system based on the six aspects of PIECES with an average score of 3.9 on a scale of 5.76% of respondents stated that they were satisfied to very satisfied, 15% were undecided/neutral, and 9% were dissatisfied. This finding indicates that the majority of users consider this application to be quite effective and worth using. Performance testing using Apptim resulted in an average response time of 2.4 seconds, CPU usage of 18%, memory usage of 170MB, and no errors (error rate 0%), indicating that the application is quite stable and runs well on user devices. It is hoped that this research can be the basis for further development of the SIGNAL application, especially in improving service aspects and overall system efficiency.

1. INTRODUCTION

Along with the development of technology, the government introduced the E-Samsat system, which allows tax payments to be made online [1]. E-Samsat is managed regionally by each province to provide solutions to some of the shortcomings of the manual [2]. However, the implementation still faces obstacles, such as the lack of digital literacy of the public, limited internet access, and variations in service quality between provinces. As a further innovative step, the Indonesian government through the POLRI Traffic Corps in the field of Information Technology launched the SIGNAL (National Digital Samsat) application in 2021 which aims to make it easier for people to get Samsat services digitally [3]. As a national digital service, SIGNAL is designed to make it easier for people to carry out motor vehicle administration, such as tax payments, cross-province STNK extensions, and other administrative services online easily, quickly, and effectively via mobile or digital devices, while supporting cross-regional payments as a national [4].

Many kind of research carried out mobile evaluation on SIGNAL, but few research that assesses utilizing the PIECES framework. The previous research analyzes the sentiment of SIGNAL reviews using the Naïve Bayes method focused on analyzing the sentiment of SIGNAL

app reviews to evaluate user perceptions [4], [7]. While the current study aims to evaluate the SIGNAL information system comprehensively using the PIECES framework, with a focus on system performance and user satisfaction levels [17].

Although SIGNAL is an official application supported by POLRI as the national Samsat supervisor, users still face several obstacles and problems. Based on various research journals and reviews on the Google Play Store, some users report technical problems, such as errors, bugs, and data errors that cause confusion. In addition, the verification process, which is considered too long and complicated, is a major complaint [3]. These problems have an impact on decreasing user comfort and trust in the SIGNAL application. This can also affect user satisfaction levels and hinder wider adoption of this digital service [18].

To address these issues, a thorough evaluation of the SIGNAL system using the PIECES framework is required. This approach will help identify the main constraints from the aspects of performance, information, economy, control, efficiency, and service, so as to provide strategic recommendations to improve the quality and reliability of the application according to user needs [5]. n addition, this research will also use additional analysis with Apptim to strengthen system testing, especially on performance aspects.

2. RELATED WORK

In recent years, a number of studies have focused on evaluating the SIGNAL app that serves as a digital platform for motor vehicle tax payments in Indonesia. These studies have mostly analyzed user sentiment and technical challenges, and provided important insights into areas for improvement. Rahman et al. [6] conducted a sentiment analysis of 1,000 user reviews of the SIGNAL application using the Naïve Bayes algorithm. This study recorded a classification accuracy of 89%, with a precision value of 95% and a recall of 85%. It also identified key sources of user dissatisfaction such as system errors, inaccurate data processing, and unresponsive customer service. While this analysis focuses on textual sentiment trends, it has not evaluated the system's overall operational performance. This shows the importance of user feedback in digital service development. Wijaya et al. [3] also applied the Naïve Bayes algorithm to 2,000 user reviews from the Google Play Store. After a manual labeling process, the data was divided into 1,665 positive reviews and 335 negative reviews. Although the accuracy of this model was lower at 63.61%, the high precision (92.19%) confirmed that most users expressed satisfaction, although there were still complaints about technical glitches and user experience.

This research only focused on sentiment classification without assessing more in-depth system factors such as service efficiency or information reliability. Another prominent study was conducted by Kacung et al. [7], who used the Support Vector Machine (SVM) algorithm to classify user sentiment. The linear kernel showed the highest accuracy of 96.2%, with additional validation through K-Fold Cross Validation resulting in an average accuracy of 97.65%. While these results are promising, users still experienced issues such as verification failures and OTP bugs, which affected the level of trust in the app [20]. These findings emphasize the need for more in-depth technical evaluation and user-oriented improvements.

3. METHOLOGY

The methodology outlines the specific technical processes and procedures that will be carried out during the course of the research as shown in Figure 1.

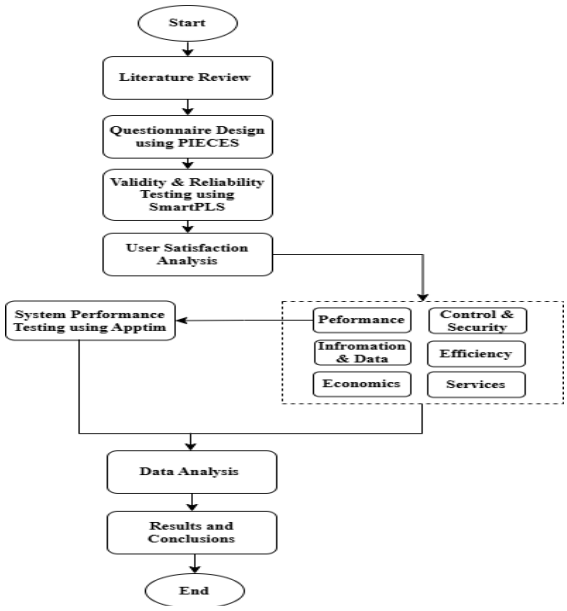

FIGURE 1. RESEARCH METHODOLOGY

3.1 Data Collection Technique

Data collection in this study was carried out through the preparation of questionnaires and their distribution to respondents who were active users of the SIGNAL application. The data collection technique in this study consists of several methods, namely: Literature study to obtain the theoretical basis, the questionnaire whose preparation stage is carried out based on the PIECES evaluation framework which includes six main aspects (Performance, information, economic, control, efficiency, service) [8].

3.2 Validity Test

The validity test in this study uses the Average Variance Extracted (AVE) method, which aims to measure the extent to which the indicators used in the questionnaire can explain the latent variables or constructs being measured [9]. These latent variables refer to constructs that cannot be measured directly, such as user satisfaction, application performance, or service quality [10]. An indicator is declared valid if the AVE value is > 0.5. Conversely, if the AVE value is <0.5 (less than 0.5), then the indicator is considered less valid and may need to be corrected or replaced with another indicator that is more representative.

3.3 Reliability Test

The reliability test is carried out using Composite Reliability (CR), which aims to measure internal consistency between indicators in each variable [11]. In

general, a variable is declared reliable if the CR value is greater than 0.7. If the CR value is <0.7 (less than 0.7), then the variable is considered less reliable and further evaluation is needed, such as replacing or revising inappropriate indicators.

3.4 User Sastifaction Level Analysis

The reliability test is carried out using Composite After conducting data quality testing, the next step is to analyze the level of satisfaction of SIGNAL application users based on the six aspects of the PIECES framework. The analysis is carried out by calculating the average level of satisfaction for each indicator in the framework. To find the average level of satisfaction can be calculated using equation (1).

$$RK = \frac{JSK}{JK} \quad (1)$$

Noted:

a. RK : Average User Satisfaction.
b. JSK : Total Questionnaire Score
c. JK : Number of Questionnaires

The results of the calculated average score are then interpreted to provide important information, including whether users are overall satisfied, which aspects still experience obstacles or limitations, as well as recommendations that can be given to improve the quality of the SIGNAL application [12]. The average value of this calculation is then classified based on the satisfaction level table using the model defined by Kaplan and Norton, with the following levels [13].

TABLE 1. AVERAGE SATISFACTION LEVEL

Answer Options	Value Range	Score
Very Dissatisfied	1.00 – 1.79	1
Dissatisfied	1.80 – 2.59	2
Undecided	2.60 – 3.39	3
Satisfied	3.40 – 4.19	4
Very Satisfied	4.20 – 5.00	5

3.5 Application Performance Testing With Apptim

To complement the analysis of user satisfaction levels, this research also conducted performance testing of the SIGNAL application using Apptim, which is a software specifically designed to test the performance of mobile applications on Android and iOS devices [14]. Testing with Apptim aims to evaluate the technical aspects of the application directly from the user's side, under real usage conditions [15]. Apptim provides various important performance metrics, including.

1) CPU Usage: Shows how much processing load the app is putting on the device during the test.
2) Memory Usage: Analyzes the memory consumption that the app uses while active.
3) Crash Reports: Detects if a crash (crash or sudden stop of the application) occurs during testing.
4) App Render Time and Network Performance: Provides information on the smoothness of the app's visual display and the quality of data communication.

The results of these tests help to provide an overview of the application's actual performance and detect potential technical issues that may not be revealed through user satisfaction surveys. Thus, the use of Apptim as a test tool adds validity to the evaluation results, as it is able to objectively assess the system from both a technical and user experience perspective [16].

3.5 Hypothesis Test

Hypothesis testing in this study was carried out using a one sample t-test, with the aim of knowing whether user perceptions of each aspect of PIECES differ significantly from the neutral value, which is 3.00. The value of 3.00 is used as a reference value because it represents a neutral or "undecided" position on a Likert scale of 1-5.

In this test, the average score of respondents' responses will be compared with the reference value, to determine whether there is a statistically significant difference. The following are the research hypotheses developed in this study:

1) First Hypothesis: Evaluation of Performance Aspects against the Reference Value (3.00). H0 : Average user perception of performance aspects = 3.00. H1 : Average user perception of the performance aspect ≠ 3.00.
2) Second Hypothesis: Evaluation of Information & Data Aspects against the Reference Value (3.00). H0 : Average user perception of information & data aspects = 3.00. H1 : Average user perception of information & data aspects ≠ 3.00.
3) Third Hypothesis: Evaluation of Economics Aspects against the Reference Value (3.00). H0 : The average user perception of the Economic aspect = 3.00. H_1 : The average user perception of the Economic aspect ≠ 3.00.
4) Fourth Hypothesis: Evaluation of Control & Security Aspects against the Reference Value (3.00). H0 : Average user perception of control & security aspects = 3.00. H_1 : The average user perception of the control & security aspect ≠ 3.00
5) Fifth Hypothesis: Evaluation of Efficiency Aspects against the Reference Value (3.00). H0 : Average user perception of efficiency = 3.00. H_1 : The average user perception of the efficiency aspect ≠ 3.00.
6) Sixth Hypothesis: Evaluation of Services Aspects against Reference Value (3.00). H0 : Average user perception of services = 3.00 H_1 : The average user perception of the services aspect ≠ 3.00.

4. RESULT AND DISCUSSION

This research combines user perception analysis and technical testing to evaluate the SIGNAL application. User feedback was collected through a PIECES-based questionnaire and analyzed with SmartPLS, showing valid and reliable results with average satisfaction scores above 3.00. For the technical evaluation Apptim was used as a tool to monitor application performance metrics during real-time interaction.

Although the majority of users are satisfied, the data shows that there are still a number of obstacles. These obstacles include slow delivery of physical documents, slow customer service response, unclear delivery status information, high memory consumption, and verification processes that are not always successful. This shows that service aspects, system efficiency, and clarity of

information still need to be improved in order to optimize the user experience [18][19].

4.1 Validity Test

The validity test is used to determine the extent to which the statement items on the questionnaire are able to measure the aspects included in the study. The test uses Average Variance Extracted (AVE) which has the criteria that a construct is declared valid if the AVE value is> 0.50. By using AVE, it can be ensured that each indicator in the instrument is able to present its variable well and is suitable for further analysis as shown in Figure 2.

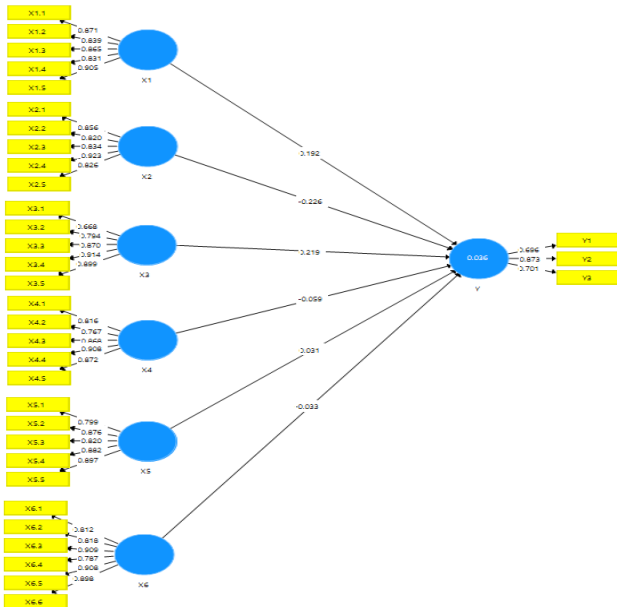

FIGURE 2. STRUCTURAL MODEL IN SMARTPLS

The results of the validity test using SmartPLS modeling show that of the 36 statement items tested, 34 statement items were declared valid with a value> 0.7. In other words, about 94.4% of the items in the instrument meet the predetermined validity criteria, so it can be ascertained that the instrument used has a very high level of reliability and relevance in measuring the intended concept.

TABLE 2. CONVERGENT VALIDITY TEST RESULTS

Variabel	Indicator	Outer Loading	AVE	Status
Performance	X1	0.871	0.744	Valid
	X2	0.839		Valid
	X3	0.865		Valid
	X4	0.831		Valid
	X5	0.905		Valid
Information & Data	X1	0.856	0.726	Valid
	X2	0.820		Valid
	X3	0.834		Valid
	X4	0.923		Valid
	X5	0.826		Valid
Economics	X2	0.794	0.695	Valid
	X3	0.870		Valid
	X4	0.914		Valid
	X5	0.899		Valid
Control & Security	X1	0.816	0.718	Valid
	X2	0.767		Valid
	X3	0.868		Valid
	X4	0.908		Valid
	X5	0.872		Valid
Efficiency	X1	0.799	0.732	Valid
	X2	0.876		Valid
	X3	0.820		Valid
	X4	0.882		Valid
	X5	0.897		Valid
Sevices	X1	0.812	0.734	Valid
	X2	0.818		Valid
	X3	0.909		Valid
	X4	0.787		Valid
	X5	0.908		Valid
	X6	0.896		Valid
Satisfaction	Y2	0.873	0.579	Valid
	Y3	0.701		Valid

Based on the results of validity testing using SmartPLS, it was found that there were 2 questionnaire statement items that were quite valid. The following table displays the results of the validity test which has a fairly valid value.

TABLE 3. VALIDITY TEST RESULTS MODERATELY

Variabel	Indicator	Outer Loading	Status
Economics	X1	0.668	Sufficiently Valid
Satifaction	Y1	0.696	Sufficiently Valid

The table above shows statement items that are quite valid, but are still retained because they have strong theoretical relevance, so they are still considered important in measuring latent constructs. In addition, although there are indicators with values below 0.7, the Average Variance Extracted (AVE) value for the construct concerned remains valid (>0.5). This shows that the construct still has a good ability to explain the variance of its indicators.

4.2 Reliability Test

The reliability test aims to assess the internal consistency of each indicator used in the research variable. In this test the test results are seen and assessed using Composite Reliability (CR). A good CR value is> 0.70, which indicates that the instrument has a high level of reliability and can be trusted in measuring the intended construct.

TABLE 4. RELIABILITY TEST RESULTS

Variabel	Composite Reliability	Status
Performance	0.9356	Realiabel
Information & Data	0.9300	Realiabel
Economics	0.9185	Realiabel
Control & Security	0.9271	Realiabel
Efficiency	0.9318	Realiabel
Services	0.9429	Realiabel
Satisfaction	0.8032	Realiabel

Based on the results of the reliability test using the Composite Reliability (CR) value, all variables in the model have a value> 0.7. This shows that all constructs are declared reliable, meaning that the indicators in each variable are able to consistently measure the intended construct.

4.3 Results Of User Satisfaction Level Analysis

1) Performance

The performance variable is used to determine the level of reliability and performance of the SIGNAL application based on user perceptions.

TABLE 5. PERFORMANCE QUESTIONNAIRE ASSESSMENT RESULTS

	Performance				
Answer	SS	S	RR	TS	STS
Score	5	4	3	2	1
Total Score	539	606	217	84	54

$$\text{RK: } \frac{(539 * 5) + (606 * 4) + (217 * 3) + (84 * 2) + (54 * 1)}{5 * 300}$$

RK: $\frac{2.695 + 2.424 + 651 + 168 + 54}{5 * 300}$

RK: 3.99

The SIGNAL application received an average score of 3.99, indicating user satisfaction in terms of performance, accessibility, and feature clarity. However, concerns remain regarding device compatibility, as some users reported issues on certain devices. Developers are advised to improve cross-device stability through broader system testing and regular technical updates.

2) Information & Data

This information & data variable is to assess the extent of information quality, data completene

TABLE 6. INFORMATION DATA QUESTIONNAIRE ASSESSMENT RESULT

	Information and Data				
Answer	**SS**	**S**	**RR**	**TS**	**STS**
Score	5	4	3	2	1
Total Score	570	592	196	83	59

RK: $\frac{(570 * 5) + (592 * 4) + (196 * 3) + (83 * 2) + (59 * 1)}{5 * 300}$

RK: $\frac{2.850 + 2.368 + 588 + 166 + 59}{5 * 300}$

RK: 4.02

The average score of 4.02 indicates user satisfaction with the clarity, accuracy, and security of information in the SIGNAL app. However, some users expressed uncertainty about the data download feature, suggesting the need for improved accessibility and clearer usage instructions. regarding data protection.

3) Economics

This Economics variable evaluates the economic value of SIGNAL application users, both in terms of cost efficiency, internet data consumption, and its usefulness in the vehicle tax payment process.

TABLE 7. ECONOMICS QUESTIONNAIRE ASSESMENT RESULTS

	Economics				
Answer	**SS**	**S**	**RR**	**TS**	**STS**
Score	5	4	3	2	1
Total Score	596	590	180	75	59

RK: $\frac{(596 * 5) + (590 * 4) + (180 * 3) + (75 * 2) + (59 * 1)}{5 * 300}$

RK: $\frac{2.980 + 2.360 + 540 + 150 + 59}{5 * 300}$

RK: 4.24

The average score of 4.24 indicates a very high level of satisfaction. SIGNAL is considered lightweight, accessible for free, and does not require additional burdensome costs, thus providing immediate economic benefits to users. However, there are still concerns regarding the efficiency of internet data consumption, indicating the need for optimization of application size and cache management to reduce data usage.

4) Control & Security

This control & security variable measures how well the application's control and security system protects user data and transaction processes.

TABLE 8. CONTROL SECURITY QUESTIONNAIRE ASSESSMENT RESULTS

	Control & Security				
Answer	**SS**	**S**	**RR**	**TS**	**STS**
Score	5	4	3	2	1
Total Score	512	590	247	94	57

RK: $\frac{(512 * 5) + (590 * 4) + (247 * 3) + (94 * 2) + (57 * 1)}{5 * 300}$

RK: $\frac{2.560 + 2.360 + 741 + 188 + 57}{5 * 300}$

RK: 3.94

The average score of 3.94 falls into the satisfied category. Most users feel that the SIGNAL app is safe to use, that their personal data is protected, and that they are not bothered by spam or advertisements. However, some respondents expressed doubts about data security and app stability, indicating that trust in data protection still needs to be strengthened. It is recommended to improve data security standards, such as implementing advanced encryption and providing regular notifications to users regarding data protection.

5) Efficiency

The Efficiency variable assesses how efficient the application is in terms of use, appearance, and ease of access to features.

TABLE 9. EFFICIENCY QUESTIONNAIRE ASSESSMENT RESULTS

	Efficiency				
Answer	**SS**	**S**	**RR**	**TS**	**STS**
Score	5	4	3	2	1
Total Score	541	617	193	94	55

RK: $\frac{(541 * 5) + (617 * 4) + (193 * 3) + (94 * 2) + (55 * 1)}{5 * 300}$

RK $= \frac{2.705 + 2.468 + 579 + 188 + 55}{5 * 300}$

RK $= 4.00$

An average score of 4.00 places this aspect in the satisfied category. Users find the application's interface helpful in facilitating transactions. A minor issue was noted in customer service response time, as some users expressed uncertainty. It is recommended to improve service responsiveness by enhancing the ticketing system and streamlining the transaction process for greater efficiency.

6) Services

The services variable aims to evaluate the quality of service provided by the SIGNAL application, including clarity of document delivery information, notifications, and ease of use of procedures.

TABLE 10. SERVICES QUESTIONNAIRE ASSESSMENT RESULTS

	Services				
Answer	**SS**	**S**	**RR**	**TS**	**STS**
Score	5	4	3	2	1
Total Score	541	617	193	94	55

RK: $\frac{(641 * 5) + (715 * 4) + (263 * 3) + (107 * 2) + (74 * 1)}{5 * 300}$

RK: $\frac{3.205 + 2.860 + 789 + 214 + 74}{5 * 300}$

RK: 4.76

This aspect received the highest average score of 4.76, indicating a very satisfied category. It reflects that the SIGNAL application provides strong service support, especially through timely notifications, simplified procedures, and responsive customer service. However, several users expressed uncertainty regarding the delivery of physical documents, mainly due to delays rather than lack of clarity. It is recommended to improve distribution efficiency and strengthen coordination with courier services, as well as provide clearer delivery time estimates within the app.

4.4 Application Performance Test Result

After testing the SIGNAL application using the Apptim tool, various performance metrics were obtained including CPU usage, memory, network, rendering, and energy consumption. This test aims to evaluate the stability, resource efficiency, and responsiveness of the application while running on Android devices. The data obtained will be the basis for analyzing the technical performance of the application under normal usage conditions. The test results can be summarized as follows.

1) Application Resource Measurement Results

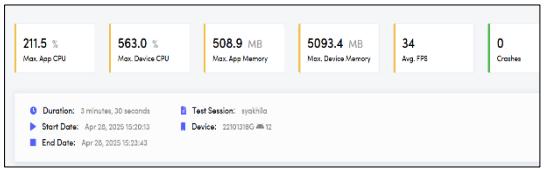

FIGURE 3. APPLICATION RESOURCE MEASUREMENT RESULTS

Performance testing showed in Figure 3 that the SIGNAL app used an average of 34.0% CPU, peaking at 211.5%, while device-wide CPU usage averaged 315.4% with a maximum of 563.0%. The app's memory usage averaged 281.4 MB and reached a maximum of 508.9 MB, exceeding the moderate threshold (>400 MB), indicating a need for optimization. Device memory usage averaged 4820.5 MB. No crashes occurred during testing, indicating stable operation. However, the app's rendering performance averaged 34 FPS below the optimal 60 FPS suggesting the interface may appear during use.

2) Recapitulation of measurement results (summary)

During the performance testing, Apptim also recorded the activities of the SIGNAL application in the form of video documentation. This video aims to provide a visual overview of the application usage process during testing, including interactions, screen transitions, and application responses to user commands. With this video, testers can evaluate application performance in more depth, both in terms of rendering and system as shown in Figure 4.

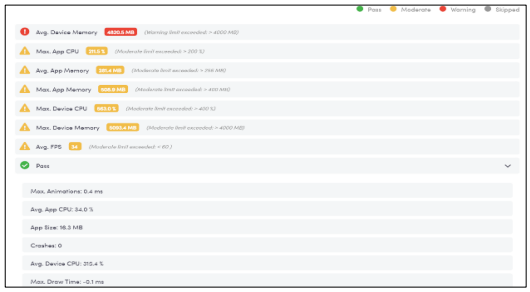

FIGURE 4. SUMMARY VIEW OF TEST RESULTS

In conclusion, although the SIGNAL application does not crash and continues to run, there are some aspects such as memory consumption and CPU usage that exceed the normal moderate limit. This shows that the application still needs optimization, especially in resource usage efficiency so that performance is more optimal and battery usage is more efficient.

3) Application Metrics Measurement Results

In the Metrics section, Apptim presents a more detailed visualization of the app's resource usage in graph form. These graphs include CPU usage, app memory, device memory, rendering performance (FPS) and app energy score over the test period. With this view, users can monitor resource usage fluctuations in real-time to identify usage patterns and potential performance issues. The following shows the results of the app's metrics measurements during the test as shown in Figure 5.

FIGURE 5. SIGNAL APPLICATION TESTING METRICS CHART

4) Testing environment configuration

In testing the performance of the SIGNAL application with Apptim, it can also identify device specifications and test environments to ensure that test results can be accurately analyzed. The following test environment information is available:

a. The Android version used on the test device.
b. The amount of RAM of the device.
c. CPU specifications include processor type, number of cores, and CPU architecture.
d. The resolution and screen density of the device.
e. The version of the application being tested as well as the application package name.
f. The host environment where the test process runs, such as the operating system and version of Apptim used. This information serves to provide context to the performance test results, ensuring that the analysis takes into account the device specifications and environmental conditions during the test.

5) Results of Monitoring Errors

In the application In addition to measuring performance, Apptim also monitors for errors or crashes that occur while the application is running. Based on the test results, no errors or crashes were found in the SIGNAL application during the test session. This shows that the application runs stably and does not experience interruptions that cause forced termination by the operating system. This information is important to show the level of stability of the application from a technical perspective during normal use. within the app.

4.5 Signal App Improvement Recommendations

Based on the evaluation results using the PIECES framework and technical performance testing with Apptim, the following are recommendations for improvement for the SIGNAL application according to each aspect as shown in Table 11.

TABLE 11. RECOMMENDATION FOR IMPROVEMENT

PIECES Aspect	Recommendation
Performance	1. Background process optimization 2. Reduced load on the home page 3. Conduct profiling to identify heavy features
Information & Data	1. Show more detailed status information 2. Add a download feature or automatic sending of digital evidence to email/WhatsApp
Economics	1. Improve the process flow so that there is no repetition 2. Reduce heavy data processing to save quota
Control & Security	1. Some respondents experienced errors when uploading documents 2. Add automatic verification (OCR) 3. Improve the error notification system, and stabilize the system during high-traffic times stabilkan sistem saat trafik tinggi
Efficiency	1. Add live chat/quick response feature 2. Strengthen input validation so that errors can be minimized
Services	1. Add shipment tracking in the app 2. Enhance automated assistance system (chatbot/live chat)

4.6 Hypothesis Test Results

The following hypothesis test presents a summary of the results of the one-sample t-test for the six aspects in the PIECES framework. This table shows the mean, standard deviation, t-value, and hypothesis testing decision based on a t-table value of 1.967.

TABLE 12. RECAPITULATION OF ONE-SAMPLE T TEST RESULTS

PIECES Aspect	Average ($\bar{x}$)	Standar Deviation (s)	t-count	t-table	Decison
Performance	3.99	0.89	19.8	1.967	H_0 is rejected (significant)
Information & Data	4.02	0.90	20.4	1.967	H_0 is rejected (significant)
Economics	4.06	0.89	21.2	1.967	H_0 is rejected (significant)
Control & Security	3.94	0.90	18.8	1.967	H_0 is rejected (significant)
Efficiency	4.00	0.90	20	1.967	H_0 is rejected (significant)
Service	3.97	0.90	19.4	1.967	H_0 is rejected (significant)

Based on the table above, all t-count values are greater than the t-table, so all null hypotheses (H_0) are rejected. This shows that users' perceptions of each aspect of PIECES are significantly different from the neutral value (3.00). Thus, it can be concluded that users are satisfied with each aspect of using the SIGNAL application.

5. CONCLUSIONS

Based on the evaluation using the PIECES framework and technical testing through Apptim, it can be concluded that most users are satisfied with the SIGNAL application, particularly in aspects such as Performance, Information & Data, and Economics. However, lower satisfaction was noted in the Services aspect, especially regarding delays in physical document delivery. Despite the generally positive response, several issues remain, including slow delivery processes, limited responsiveness from customer service, high memory usage, and verification-related problems. Technical testing showed stable performance with fast startup times, acceptable CPU and memory usage, and no crashes. To enhance user experience, improvements are recommended in post-transaction services, performance optimization, delivery status clarity, and service responsiveness. Overall, while the SIGNAL app meets most user expectations, targeted enhancements are still needed to ensure a more reliable and efficient service.

REFERENCES

[1] I. Puspitasari *et al.*, "Edukasi Pembayaran Pajak Melalui Implementasi E-Samsat," vol. 5, no. 2, pp. 219–229, 2022.

[2] W. Hertiarani, "Implementasi Kebijakan E-Samsat Di Jawa Barat," *J. Ilmu Adm. Media Pengemb. Ilmu dan Prakt. Adm.*, vol. 13, no. 3, pp. 419–440, 2020.

[3] N. Nurzaman, N. Suarna, and W. Prihartono, "Analisis Sentimen Ulasan Aplikasi Threads Di Google Playstore Menggunakan Algoritma Naïve Bayes," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 8, no. 1, pp. 967–974, 2024, doi: 10.36040/jati.v8i1.8708.

[4] Siska Rafitanuri, Nur Arsyida, and Rizky Gunawan, "Analisis Tingkat Kepuasan Masyarakat Terhadap Pelayanan Pembayaran Pajak Kendaraan Bermotor Berbasis Aplikasi Signal Di Kantor Samsat Kota Tanjungpinang," *J. Hukum, Polit. Dan Ilmu Sos.*, vol. 1, no. 3, pp. 92–103, 2022, doi: 10.55606/jhpis.v1i3.537.

[5] A. E. Karlinda, P. Azizi, and M. F. Sopali, "Pengaruh pengalaman kerja, prestasi kerja, pendidikan dan pelatihan terhadap pengembangan karir pada PT. PLN (persero) kota padang rayon kuranji," *J. Inf. Syst. Applied, Manag. Account. Res.*, vol. 5, no. 2, pp. 523–531, 2021, doi: 10.52362/jisamar.v5i2.

[6] I. F. Rahman, A. N. Hasanah, and N. Heryana, "Analisis Sentimen Ulasan Pengguna Aplikasi Samsat Digiital Nasional (Signal) Dengan Menggunakan Metode Naïve Bayes Classifier," *J. Inform. dan Tek. Elektro Terap.*, vol. 12, no. 2, pp. 963–969, 2024, doi: 10.23960/jitet.v12i2.4073.

[7] S. Kacung, C. Pamungkas Putra Bagyana, and D. Cahyono, "Analisis Sentimen Terhadap Layanan Samsat Digital Nasional (Signal) Menggunakan Metode Svm," *J. Mnemon.*, vol. 7, no. 1, pp. 118–122, 2024, doi: 10.36040/mnemonic.v7i1.9557.

[8] N. L. Kakihary, "Pieces Framework for Analysis of User Saticfaction Internet of Things-Based Devices," *J. Inf. Syst. Informatics*, vol. 3, no. 2, pp. 243–252, 2021, doi: 10.33557/journalisi.v3i2.119.

[9] N. P. Astuti and R. Bakri, "Pelatihan Pengolahan Data Menggunakan Aplikasi Smart-PLS 3 Secara Online di Masa Pandemik Covid 19," *CARADDE J. Pengabdi. Kpd. Masy.*, vol. 4, no. 1, p. 614, 2021.
[10] M. Riefky and W. N. Hamidah, "SNHRP-II," 2019.
[11] R. D. Hanifa, Mustafid, and A. R. Hakim, "Departemen Statistika, Fakultas Sains dan Matematika, Universitas Diponegoro," *J. Gaussian*, vol. 10, no. 2, pp. 279–292, 2021.
[12] G. Esmemed, G. Testiana, and F. Nopriani, "Analisis Kepuasan Pengguna Terhadap Penggunaan Aplikasi Transportasi Online Dengan Metode Unified Theory Of Acceptance And Use Of Technology," *J. Comput. Inf. Syst. Ampera*, vol. 2, no. 2, pp. 78–93, 2021, doi: 10.51519/journalcisa.v2i2.62.
[13] R. Prayogi, K. Ramanda, C. Budihartanti, and A. Rusman, "Penerapan Metode PIECES Framework Dalam Analisis dan Evaluasi Aplikasi M-BCA," *J. Infortech*, vol. 3, no. 1, pp. 7–12, 2021, doi: 10.31294/infortech.v3i1.10122.
[14] D. Gavilanes, "Apptim Review Mobile Testing," Abstracta Blog. [Online]. Available: https://abstracta.us/blog/mobile-testing/apptim-review/, 2020.
[15] S. Palamarchuk, "Mobile App Performance with apptim," DEV Community. [Online]. Available: https://dev.to/sofiapalamarchuk/how-to-test-mobile-app-performance-with-apptim-3436. 2020.
[16] Q. Madness, "How to use apptim for mobile performance testing" QA Madness Blog. [Online]. Available: https://www.qamadness.com/how-to-use-apptim-for-mobile-performance-testing/%0A%0A, 2020.
[17] S. Anjelita, R. Fitria, and A. Pratama, "Evaluasi Usability Dan Functionality Website Linkedin Menggunakan ISO/IEC 25010," *Jurnal Sistem Informasi Musi Rawas*, vol. 10, no. 1, pp. 96–111, Jun. 2025.
[18] A. Hussain and R. Fitria, "Mobile Flight and Hotel Booking Application: A Heuristic and UX Test", *JTEC*, vol. 10, no. 1-11, pp. 93–101, Feb. 2018.
[19] R. Fitria, R. Meiyanti, H. A. K. Aidilof, A. Ruzanna, W. Hamsi, and I. Na'syakban, "Developing Prototype Model Based on Analysis of The People at The Center of Mobile App Development (PACMAD) on The Panel Harga Pangan Application", *jinav j. inf. vis.*, vol. 5, no. 2, pp. 194-205, Dec. 2024.
[20] R. Diani, R. Fitria, and N. Sitti Nurfebruary, "Evaluasi Kualitas Penggunaan Aplikasi traveloka menggunakan model FURPS," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 9, no. 5, pp. 7659–7666, Jul. 2025. doi:10.36040/jati.v9i5.14812.

AUTHORS

Rahma Fitria
Rahma Fitria, B.Tech(Hons), M.Sc., was born in Banda Aceh, 12 June 1987. Currently, the author lives in Lhokseumawe City, Aceh Province. The author's current activity is teaching at the undergraduate level of the Faculty of Engineering, Information Systems Programmer, University of Malikussaleh by teaching Software Engineering courses, Software Testing and Quality and Audit of Information Systems. The author's research interests are in human-computer interaction, usability testing, software testing & quality, and other information systems-related fields.

Amanda Syakhila
She is currently an undergraduate student in the Information Systems program at Malikussaleh University. Her research interests include digital public services, system evaluation, and user experience analysis. She is currently conducting a study focused on evaluating digital government applications from both user and technical perspectives.

Desvina Yulisda
Desvina Yulisdia, S.ST., M.S., is a lecturer at the Information Systems Program, Faculty of Engineering, Universitas Malikussaleh. She was born in Bener Meriah on December 12, 1989, and currently resides in Lhokseumawe, Aceh. She obtained her master's degree in Computer Science and Engineering from National Chiao Tung University, Taiwan. Her research interests include user interface design, image-based classification, and digital media in information systems.

Azham Hussain
The Professor of Software Engineering at UUM School of Computing. He is the founder and head of Human-Centered Computing Research Group which is affiliated with the Software Technology Research Platform Center at School of Computing, Universiti Utara Malaysia. He is a member of the US-based Institute of Electrical and Electronic Engineers (IEEE), and actively involved in both IEEE Communications and IEEE Computer societies. He is published in the areas of software testing, user behaviours, group collaboration, and ubiquitous and mobile technology design. He has authored and co-authored more than 200 refereed technical publications, served as reviewer and referee for refereed journals and conferences on computing as well as the examiner. In 2012, PhD, Computer Science, University of Salford, United Kingdom 2000, Masters, Information Technology, Universiti Utara Malaysia, Universiti Utara Malaysia 1999, Bachelor Degree, Computer Science, Universiti Teknologi Malaysia, Universiti Teknologi Malaysia

Arafat Febriandirza

He is an Indonesian academic and researcher in informatics engineering, currently affiliated with the Faculty of Engineering, Universitas Muhammadiyah Prof. Dr. Hamka (UHAMKA). He holds a Ph.D. in Communication and Transportation Engineering from Wuhan University. He is a lecturer in the Department of Informatics Engineering, Faculty of Computer Science/Engineering at UHAMKA. His research interests include intelligent transportation systems, application design (such as Kotlin-based applications), sentiment analysis, and the application of machine learning methods (e.g., Naïve Bayes, Support Vector Machine). BRIN Collaboration: He has been associated with research initiatives involving the National Research and Innovation Agency (BRIN), specifically in informatics and data science.