Online-Journals.org

iJOE International Journal of Online and Biomedical Engineering

iJOE | eISSN: 2626-8493 | Vol. 20 No. 1 (2024) | OPEN ACCESS

https://doi.org/10.3991/ijoe.v20i01.45029

PAPER

# Evaluation of a Chatbot Powered by ChatGPT for the Preliminary Diagnosis of Dengue

**Raúl Jáuregui-Velarde[1], Domingo Hernández Celis[2], Pedro Molina-Velarde[3], Cesar Yactayo-Arias[4], Laberiano Andrade-Arenas[5](✉)**

[1]Universidad Privada Norbert Wiener, Lima, Perú

[2]Universidad Nacional Federico Villarreal, Lima, Perú

[3]Universidad Tecnológica del Perú, Lima, Perú

[4]Universidad Continental, Lima, Perú

[5]Universidad de Ciencias y Humanidades, Lima, Perú

landrade@uch.edu.pe

**ABSTRACT**

Dengue poses a public-health challenge in several regions. Early detection is essential to reduce the impact of a condition, and artificial intelligence (AI) and chatbots provide new opportunities to enhance diagnosis. This study evaluates the effectiveness of integrating a chatbot with ChatGPT version GPT-3.5 for the preliminary diagnosis of dengue and its contribution to timely detection. To evaluate this, two types of tests were conducted using a dataset of 30 dengue cases. In the first test, the chatbot was evaluated without being trained with dengue symptoms. In the second test, however, the model was trained using dengue symptoms obtained from official websites such as the World Health Organization (WHO) and the Pan American Health Organization (PAHO). The performance of the chatbot was evaluated using the confusion matrix, performance metrics, and user satisfaction. The results of the second test showed impressive performance, with accuracy, sensitivity, and specificity of 100%. This surpassed the first test, which achieved accuracy, sensitivity, and specificity of 83%, 80%, and 90%, respectively. In addition, 15 users reported positive satisfaction, with an overall average rating of 4.25 out of 5. In conclusion, these results highlight the effectiveness of the chatbot as a valuable public health tool for the early detection and management of dengue. It is important to note that, despite the remarkable diagnostic results of the chatbot integrated with ChatGPT, it does not replace medical judgment.

**KEYWORDS**

artificial intelligence (AI), chatbot, chat generative pre-trained transformer (ChatGPT), dengue, dengue diagnosis

## 1 INTRODUCTION

Dengue, a disease transmitted by infected Aedes aegypti mosquitoes, is a significant public-health challenge in numerous tropical and subtropical regions worldwide [1], [2]. The World Health Organization (WHO) classifies dengue into three levels: probable dengue, dengue with warning signs, and severe dengue. If not treated in a timely manner, severe dengue can lead to serious consequences such as plasma loss, excessive

Jáuregui-Velarde, R., Celis, D.H., Molina-Velarde, P., Yactayo-Arias, C., Andrade-Arenas, L. (2024). Evaluation of a Chatbot Powered by ChatGPT for the Preliminary Diagnosis of Dengue. *International Journal of Online and Biomedical Engineering (iJOE)*, 20(1), pp. 58–73. https://doi.org/10.3991/ijoe.v20i01.45029

Article submitted 2023-09-16. Revision uploaded 2023-10-27. Final acceptance 2023-10-27.

bleeding, or organ damage, resulting in high rates of morbidity and mortality [3]. In addition, this disease poses a significant threat, with approximately 390 million infections per year [4]. It is also responsible for around 25,000 deaths annually, putting more than 40% of the global population at risk. In recent decades, the virus has spread to more areas and increased in frequency. This can be attributed to the rise in international travel and trade, increasing global temperatures [5], uncontrolled population growth, and unmanaged urbanization [6]. Furthermore, considering the variable pattern of the disease that presents itself year after year, it is crucial to have tools that aid in forecasting [7] and allow for early and precise diagnosis of dengue, which becomes a priority.

Dengue diagnosis has traditionally relied on laboratory techniques. These tests require specialized infrastructure and trained personnel, and they often involve long waiting periods for results. This limits their accessibility, particularly in rural and low-income areas where the incidence of dengue is high. Faced with this problem, we propose the development of a chatbot for the preliminary diagnosis of dengue, integrated with a chat generative pre-trained transformer (ChatGPT), an artificial intelligence (AI)-based language model [8]. The proposed solution aims to leverage chatbot technology and the natural language processing capabilities of ChatGPT to offer a user-friendly, efficient, and precise alternative for early disease detection.

This study is based on the urgent need to improve diagnostic methods for dengue, particularly in contexts with limited resources. The proposal of integrating a chatbot with ChatGPT shows promise as a solution to expand the scope of early diagnosis, reduce waiting times, and provide a valuable support tool for both health professionals and users seeking self-diagnosis. Chatbots in medicine mimic human communication and, thanks to AI, are proving useful in healthcare by handling complex dialogues and being flexible [9]. This will streamline clinical decision-making and improve case management, thereby reducing the burden on healthcare systems.

The aim of this study is to design, develop, and evaluate the effectiveness of a chatbot integrated with ChatGPT version GPT-3.5 and implemented within a web application. Furthermore, this technology aims to demonstrate its feasibility and usefulness in the field of public health. Its purpose is to make a significant contribution to the early detection of dengue, facilitate proper case management, and reduce the burden of this disease in affected communities. With this innovative tool, we hope to strengthen dengue prevention and treatment efforts and provide an effective solution to the challenges associated with its spread. The structure of the paper is as follows: Section 2 presents the literature review; Section 3 presents the methodology used; Section 4 presents the case study; Section 5 presents the results; Section 6 presents the discussion; and Section 7 presents the conclusion.

## 2 LITERATURE REVIEW

This section aims to enhance the study by gathering pertinent information about dengue, the functionality of ChatGPT, and previous research related to the topic. This literature review will establish a strong basis for the design and development of the application. It also allows for deepening knowledge about the symptoms, diagnosis, and treatment of dengue, as well as exploring the possibilities and limitations of using ChatGPT in the medical field.

### 2.1 Epidemiological situation of dengue and its diagnostic challenges

Dengue is a viral disease transmitted by mosquito bites and characterized by a high prevalence and rapid spread among humans. Its geographic range is constantly

expanding, and this phenomenon is observed in numerous tropical regions worldwide [10]. Dengue usually presents as a febrile illness that tends to resolve spontaneously. However, in some cases, more severe forms of dengue can occur, with symptoms of hemorrhage and shock, especially in children. These variants can be fatal if not properly managed, and no effective antiviral treatment has been found to combat them [11]. According to Liao [12], individuals affected by dengue fever often present with a fever and rash. However, they may also occasionally experience general symptoms such as headaches, muscle aches, joint pain, and digestive problems that are not specific to dengue fever. Because the symptoms and manifestations of dengue range from a febrile illness to the appearance of a hemorrhagic syndrome [13], some may be similar to those of other diseases.

In a nutshell, dengue is a viral disease that spreads rapidly through the bite of Aedes aegypti mosquitoes. It can vary in severity, ranging from a mild fever to severe, life-threatening forms. The lack of effective antiviral treatment and the similarity of its symptoms to those of other diseases are important challenges in diagnosing and managing it.

## 2.2 Implications and potential of ChatGPT in clinical practice

ChatGPT, developed by OpenAI, is a chatbot driven by a conversational AI interface. Considered one of the most sophisticated implementations of AI, it has attracted worldwide attention [14] due to its remarkable human-like expression and reasoning ability [15]. According to Cascella et al. [16], ChatGPT is a large language model (LLM) trained on a vast text dataset with the aim of enabling dialogue with users. In the clinical field, as indicated in [17], this technology has gained increasing interest because of its wide range of applications. In addition, its ability to intelligently answer questions and provide reliable information about diseases and medical consultations is remarkable. It is also successful in generating clinical letters and other documents, thereby improving the efficiency and accuracy of health professionals. Furthermore, the study conducted by [18] has shown that AI chatbots, such as ChatGPT, have the capability to generate a distinct diagnostic list for the most prevalent complaints. However, in [19], they emphasize the importance of understanding that ChatGPT should not replace healthcare professionals. Rather, it should be seen as a complementary tool that supports and enhances the work of medical experts, enabling them to provide more effective and higher-quality care to their patients.

ChatGPT is characterized by its remarkable capacity for human-like communication and reasoning. Its application in clinical settings has generated growing interest due to its usefulness in providing reliable information and enhancing the efficiency of medical professionals. However, it is important to remember that, while it is useful for diagnosis, it should not replace medical professionals. Instead, it should complement their work to provide more effective and higher-quality care to patients.

## 2.3 Comparison of previous studies on the use of ChatGPT in medical diagnosis

Through a literature review, several studies evaluating to the use of AI conversational language models, such as ChatGPT, in the diagnosis of various diseases have been conducted. Some of these studies are discussed below, while others are compared in a table that summarizes key findings.

In an evaluation study [20], Balas and Ing compared the use of ChatGPT, an AI conversational language model, with the existing tool Isabel Pro for diagnosing

ophthalmic diseases. Using case report descriptions, they evaluated the ability of both systems to formulate tentative and differential diagnoses. ChatGPT was able to identify the correct diagnosis in 9 out of 10 cases and provide accurate lists of differential diagnoses in all cases. Isabel Pro, on the other hand, correctly identified the correct provisional diagnosis in only 1 out of 10 cases but included the diagnosis in 7 out of 10 differential diagnosis lists. The researchers concluded that conversational AI models, such as ChatGPT, have promising value in diagnosing eye diseases, especially in primary care settings. However, more research is needed to determine the optimal way to incorporate them into medical practice.

Furthermore, researchers [17] propose a novel approach to improving medical diagnosis by combining ChatGPT with large language models (LLM) and vision language models (VLM). Their approach is based on medical image classification using CLIP and ChatGPT. The findings indicate that language models can play a critical role in generating additional clues and insights to enhance the quality of image-based diagnosis. This highlights the effectiveness of integrating language models and visual models in medical applications. On the other hand, the study [21] focuses on developing a decision support system that utilizes ChatGPT to automatically generate diagnoses based on medical images. Their results demonstrate improvements in medical decision-making, cost reduction, and the empowerment of healthcare professionals. Furthermore, they are supported by experiments that demonstrate promising performance in automated diagnosis.

Taken together, these studies highlight the importance of ChatGPT in the field of medical diagnosis. Study [20] demonstrates the value of primary ophthalmic care, while study [17] proposes an innovative approach to improve diagnostic accuracy by combining speech and visual models. The study [21] highlights the potential of ChatGPT in the development of decision support systems that can benefit both physicians and patients. These studies demonstrate how ChatGPT can serve as a valuable tool for enhancing the quality and efficiency of medical diagnosis across different contexts.

Similarly, other researchers have investigated and presented solutions using ChatGPT in various medical scenarios. Below is a table that concisely compares the key findings and approaches of these studies (see Table 1).

**Table 1.** Comparison of previous studies on the use of ChatGPT in medical diagnosis

| Source | Approach | Findings | Limitations and Considerations |
|---|---|---|---|
| [22] | Evaluation of the diagnostic accuracy of ChatGPT with 50 clinical cases, including 10 rare presentations. | • ChatGPT 4 shows accuracy in common cases.<br>• Eight or more suggestions are required to diagnose rare diseases and resolve 90% of cases.<br>• In precision, ChatGPT 4 has greater precision than ChatGPT 3.5.<br>• Usefulness of ChatGPT in difficult medical cases. | • Caution should be exercised when used by non-professionals. |
| [23] | To evaluate the reliability of ChatGPT in the diagnosis and treatment of patients. They analyzed ten case reports published in Korea. | • ChatGPT correctly answered four cases using the patient's symptoms, findings, and medical history.<br>• Accuracy increases to 7 out of 10 with additional data from laboratory, pathological, and radiological results. | • Users must evaluate the validity of the answers.<br>• Inappropriate content present in some responses. |
| [24] | Evaluation of ChatGPT to disseminate knowledge about gastric cancer, provide recommendations and interpret endoscopy reports. | • ChatGPT GPT-4 achieves high precision (91.3%) and consistency (95.7%) in gastric cancer tests.<br>• Effective in the analysis of endoscopy reports. | • Promising in generating consultation recommendations and analyzing endoscopy reports. |

To summarize, the comparison table highlights the key aspects of the studies, including the approach, findings, limitations, and considerations of ChatGPT in the field of medical diagnosis. Despite its remarkable performance in diagnosing common cases and generating recommendations, it faces challenges when it comes to rare cases, and the quality of responses may vary. The need for caution and validation in the interpretation of responses is an important consideration in their use in medical practice.

Based on these findings, the present study aims to evaluate the effectiveness of a chatbot that is specifically integrated with the ChatGPT model for diagnosing dengue. The implementation of this chatbot will take place in a web application, which represents a significant advancement in the utilization of natural language technology in medical practice. It is essential to continue the research and development of this type of tool, ensuring a strong scientific and ethical foundation for its use. This will effectively and safely contribute to the enhancement of medical care and the well-being of patients.

# 3 METHOD

## 3.1 Chatbot development method

This study followed a process consisting of five phases, which are detailed below in relation to the development of the chatbot integrated with ChatGPT (see Figure 1).

**User interface design and development.** In the first phase, the design and development of the user interface are addressed. This involves starting with the design of the chatbot's appearance in the web application. This involves defining the key elements of the interface, such as the chat window and the text input bar, as well as any other necessary components. For this purpose, technologies such as Flask are used. Flask is a framework that utilizes the Python language and is known for its easy-to-understand code [25]. HTML and JavaScript are also essential in web application systems, as they provide dynamic interactions on web pages [26]. Additionally, Bootstrap is used for presentation purposes. Bootstrap is a framework that ensures websites are designed to be responsive, allowing for optimal viewing on various screen resolutions, including mobile devices [27].

**Integration of the ChatGPT model into the application.** The ChatGPT model is connected to the previously developed user interface. This step entails leveraging the APIs provided by the OpenAI platform to enable bidirectional communication between the chatbot and the application, streamlining the transmission and reception of messages. For the integration, a programming language is used. In this case, Python is the chosen language, which has become the most popular computer language [28], [29].

**Implementation of diagnostic functionality.** The necessary logic is implemented, which includes the following points: a) Specify the context or topic, clearly defining the scope in which the chatbot must provide answers. b) Modify the prompt by adding relevant information about the context or topic in the message sent to the ChatGPT model. This includes providing an introduction or a specific question related to dengue diagnosis to guide the model. c) Add constraints and guidelines, imposing additional restrictions to ensure that the chatbot solely focuses on the topic of dengue and refrains from addressing other subjects.

**Test and evaluation.** It involves an exhaustive testing process to ensure the proper functioning of the application. Likewise, the accuracy and quality of the responses generated by the chatbot are verified. In addition, diagnostic tests are performed using various cases to validate the overall effectiveness of the application.

**Deployment.** The deployment phase completes the process once development and testing have been successfully completed. In this phase, the application is deployed on a web server, allowing users to access and use it in real time.

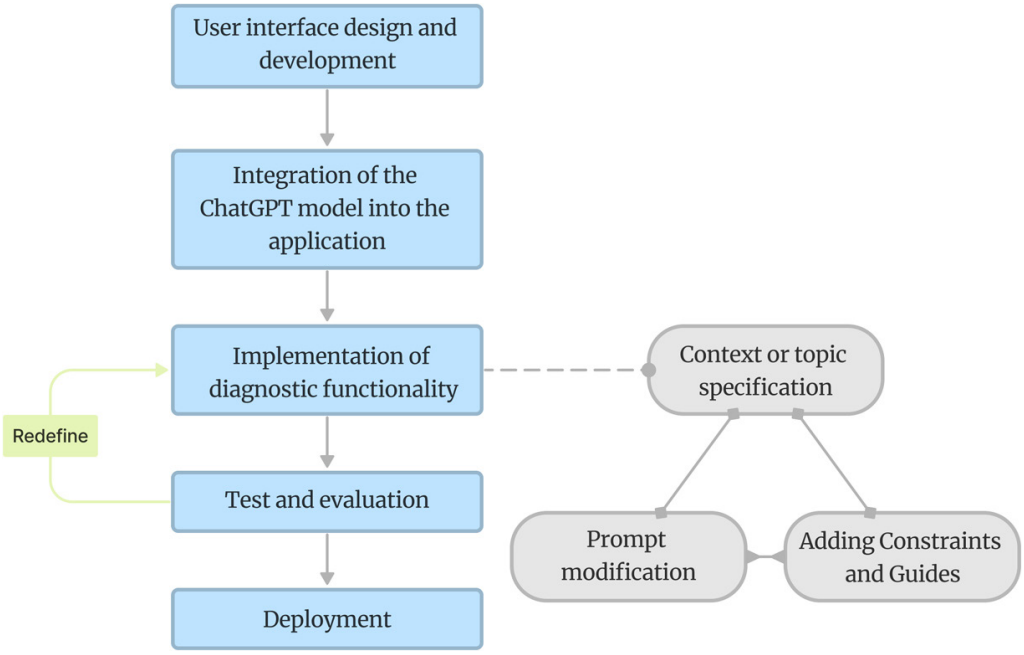


**Fig. 1.** Development phases of the proposed chatbot

## 3.2 Chatbot feasibility evaluation method

To ensure the feasibility of the dengue diagnosis chatbot, we conducted an exhaustive analysis using the confusion matrix. A matrix that presents an exhaustive analysis of the results divides them into four categories [30]: true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN) (see Table 2). Subsequently, several metrics, including accuracy, sensitivity, and specificity, were evaluated to obtain a comprehensive assessment of the reliability of the system. These results were instrumental in validating the efficacy and accuracy of the chatbot in accurately identifying dengue cases, thus making significant progress in the field of AI-assisted medical diagnosis.

**Table 2.** Confusion matrix

| Actual | Predicted | |
|---|---|---|
| | Positive | Negative |
| Positive | TP | FP |
| Negative | FN | TN |

$$Accuracy = \frac{(TP + TN)}{(TP + FP + TN + FN)} \quad (1)$$

$$Sensitivity = \frac{TP}{(TP + FN)} \quad (2)$$

$$Specificity = \frac{TN}{(TN + FP)} \quad (3)$$

### 3.3 User satisfaction evaluation method

To assess user satisfaction with the chatbot, a survey was conducted using an online questionnaire. The questionnaire consisted of ten questions based on a Likert scale, which measured participants' satisfaction levels with the chatbot's various functionalities and its overall feasibility. Subsequently, the mean and standard deviation (SD) were calculated to determine the level of satisfaction. The evaluation of satisfaction allows us to identify areas for improvement and assess the effectiveness of the chatbot as a complementary tool in medical diagnosis. This assessment significantly contributes to its potential application in future public-health contexts.

## 4 CASE STUDY

This section of the document provides a detailed explanation of the fundamental phases involved in creating the chatbot. It aims to provide a deeper and more precise understanding of how this innovative tool for the preliminary diagnosis of dengue disease was conceived and developed.

### 4.1 User interface design and development

The chatbot in question has been carefully designed and developed with a primary focus on user experience. It was conceived as a simple, elegant, and adaptable user interface tool for mobile devices, as shown in Figure 2. In addition, this chatbot has remarkable capabilities, specifically designed to integrate the ChatGPT model through its API. This integration gives the application an exceptional ability to provide contextual and intelligent responses when users enter specific symptoms in their interactions. The purpose is to diagnose possible cases of dengue. However, it is important to emphasize that the chatbot is not limited to this diagnostic function. It also allows users to ask questions related to dengue.

Likewise, the design of the chatbot includes various navigation menus that emphasize that the chatbot is not intended to replace the expertise of a medical professional. These menus guide users on how to interact with the tool and emphasize the importance of seeking a physician's opinion if the chatbot suggests a possible case of dengue. This approach reinforces the responsibility and critical role of the physician in medical decision-making, ensuring a thorough and accurate assessment of the patient's clinical situation.

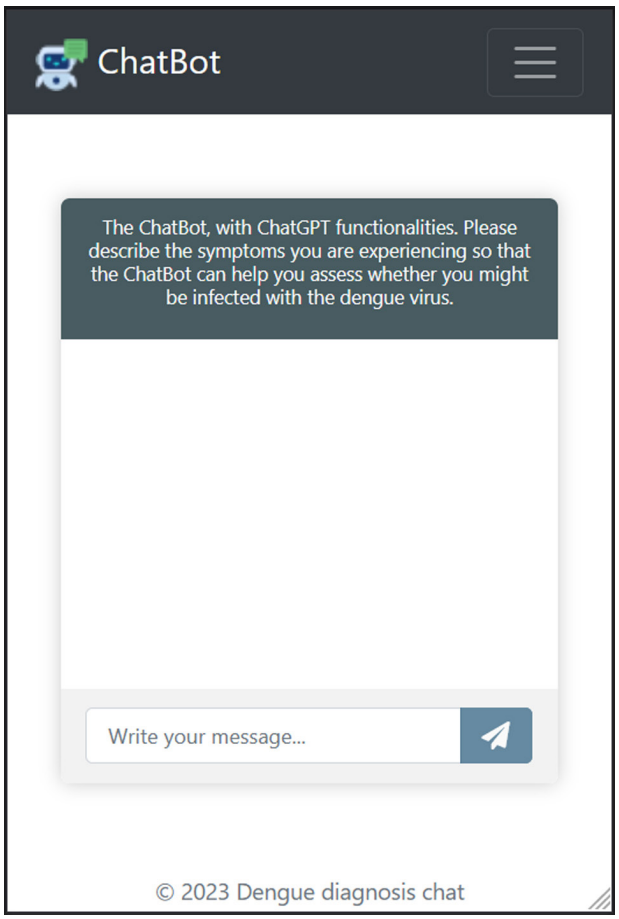


**Fig. 2.** Chatbot user interface

## 4.2 ChatGPT API configuration

The configuration involves integrating the ChatGPT API into the dengue diagnostic system. Table 3 outlines the necessary steps and procedures for facilitating interaction between the chatbot and ChatGPT. This interaction enables the system to generate diagnostic responses and recommendations using the information provided by users.

**Table 3.** ChatGPT API configuration steps

| Steps | Description | Action |
|---|---|---|
| **1.** Get access to the ChatGPT API | Obtain access to the ChatGPT API for its use in the dengue diagnosis project. | Obtain authentication credentials and enable the API. |
| **2.** Library Installation | Install the OpenAI library in the development environment. | Use 'pip' to install the library. |
| **3.** Obtaining the API Key | Obtain the ChatGPT API key from the OpenAI platform, which is necessary for authenticating API calls. | Generate and securely store the API key, |
| **4.** Interaction Code | Write the Python code that will enable the chatbot to interact with the ChatGPT API. | Import the OpenAI library and configure the API key.<br>Select and define the ChatGPT model (engine).<br>Define the maximum number of tokens. |

### 4.3 Implementation of the diagnostic context and chatbot restrictions

Once the process of configuring the ChatGPT model to send requests and receive responses using the API provided by Open AI has been completed, it is followed up with integration with the application to present these responses to the user. This step is performed to add the responses to the chat history, display them in the user interface, or take any other necessary action. However, before performing this integration, certain additional features and restrictions must be implemented and configured.

**Conversational context specification.** In Figure 3, the context or topic on which the chatbot will provide answers is explicitly presented, thus establishing a defined approach to addressing dengue. This information is stored in the variable "DENGUE_PROMPT," which facilitates effective communication with the model. The code snippet also displays the "prompt" variable, which stores a specific text used to initiate an interaction or conversation with the ChatGPT language model regarding the topic of dengue. It is important to note that this configuration prepares the model to answer questions or provide information about this topic. The variable DENGUE_PROMPT contains a message inviting the user to ask questions or provide details about dengue. Then, when the user sends a message using the "message" variable, that message is combined with the text in DENGUE_PROMPT using the format "User," resulting in a simulated conversation. This interaction simulation is used as input for the ChatGPT model, which then generates a response based on the context provided by both the prompt and the message sent by the user.

```
# Specific prompt for issues related to Dengue
DENGUE_PROMPT = "Dengue is a disease transmitted by the
Aedes mosquito. Please ask me a question or provide details
about your situation."

    prompt= DENGUE_PROMPT + "\n\nUser: " + message,
```

**Fig. 3.** Specify the conversational context and start of interaction with the model

**Generated response check.** Figure 4 shows the response message stored in the variable "NO_DENGUE_RESPONSE," which is used when the model generates a response that is unrelated to dengue. It also illustrates the verification of whether the response generated by the model is related to the problem of dengue. This verification is done by using the "if" condition, which determines whether there are any references to dengue in the generated response. If no mention of dengue is found in a specific phrase, is used to inform the user that the chatbot is not qualified to answer questions unrelated to dengue.

```
# Reply message for topics not related to Dengue
NO_DENGUE_RESPONSE = "I'm sorry, I'm not qualified to
answer that question. However, I can provide information
about dengue and its symptoms."

# Check if the answer addresses dengue related issues
if "dengue" not in reply.lower():
    reply = NO_DENGUE_RESPONSE
```

**Fig. 4.** Defined response and verification of the generated response

# 5 RESULT

## 5.1 About user interaction

Figure 5a provides an illustrative visualization of the interaction between the chatbot and the user. In this process, when the user interacts with the chatbot and provides information describing the symptoms they are experiencing, the system promptly responds by providing a possible diagnosis related to dengue, accompanied by recommendations and actions that the user could take. This chatbot's ability to analyze and provide medical diagnoses based on the user's symptoms highlights its potential as a valuable tool in healthcare and virtual assistance.

Figure 5b broadens the perspective by illustrating how the chatbot manages queries that are beyond the scope of dengue. When the user asks questions that are not related to the disease in question, the chatbot responds in an ethical and transparent manner. It indicates that it is not authorized to respond to issues unrelated to dengue. This functionality of the chatbot is essential to prevent the dissemination of incorrect or inappropriate medical information beyond its area of expertise. This ensures the integrity and accuracy of the answers it provides. Taken together, these visual representations effectively illustrate the chatbot's ability to address specific medical questions and its focus on providing reliable and relevant information about dengue.

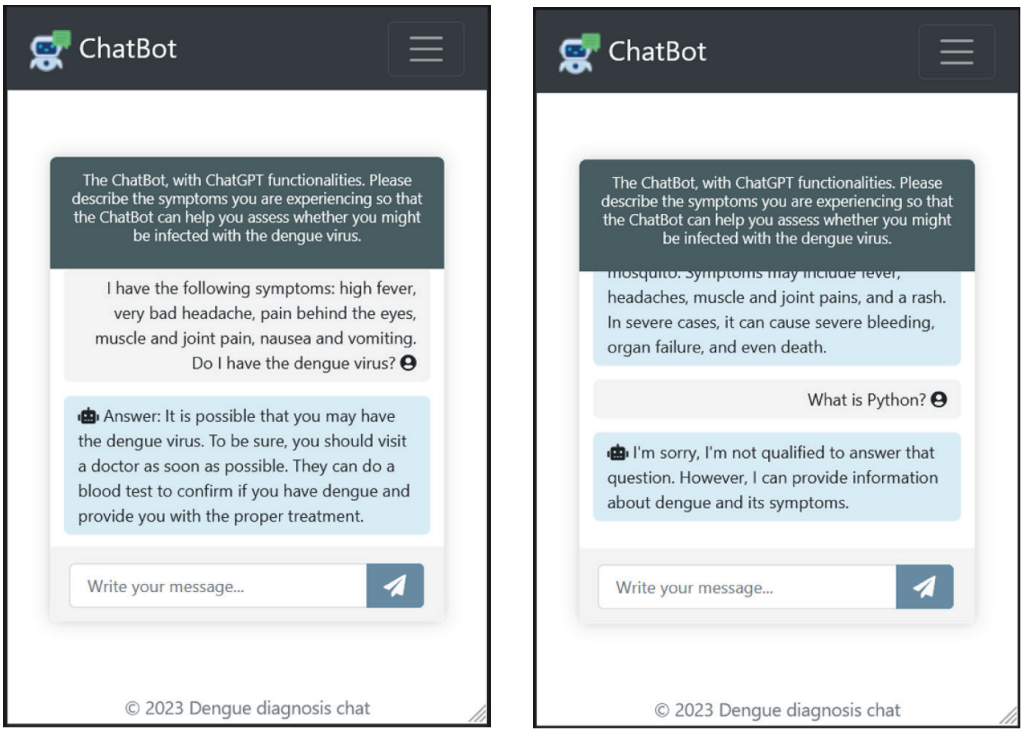


**Fig. 5.** Chatbot: (a) Dengue diagnosis and (b) Response to queries on issues not related to dengue

## 5.2 About the evaluation of the effectiveness of the Chatbot

The evaluation of the chatbot's effectiveness was conducted through two tests, both utilizing the same set of 30 dengue cases (20 positive and 10 negative). The first

test involved evaluating the chatbot without training the model with dengue symptoms, while the second test was conducted with the model that had been trained beforehand using symptoms of the disease. The Chatbot with the GPT-3.5 model was trained using prompts that included dengue symptoms provided by recognized health organizations, such as the WHO and PAHO. These symptoms were obtained from the official websites of these organizations.

In addition, the following criteria were considered when classifying the chatbot responses in the confusion matrix: a) true positives, when the chatbot correctly diagnosed dengue as a possible disease; b) false positives, when the chatbot diagnosed dengue along with other similar diseases as possible disease cases; c) true negatives, when the chatbot correctly identified that the disease was not dengue; d) false negatives, when the chatbot misdiagnosed that it was not dengue, although dengue was a possible disease.

**Comparison of confusion matrix between the results of the first and second diagnostic test.** The results obtained from the evaluation of the chatbot's effectiveness in diagnosing dengue show significant differences between the first and second tests.

Figure 6a shows the analysis of the confusion matrix for the first test. As can be seen, out of a total of 20 positive cases, the application successfully identified 16 of them as possible cases of dengue, but there were 4 cases of false negatives. Similarly, in the context of the 10 negative cases, 9 of them were correctly identified as true negatives. However, there was one instance of a false positive, where one of the negative cases was identified as a possible case of dengue.

On the other hand, Figure 6b shows the analysis of the confusion matrix corresponding to the second test, in which the chatbot was trained using the symptoms of the disease. As can be seen, the chatbot was able to correctly identify both positive and negative cases among the 30 cases analyzed in the initial test. Of the 20 positive cases, it correctly classified all 20 as possible cases of dengue. Similarly, in relation to the 10 negative cases, he also made an accurate diagnosis by classifying all 10 as true negatives.

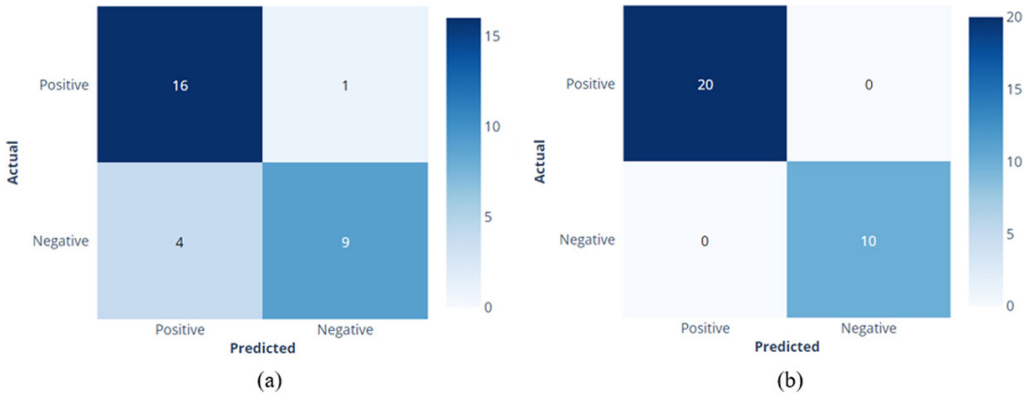


**Fig. 6.** Confusion matrix results (a) of the first test and (b) of the second test

**Comparison of chatbot performance metrics between the results of the first and second diagnostic tests.** As shown in Figure 7, there are noticeable and significant differences in the results between the first and second tests of the chatbot. In the first test, the chatbot demonstrated promising performance with an accuracy rate of 83%; however, there was still room for improvement. The sensitivity of 80% indicates that it was able to correctly identify the majority of positive dengue

cases, although there were still false negatives, which pose potential risks to patients. On the other hand, a specificity of 90% indicates that the chatbot correctly classified most of the negative cases. However, there were some false positives among the negative cases, which could lead to unnecessary concerns. However, in the second test, after training the model with dengue symptoms, the chatbot achieved a perfect level of performance. 100% accuracy, sensitivity, and specificity demonstrate an exceptionally precise performance. This means that the chatbot made no errors in classifying dengue cases as positive or negative. 100% accuracy indicates that all diagnoses made by the chatbot were correct, and both 100% sensitivity and specificity indicate that there were no incorrect positive or negative assessments. These significant improvements in the second test demonstrate the importance of model training and underscore its potential as a powerful and precise tool for AI-assisted diagnosis of dengue. Although these results are very encouraging, it is essential to further evaluate and validate the model with larger and more diverse samples to ensure its robustness and applicability in various clinical settings.

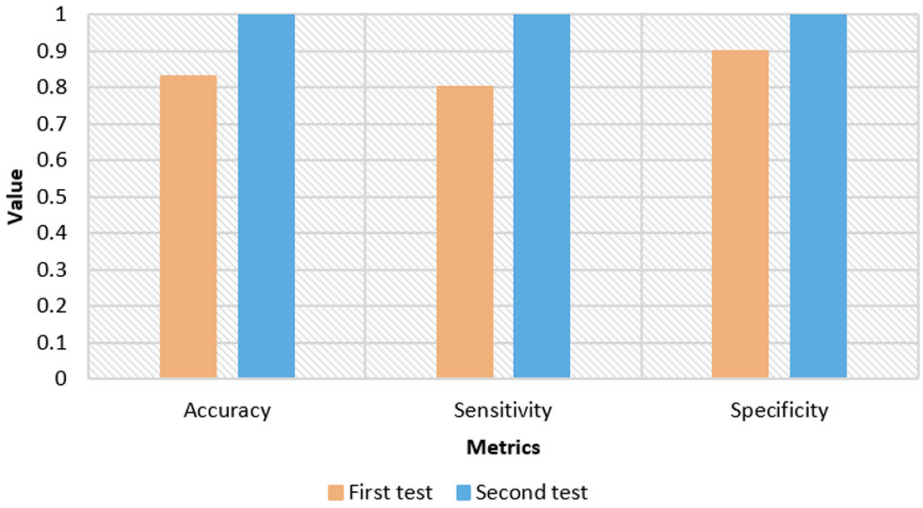


**Fig. 7.** Comparison of chatbot performance metrics: first test vs. second test

## 5.3 About user satisfaction evaluation

A total of 15 users provided a rating to express their level of satisfaction with the effectiveness of the chatbot. Table 4 displays the analysis of user satisfaction regarding the use of the chatbot integrated with ChatGPT for dengue detection. It demonstrates an overall satisfaction average of 4.25 with a SD of 0.623. These results indicate that, overall, users have a positive perception of the chatbot's effectiveness and usefulness in addressing issues related to dengue. The average score of over 4 indicates that the majority of users were highly satisfied with the chatbot. They found it effective in providing information, identifying symptoms, giving recommendations, and offering adequate knowledge about dengue. However, the relatively high SD of 0.623 indicates that there is some variability in users' responses. This variability could be attributed to differences in individual expectations or usage experiences. These results support the feasibility of using this chatbot as an effective tool to address preliminary diagnoses and concerns related to dengue. However, it is recommended to pay greater attention to the clarity of the information provided and the accuracy of responses in order to further improve user satisfaction.

**Table 4.** User evaluation results

| No | Question | Average | SD |
|---|---|---|---|
| 1 | The chatbot interface was intuitive and easy to use. | 5.00 | .000 |
| 2 | The chatbot provided clear and understandable information about dengue. | 3.60 | .507 |
| 3 | The accuracy of the chatbot's responses was satisfactory. | 3.87 | .352 |
| 4 | The chatbot showed adequate understanding of my questions. | 4.00 | .000 |
| 5 | The chatbot answered all my dengue related questions. | 4.67 | .488 |
| 6 | The chatbot was able to accurately detect my dengue symptoms. | 3.73 | .458 |
| 7 | The chatbot's response speed was adequate for my needs. | 5.00 | .000 |
| 8 | The chatbot demonstrated solid knowledge about dengue. | 3.80 | .414 |
| 9 | The chatbot provided useful recommendations for the care and treatment of dengue. | 4.60 | .507 |
| 10 | Overall, I am satisfied with the experience of using the chatbot for dengue diagnosis. | 4.20 | .414 |
| Total | | 4.25 | .623 |

## 6 DISCUSSIONS

The results of this study on the integration of ChatGPT with a chatbot for the preliminary diagnosis of dengue and addressing related issues show a promising outlook. In the first test, a diagnosis of dengue was achieved with an accuracy, sensitivity, and specificity of 83%, 80%, and 90%, respectively. This level of performance, although not perfect, suggests a significant ability to identify dengue patients. However, a noteworthy outcome is the substantial improvement in the second test after training the model with specific disease symptoms using a targeted prompt. Here, we achieved 100% accuracy, sensitivity, and specificity, indicating that the chatbot has become highly accurate in diagnosing dengue after adaptation.

When comparing the results with previous studies in the medical field, it is notable that the study conducted by the author [22] revealed that ChatGPT's accuracy rate in diagnosing diseases was variable, similar to the results of the first test in this study. Furthermore, researchers [21] found that ChatGPT-4 was able to correctly identify common cases but faced challenges with rare diseases. This highlights the importance of training models specifically for certain conditions, as was done in the second test of this study. On the other hand, the study [20] demonstrated the effectiveness of ChatGPT in ophthalmic diagnosis. Similarly, the chatbot in this study demonstrated effectiveness after training the model with specific dengue symptoms.

In this sense, the results of this study have significant practical implications. They suggest that the chatbot integrated with ChatGPT can be a valuable tool for preliminary medical care and the diagnosis of diseases such as dengue. This could improve access to healthcare in remote or resource-limited areas where specialized healthcare may be scarce. In addition, the high level of user satisfaction, with an average score of 4.25 out of 5, suggests that this technology has been positively received.

A significant limitation of this study is the small sample size of the test, which could impact the generalizability of the results. Moreover, while the model achieved high accuracy in the second test after being trained with specific dengue symptoms, this

adaptation restricts the chatbot's usefulness to other diseases or medical conditions. It is suggested that future work address these limitations and explore strategies to improve the versatility of chatbots in medical diagnosis.

## 7 CONCLUSION

This study successfully designed, developed, and evaluated a chatbot integrated with ChatGPT version GPT-3.5 for the preliminary diagnosis of dengue fever, yielding promising results. The study objective was satisfactorily achieved, providing an innovative and effective technological solution in the field of public health. Users expressed positive satisfaction with the chatbot. In addition, the chatbot's excellent performance in the second test, achieving 100% accuracy, sensitivity, and specificity, highlights its effectiveness in identifying positive and negative cases of the disease. As early detection of dengue is crucial for effectively managing and reducing of disease burden in affected communities, this tool can have a significant impact on streamlining the medical care process and improving clinical outcomes. In addition, the chatbot can contribute to raising awareness and educating people about dengue, promoting preventive measures, and healthcare practices. This study lays the groundwork for future advancements in the utilization of AI in disease detection and management, which has the potential to greatly influence public health. Although the diagnostic results of the chatbot integrated with ChatGPT are impressive, they should not replace the critical judgment of healthcare professionals. Moreover, its implementation must be done carefully and responsibly, always ensuring proper medical supervision.

## 8 REFERENCES

[1] P. K. R. Kalluru *et al.*, "Aminotransferases in relation to the severity of dengue: A systematic review," *Cureus*, vol. 15, no. 5, 2023. https://doi.org/10.7759/cureus.39436

[2] B. Hnusuwan, S. Kajornkasirat, and S. Puttinaovarat, "Dengue risk mapping from geospatial data using GIS and data mining techniques," *International Journal of Online and Biomedical Engineering (iJOE)*, vol. 16, no. 11, pp. 57–79, 2020. https://doi.org/10.3991/ijoe.v16i11.16455

[3] A. Tayal, S. K. Kabra, and R. Lodha, "Management of dengue: An updated review," *Indian Journal of Pediatrics*, vol. 90, no. 2, pp. 168–177, 2023. https://doi.org/10.1007/s12098-022-04394-8

[4] C.-H. Huang, Y.-T. Tsai, S.-F. Wang, W.-H. Wang, and Y.-H. Chen, "Dengue vaccine: An update," *Expert Review of Anti-infective Therapy*, vol. 19, no. 12, pp. 1495–1502, 2021. https://doi.org/10.1080/14787210.2021.1949983

[5] C. J. Stica, R. A. Barrero, R. Z. Murray, G. J. Devine, M. J. Phillips, and F. D. Frentiu, "Global evolutionary history and dynamics of dengue viruses inferred from whole genome sequences," *Viruses*, vol. 14, no. 4, p. 703, 2022. https://doi.org/10.3390/v14040703

[6] N. Ahmed, M. Shoaib, A. Ishaq, and A. Wahab, "Role of expert systems in identification and overcoming of dengue fever," *International Journal of Advanced Computer Science and Applications*, vol. 8, no. 10, 2017. https://doi.org/10.14569/IJACSA.2017.081011

[7] P. Somboonsak, "Development innovation to predict dengue affected area and alert people with smartphones," *International Journal of Online and Biomedical Engineering (iJOE)*, vol. 16, no. 2, pp. 62–79, 2020. https://doi.org/10.3991/ijoe.v16i02.12425

[8] A. Stojanov, "Learning with ChatGPT 3.5 as a more knowledgeable other: An autoethnographic study," *International Journal of Educational Technology in Higher Education*, vol. 20, no. 1, 2023. https://doi.org/10.1186/s41239-023-00404-7

[9] L. Xu, L. Sanders, K. Li, and J. C. L. Chow, "Chatbot for health care and oncology applications using artificial intelligence and machine learning: Systematic review," *JMIR Cancer*, vol. 7, no. 4, p. e27850, 2021. https://doi.org/10.2196/27850

[10] M. E. H. Kayesh, I. Khalil, M. Kohara, and K. Tsukiyama-Kohara, "Increasing dengue burden and severe dengue risk in Bangladesh: An overview," *Tropical Medicine and Infectious Disease,* vol. 8, no. 1, p. 32, 2023. https://doi.org/10.3390/tropicalmed8010032

[11] P. Foucambert *et al.*, "Efficacy of dengue vaccines in the prevention of severe dengue in children: A systematic review," *Cureus*, 2022. https://doi.org/10.7759/cureus.28916

[12] Z. Liao, J. Wang, Z. Shi, L. Lu, and H. Tabata, "Revolutionary potential of ChatGPT in constructing intelligent clinical decision support systems," *Annals of Biomedical Engineering*, 2023. https://doi.org/10.1007/s10439-023-03288-w

[13] M. E. Villanueva-Aguilar, L. del C. Rizo-de-la-Torre, M. del P. Granados-Muñiz, A. Montoya-Fuentes, and H. Montoya-Fuentes, "The genetic variant TNFA (rs361525) is associated with increased susceptibility to developing dengue symptoms," *Viral Immunology*, vol. 36, no. 3, 2023. https://doi.org/10.1089/vim.2022.0093

[14] A. Tlili *et al.*, "What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education," *Smart Learning Environments*, vol. 10, no. 1, 2023. https://doi.org/10.1186/s40561-023-00237-x

[15] Q. Lyu *et al.*, "Translating radiology reports into plain language using ChatGPT and GPT-4 with prompt learning: Results, limitations, and potential," *Visual Computing for Industry, Biomedicine, and Art*, vol. 6, no. 1, 2023. https://doi.org/10.1186/s42492-023-00136-5

[16] M. Cascella, J. Montomoli, V. Bellini, and E. Bignami, "Evaluating the feasibility of ChatGPT in healthcare: An analysis of multiple clinical and research scenarios," *Journal of Medical Systems*, vol. 47, no. 1, p. 33, 2023. https://doi.org/10.1007/s10916-023-01925-4

[17] J. Liu, T. Hu, Y. Zhang, X. Gai, Y. Feng, and Z. Liu, "A ChatGPT aided explainable framework for zero-shot medical image diagnosis," *ArXiv*, 2023. https://doi.org/10.48550/arxiv.2307.01981

[18] T. Hirosawa, Y. Harada, M. Yokose, T. Sakamoto, R. Kawamura, and T. Shimizu, "Diagnostic accuracy of differential-diagnosis lists generated by generative pretrained transformer 3 Chatbot for clinical vignettes with common chief complaints: A pilot study," *International Journal of Environmental Research and Public Health*, vol. 20, no. 4, p. 3378, 2023. https://doi.org/10.3390/ijerph20043378

[19] F. Tustumi, N. A. Andreollo, and J. E. de Aguilar-Nascimento, "Future of the language models in healthcare: The role of ChatGPT," *ABCD. Arquivos Brasileiros de Cirurgia Digestiva*, vol. 36, p. e1727, 2023. https://doi.org/10.1590/0102-672020230002e1727

[20] M. Balas and E. B. Ing, "Conversational AI models for ophthalmic diagnosis: Comparison of ChatGPT and the Isabel Pro Differential Diagnosis Generator," *JFO Open Ophthalmology*, vol. 1, p. 100005, 2023. https://doi.org/10.1016/j.jfop.2023.100005

[21] A. Alzahem, S. Latif, W. Boulila, and A. Koubaa, "Unlocking the potential of medical imaging with ChatGPT's intelligent diagnostics," *ArXiv*, 2023. https://doi.org/10.48550/arxiv.2305.07429

[22] L. Mehnen, S. Gruarin, M. Vasileva, and B. Knapp, "ChatGPT as a medical doctor? A diagnostic accuracy study on common and rare diseases," *medRxiv*, 2023. https://doi.org/10.1101/2023.04.20.23288859

[23] S. Huh, "Can we trust AI chatbots' answers about disease diagnosis and patient care?" *Journal of the Korean Medical Association*, vol. 66, no. 4, pp. 218–222, 2023. https://doi.org/10.5124/jkma.2023.66.4.218

[24] J. Zhou, T. Li, S. J. Fong, N. Dey, and R. G. Crespo, "Exploring ChatGPT's potential for consultation, recommendations and report diagnosis: Gastric cancer and gastroscopy reports' case," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 8, no. 2, pp. 7–13, 2023. https://doi.org/10.9781/ijimai.2023.04.007

[25] M. R. Mufid, A. Basofi, M. U. H. Al Rasyid, I. F. Rochimansyah, and A. rokhim, "Design an MVC Model using python for flask framework development," in *2019 International Electronics Symposium (IES)*, IEEE, 2019, pp. 214–219. https://doi.org/10.1109/ELECSYM.2019.8901656

[26] H. Qi, N. Funabiki, K. H. Wai, X. Lu, H. H. S. Kyaw, and W.-C. Kao, "An implementation of element fill-in-blank problems for code understanding study of JavaScript-based web-client programming," *International Journal of Information and Education Technology*, vol. 12, no. 11, pp. 1179–1184, 2022. https://doi.org/10.18178/ijiet.2022.12.11.1736

[27] O. López-Gorozabel, E. Cedeño-Palma, J. Pinargote-Ortega, W. Zambrano-Romero, and M. Pazmiño-Campuzano, "Bootstrap as a tool for web development and graphic optimization on mobile devices," in *Advances in Intelligent Systems and Computing*, 2021, vol. 1326, pp. 290–302. https://doi.org/10.1007/978-3-030-68080-0_22

[28] H.-C. Ling, K.-L. Hsiao, and W.-C. Hsu, "Can students' computer programming learning motivation and effectiveness be enhanced by learning Python language? A multi-group analysis," *Frontiers in Psychology*, vol. 11, 2021. https://doi.org/10.3389/fpsyg.2020.600814

[29] J. M. Sierra-Fernandez, O. Florencias-Oliveros, M. J. Espinosa-Gavira, J. C. Palomares-Salas, A. Agüera-Pérez, and J. J. González-de-la-Rosa, "Reconfigurable web-interface remote lab for instrumentation and electronic learning," *International Journal of Online and Biomedical Engineering (iJOE)*, vol. 16, no. 14, p. 69–80, 2020. https://doi.org/10.3991/ijoe.v16i14.16951

[30] D. Mamadou, K. J. Ayikpa, A. B. Ballo, and B. M. Kouassi, "Cocoa pods diseases detection by MobileNet confluence and classification algorithms," *International Journal of Advanced Computer Science and Applications*, vol. 14, no. 9, 2023. https://doi.org/10.14569/IJACSA.2023.0140937

## 9 AUTHORS

**Raúl Jáuregui-Velarde** has a Bachelor's degree in Systems and Computer Engineering. He has a background in computer system design, with a focus on artificial intelligence applications, machine learning, and data science. His research interests are in the area of computer science (E-mail: jaureguivraul@gmail.com).

**Domingo Hernández Celis** is a Doctor in Accounting; Doctor in Economics; Doctor in Administration; Master in Accounting and Financial Auditing; Certified Public Accountant; Independent auditor. General Manager of Microconsult-DHC & Asociados. Normal, distance and virtual undergraduate teacher; master's teacher; doctoral professor; Financial Advisor, with more than 30 years of professional practice and 20 years in teaching work (E-mail: dhernandez@unfv.edu.pe).

**Pedro Molina-Velarde** is a Systems Engineer, with a Master's degree in administration. Software development specialist. He is also an undergraduate teacher at private universities (E-mail: pmolina@utp.edu.pe).

**Cesar Yactayo-Arias** is a professional in administration and Master's in university teaching. Extensive teaching experience in higher education. In addition to experience in educational management is pursuing doctoral study in administration (E-mail: cyactayo@continental.edu.pe).

**Laberiano Andrade-Arenas** is a Systems Engineer, dedicated to scientific research. In addition, he is a teacher of Thesis courses, Scientific Research Methodology (E-mail: landrade@uch.edu.pe).