Artificial Intelligence-Based Mobile Application for Predicting Asthma Attacks

Larri R. Estrada-Leon [1], Segundo E. Cieza-Mostacero [1]

[1] *César Vallejo University, Avenida Larco 1770, Trujillo, Perú*

***Abstract* – This study presents applied research using a pure experimental design, with the objective of improving the prediction process of asthma attacks through a mobile application based on Machine Learning. For this purpose, software development tools such as Flutter, FastAPI, Render and Firebase were used. The results indicate an increase in the Number of New Cases to 98 cases, along with a decrease in the Time to Detection to 17 milliseconds in the experimental group compared to the control group. In addition, an 80% increase in the detection accuracy was observed. The comparison between both groups led to the acceptance of the alternative hypotheses, statistically demonstrating the effectiveness of the machine learning-based mobile application in improving the asthma attack prediction process. In conclusion, the remarkable improvements in the prediction of asthmatic attacks are highlighted, increasing both the number of new cases and detection accuracy while also reducing the detection time.**

***Keywords* – Machine Learning, mobile app, learning, health, asthma.**

DOI: 10.18421/TEM153-08
https://doi.org/10.18421/TEM153-08

Corresponding author: Larri R. Estrada-Leon,
César Vallejo University, Avenida Larco 1770, Trujillo, Perú
Email: lestradal@ucvvirtual.edu.pe

Received: 31 July 2025.
Revised: 16 February 2026.
Accepted: 23 February 2026.
Published: 27 August 2026.

1. Introduction

According to data from the World Health Organization (WHO), it is estimated that around 30 million people worldwide suffer from asthma, and approximately 210 million live with chronic obstructive pulmonary disease (COPD). In India, projections suggest that between 15 and 25 million citizens are affected by this respiratory condition, posing a significant challenge for the healthcare system, which must address the early detection of the disease in its vast population on a daily basis [1]. In the United Kingdom, approximately 5.4 million people are diagnosed with asthma, meaning that roughly one in ten individuals may experience an asthma attack at some point in their lives. Some of these attacks can be fatal, resulting in around 1,400 deaths each year [2]. For this reason, it is essential that asthma prediction models are highly sensitive and specific, in order to reduce mortality risk and avoid the unnecessary prescription of medications [3].

In Peru, the National Center for Epidemiology, Prevention and Disease Control (CENEP) reported 53,705 asthma cases as of July 2022. The Piura region was the most affected, accounting for 3,419 cases. In response, pulmonologist Dr. Alfredo Pachas warned of an unusual drop in temperature and a rise in respiratory cases, estimating that 90% of the population was experiencing pulmonary health issues [4].

Asthma is a complex disease influenced by environmental and genetic factors, allergens, pollution, stress, and infections [5]. Its manifestation varies widely among patients, who may go long periods without symptoms or experience sudden and severe attacks [6]. Moreover, current treatments for asthma attacks have changed little over time and are generally broad in scope, focused on immediate outcomes, and used in various contexts from self-care to hospital settings [7]. This lack of personalization and innovation in treatment may contribute to a gradual increase in asthma-related mortality.

Machine Learning (ML) has emerged as a leading technology, enabling models to learn autonomously from past data without human intervention [8]. In this context [9], ML refers to the ability of a system to improve performance through algorithms that learn to carry out specific tasks effectively.

To address this issue, a mobile application powered by Machine Learning is proposed as an effective solution to reduce undetected asthma attacks, helping patients assess their risk and receive timely treatment. As a result, the research question guiding this study is:

How does the use of a Machine Learning-based mobile application influence the improvement of asthma attack prediction in Trujillo?

The theoretical foundation of this study is supported by pioneering work in Machine Learning, which emphasized the relevance of predictive learning for identification and decision-making in healthcare. Other definitions [10] describe it as the field that studies algorithms capable of improving autonomously using past experience, encompassing a broad range of techniques aimed at learning from available data. This study also builds on prior research [11] focused on asthma attack prediction to assess whether previously proposed concepts remain valid or can be expanded.

Methodologically, observation forms were used to analyze changes in indicators such as the number of new cases, detection time, and prediction accuracy. From a practical standpoint, the study seeks to optimize the asthma attack prediction process in Trujillo through a machine learning-based mobile application. On a social level, the goal is to support the timely detection of asthma episodes, ultimately improving the well-being of Trujillo's population through faster and more effective interventions.

The general objective of this research is to improve the asthma attack prediction process through the use of a machine learning-based mobile application. Specific objectives include increasing the number of newly detected cases, reducing detection time, and improving the accuracy of predictions. Therefore, the general hypothesis is:

If a Machine Learning-based mobile application is used, then the asthma attack prediction process will significantly improve.

2. Methodology

The study led by [12] focused on developing a predictive machine learning model to forecast the occurrence of new asthma attacks in adults over a 12-week period. To achieve this, 360 patients diagnosed with asthma were enrolled, all of whom had used a smart inhaler connected to a mobile app.

The data showed that 64 individuals experienced at least one new episode, totaling 78 exacerbation events during the study period. The study concluded that such technological solutions significantly enhance the ability to predict asthma attacks, enabling timely clinical interventions. Additionally, it emphasized the role of digital technologies in anticipating medical events and supporting preventive actions through specialized tools.

In a complementary study, [13] aimed to establish a general framework for standardizing the recording of diagnostic data to facilitate disease prediction using machine learning techniques. The results showed that the Random Forest algorithm achieved an accuracy of 86.73% with a detection time of 6.3 seconds, while the Decision Tree algorithm reached an accuracy of 85.81% and a significantly shorter detection time of 0.11 seconds. Consequently, both models were deemed capable of providing an effective balance between accuracy and speed, supporting early clinical identification and decision-making in healthcare. The study also demonstrated the feasibility of reducing response time without compromising accuracy, which is beneficial for the timely detection of asthma attacks.

Furthermore, the research conducted by [14] implemented a predictive system for heart disease using machine learning techniques. The model was built using 1,670 medical records of individuals aged between 30 and 79. The results showed that the Random Forest algorithm delivered the highest performance, achieving a diagnostic accuracy of 93.8%. Based on these findings, the study concluded that the proposed approach not only offers strong predictive capabilities but can also be deployed via online platforms, facilitating the early identification of cardiovascular conditions. This research provides compelling evidence of the effectiveness of machine learning models in disease prediction, particularly in the context of cardiovascular health.

3. Materials and Methods

This section outlines the experimental methodology used, specifying the variables evaluated and the procedure followed to determine the effectiveness of the machine learning-based mobile application in predicting asthma attacks. It also describes the research design, participant selection and assignment criteria, data collection instruments, and the interventions carried out with both the experimental and control groups, aiming to compare their results.

3.1. Research Type and Design

The research adopted an applied research approach, consistent with the view of [15], who states that such studies use prior knowledge to generate tangible benefits for the target groups while contributing new insights to the field. A pure experimental design was employed, as defined by [16], which emphasizes two key aspects of validation: internal validity, ensuring that observed effects are attributable to the independent variable, and external validity, which supports the generalization of results to other contexts. Participants were randomly assigned to two groups: an experimental group composed of users of the mobile application and a control group consisting of individuals who did not use the application. This design allowed for an effective comparison of outcomes.

3.2. Variables and Operationalization

Machine learning refers to the study of computer algorithms capable of autonomously improving their performance through experience. This field encompasses various computational techniques designed to learn from available data to optimize results and make accurate predictions [10].

The prediction process is understood as an analytical decision tree used to estimate future events by analyzing past data, drawing on statistical models, data mining, and machine learning algorithms [17]. This process was assessed using an observation form that included indicators such as the number of detected cases, detection time, and accuracy level, all measured on a ratio scale.

3.3. Population, Sample, and Sampling

The study focused on individuals with asthma who were at risk of experiencing an asthma attack in the city of Trujillo. Due to limited available data, the exact number of individuals at risk in this area has not been determined.

A convenience sample of 30 asthmatic individuals from the at-risk population was selected. This sampling method enables efficient participant selection based on predefined criteria. Inclusion criteria covered both male and female participants aged 18 or older.

Individuals under the minimum age requirement were excluded. It is important to note that convenience sampling may introduce bias, which could affect the generalizability of the findings. This limitation should be considered when interpreting the study results.

3.4. Data Collection Techniques and Instruments

During the study, observation forms were used to record changes in key indicators, such as the number of new cases, detection time, and prediction accuracy. Additionally, the technique of indirect observation was applied as part of the data collection strategy. As noted by [16], this approach allows researchers to gather valuable information by observing the phenomenon in its natural setting without directly interfering with its progression.

3.5. Procedures

Data was collected in the pulmonology department of a clinic using observation forms to track changes in indicators such as the number of new cases, detection time, and accuracy level. Participants were randomly selected from at-risk patients waiting for their pulmonology consultation and were divided into two groups: an experimental group that used the mobile application, and a control group whose data were recorded manually.

In addition, to measure the number of new cases, the number of new people admitted to the clinic with a high probability of suffering an asthma attack was counted, differentiating between those who were evaluated using the mobile application and those who were evaluated manually. Priority was given to patients with a high probability, which optimized the use of time in the waiting rooms. Regarding detection time, the control group manually recorded the start and end times in the pulmonologist's module, while the mobile application automatically recorded the time it took for the model to generate a prediction.

Finally, to evaluate the accuracy rate, an observation form was used to analyze 30 medical records of asthma patients. The data contained in these records were compared with the predictions generated by the mobile application in order to assess its diagnostic accuracy. Furthermore, as noted in [18], machine learning algorithms play a key role in disease diagnosis by analyzing medical records and facilitating early detection and identification of health conditions. This method enabled a comprehensive evaluation of the software's diagnostic performance, and the observation forms provided detailed insights into the experience of both the control and experimental groups.

3.6. Development Trough the Mobile-D Methodology

Among the methodologies used for the development of information systems in mobile applications [19], the Mobile-D (MD) methodology stands out due to its iterative approach in each phase and its emphasis on product functionality. MD is an agile framework specifically designed for mobile app developers. It integrates well-established agile practices while prioritizing adaptability and functional value over rigid planning [20]. The Mobile-D methodology is structured into five stages: exploration, initialization, production, stabilization, and testing.

3.6.1. Exploration

In the initial phase of the project, the key stakeholders were identified, a preliminary draft of the system requirements was prepared, and the core functionalities were defined. In addition, the development tools were selected, use cases were specified, and the primary work sequence was organized. The software architecture was also designed, along with the diagrams needed to support the development process.

Regarding the technology stack, compatibility was prioritized throughout development. The machine learning model was trained using Visual Studio Code v1.101 with the Jupyter-Lab v2024.11.0 extension. A Firestore database hosted on Firebase was used to store questionnaire responses.

For the mobile application, Flutter SDK v3.27.2 was used in combination with Android Studio Koala 2024.1.1 Patch 1, ensuring compatibility with both iOS and Android. To enable communication between the app and the ML model, an API was developed using FastAPI v0.115.0 and deployed on a server. Through the application, users could complete a risk assessment questionnaire, and the ML model would analyze the responses to optimize the prediction of asthma attacks, as illustrated in Figure 1.

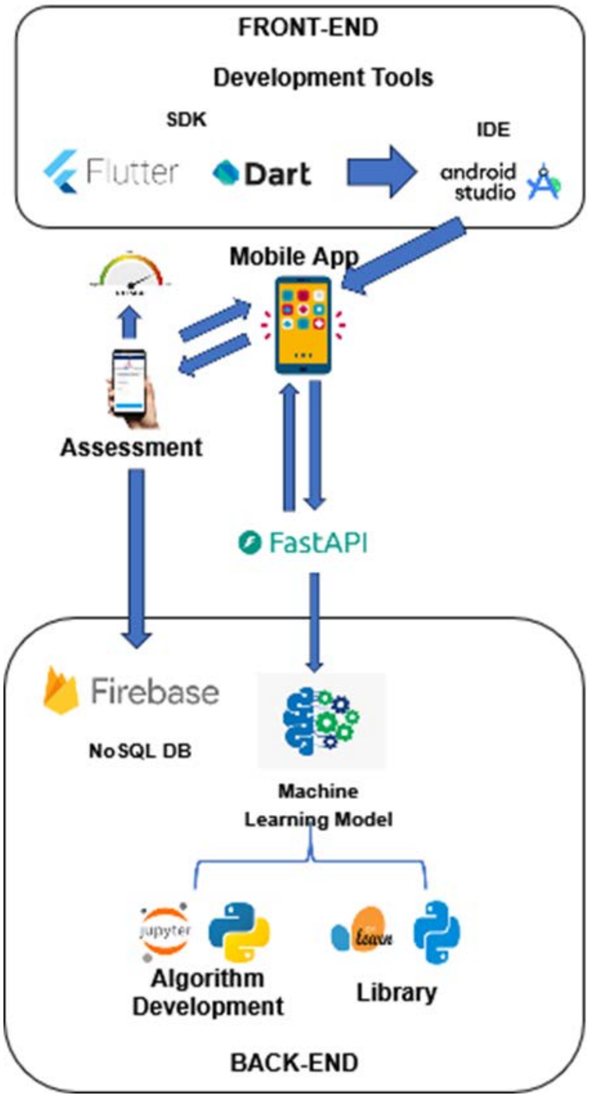

Figure 1. Software architecture

3.6.2. Initialization

The necessary resources were organized to move forward with the next phase. A key component of this stage was the creation of prototypes using Figma Desktop v125.4.7, which enabled a visual representation of the system's structure, functionality, and user experience. These prototypes supported early testing and iterative improvements, incorporating feedback collected through a survey administered to the study population. The comments focused on aspects such as usability, design, interaction, and overall applicability, helping to identify areas for refinement. As a result, the final prototypes of the mobile application were enhanced and optimized, as shown in Figure 2.

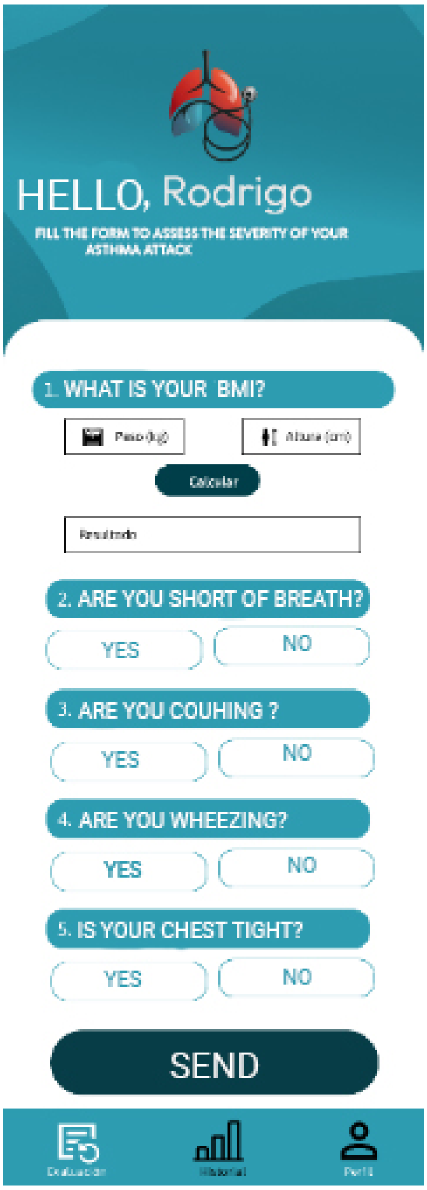

Figure 2. Risk assessment prototype

The main view prototype highlights the application's custom-designed logo, created from scratch to avoid issues with public image usage and to ensure compliance with copyright regulations. The questions used for assessing the risk of asthma attacks were developed and validated by medical experts, ensuring they are clear, consistent, and easy for patients to understand.

3.6.3. Production

During this stage, all required dependencies were installed in the integrated development environment (IDE), ensuring proper version management for Flutter, Android SDKs, and relevant libraries. The user interfaces were then built based on the previously validated prototypes, keeping alignment with the original design specifications.

To promote scalability and facilitate code maintenance, a clean architecture was implemented, incorporating dependency injection and the repository pattern. This structure ensured a clear organization of responsibilities between the presentation, domain, and data layers, facilitating efficient communication, easier testing, and simplified maintenance.

Regarding the machine learning component, a dataset from Kaggle was selected, containing various risk factors associated with asthma attacks. To tailor the dataset to the Peruvian context, only relevant factors for the local population were retained, as validated by medical experts. Due to class imbalance in the dataset, the Random Over-Sampler technique was applied to increase instances of the minority class, thereby balancing the data and enhancing the model's predictive performance, as illustrated in Figure 3.

```
# Random Over Sampler
from imblearn.over_sampling
import RandomOverSampler
oversampler = RandomOverSampler
(random_state=42)
balanced_data, balanced_labels = oversampler.
fit_resample(x_risk_factors,y_target)
healthy_patients = balanced_labels.sum()
at_risk_patients = balanced_labels.shape[0]
- healthy_patients
print(f'Healthy Patients:
      {healthy_patients} and
      At-Risk Patients:
      {at_risk_patients}')
```

Figure 3. Random Over Sampler (ROS)

Additionally, to implement the Random Over-Sampling (ROS) technique, it is first necessary to import it from the imblearn.over_sampling module, which is part of the imbalanced-learn library and must be installed beforehand. ROS is initialized by creating an instance of the RandomOverSampler class, setting the random_state parameter to a fixed value (42) to ensure reproducibility during training. The fit_resample method is then applied to the input data (x_risk_factors and y_target), adjusting and balancing the dataset by generating additional synthetic samples. This procedure yields the newly resampled sets: dataROS and targetROS. With the balanced dataset, a Random Forest (RF) model recognized in the literature for its effectiveness in pulmonary risk assessments was selected and trained. To ensure optimal performance, the balanced dataset (dataROS and targetROS) was split into 70% for training and 30% for testing. Additionally, the test set was further segmented to facilitate model validation, as shown in Figure 4.

```
#Random Forest (RF)
rf_ros_dataTrain,rf_ros_dataTemp,
rf_ros_targetTrain, rf_ros_targetTemp=
train_test_split(dataROS,targetROS,
test_size=0.30,random_state=42)
rf_ros_dataVal, rf_ros_dataPrueba,
rf_ros_targetVal, rf_ros_targetPrueba =
train_test_split(rf_ros_dataTemp,
rf_ros_targetTemp, test_size=0.50,
random_state=42)
start_train_time=time.time()
ros_modelRF=RandomForestClassifier
(random_state=42,criterion='gini',
max_features='sqrt' )
ros_modelRF.fit(rf_ros_dataTrain,
rf_ros_targetTrain)
end_train_time=time.time()
time_training_model=
(end_train_time-start_train_time)
print(f"Tiempo del entrenamiento
del modelo RF: {time_training_model}")
```

Figure 4. Random Forest (RF) model training

The model training began by splitting the dataROS dataset and targetROS labels into two portions: 70% for training and 30% reserved for validation and testing. This temporary set was then divided into two equal subsets one for validation (rf_ros_dataVal, rf_ros_targetVal) and another for testing (rf_ros_dataTest, rf_ros_targetTest). The ros_modelRF was created using the RandomForestClassifier class from the sklearn.ensemble library, a choice supported by studies highlighting its robustness in managing multiple decision trees and its effectiveness in medical risk assessment. Additionally, the training time was recorded to evaluate computational performance and measure the model's efficiency.

To assess the performance of the machine learning model, several metrics were calculated, including precision, accuracy, recall, F1 score, specificity, and the area under the ROC curve (AUC), as shown in Figure 5.

```
#Metricas Modelo Random Forest
def metrics_model_rf_ros(x, y):
  predictions =
  ros_modelRF.predict(x)
  accuracy_metric =
  accuracy_score(y, predictions)
  precision_metric =
  precision_score(y, predictions)
  recall_metric =
  recall_score(y, predictions)
  f1_metric =
  f1_score(y, predictions)
  tn, fp, fn, tp =
  confusion_matrix(y, predictions)
  .ravel()
  specificity_metric =
  tn / (tn + fp)
  confusion_mat =
  confusion_matrix(y, predictions)
  return {
    'Accuracy': accuracy_metric,
    'Precision': precision_metric,
    'Recall': recall_metric,
    'F1 score': f1_metric,
    'Specificity': specificity_metric
    }, confusion_mat

rf_ros_train_results, train_mat_conf_rf_ros=
metrics_model_rf_ros(
rf_ros_dataTrain, rf_ros_targetTrain)
print(f'**METRICAS DE ENTRENAMIENTO RF
:\n{rf_ros_train_results}')
rf_rus_probs = ros_modelRF.predict_proba
(rf_ros dataPrucba[:,1]
fpr.tpr.thresholds = roc_curve
(rf_ros_targetPrueba.rf_ros_probs)
roc_auc = auc(fpr.tpr)
print(f'**AUC curve:\n {roc_auc}')
```

Figure 5. Random Forest (RF) model performance metrics

The code snippet defines a function called metrics_model_rf_ros, which takes input data (x, y) and generates predictions using the Random Forest model trained on the balanced dataset created through the ROS technique. Based on these predictions, key performance metrics are calculated: 97% precision, 98% accuracy, 100% recall, 98% F1 score, and 97% specificity.

These metrics are stored in a dictionary along with the confusion matrix and are printed to assess the effectiveness of the machine learning model. Finally, the area under the ROC curve (AUC) is calculated using the predicted probabilities on the balanced dataset, yielding a value of 0.98. This result demonstrates the model's strong ability to distinguish between healthy individuals and those at risk, as illustrated in Figure 6.

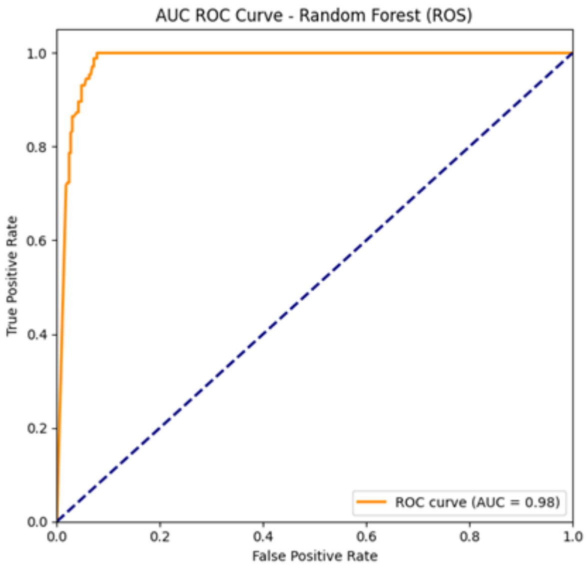

Figure 6. Random Forest (RF) model AUC ROC result

After training and evaluating the model, it was integrated into an API developed with FastAPI, enabling efficient and seamless communication between the mobile application and the model. This integration was implemented through a dedicated endpoint that facilitates data exchange in JSON format, thereby ensuring real-time risk assessment, as illustrated in Figure 7.

```
@app.post("/evaluation")
async def add_evaluation_user(request: Request):
    req_data = await request.json()
    BMI=req_data.get('questionIMC')
    Wheezing=req_data.get('questionWheezing')
    ShortnessOfBreath=req_data.get('questionShortnessOfBreath')
    ChestTightness=req_data.get('questionChestTightness')
    Coughing=req_data.get('questionCoughing')

    params_news_df = pd.DataFrame([[
        BMI, Wheezing,
        ShortnessOfBreath,
        ChestTightness,
        Coughing]],
    columns=["BMI", "Wheezing",
             "ShortnessOfBreath",
             'ChestTightness',
             'Coughing'])

    result_model = model_training
    .predict(params_news_df)

    if result_model[0] == 0:
        mensaje = "Su evaluación indica una " \
        "baja probabilidad de crisis asmatica."
    else:
        mensaje = "Su evaluación indica una " \
        "alta probabilidad de crisis asmatica."
    return JSONResponse(
        content={'AsthmaDiagnosis':
                 str(result_model),
                 "message": mensaje,
                 })
```

Figure 7. Random Forest (RF) model API endpoint

The code snippet defines a POST endpoint in FastAPI, accessible via the /evaluation route, which receives user data in JSON format. This data includes variables such as body mass index (BMI), wheezing, shortness of breath, chest tightness, and cough, sent from the mobile application. The API extracts this information using await request.json(), allowing each value to be accessed individually through the .get() method .

These values are then organized into a DataFrame called params_news_df, structured to be compatible with the trained model (model_training). The model processes this input and returns a binary prediction: 0 indicates a low probability of an asthma attack, while 1 indicates a high probability. Finally, the API responds with a JSON object via JSONResponse, which includes the prediction under the key AsthmaDiagnosis and an informative message for the user regarding their evaluated risk level.

3.6.4. Stabilization and Software Testing

In the final stages of mobile application development, all technological components defined in the software architecture were integrated to ensure efficient and consistent system interaction. This process enabled measurement of the API's response time during communication with the mobile application and ensured proper handling of data in the core risk evaluation process.

These validations were conducted through functional testing. To verify the performance of the asthma attack prediction model, the API was tested using the Postman tool. A POST request was sent to the /evaluation endpoint, hosted on the Render server, with JSON data representing input values from the mobile app, such as BMI and responses related to asthma symptoms. During the test, low-risk values were submitted, and the API successfully responded with a 200 OK status code, returning a message indicating a low probability of an asthma attack, along with the model's prediction, start and end processing times, and diagnostic result. This test confirmed that the API correctly handles valid requests and provides accurate predictions based on the received data. The APK file of the application is available at [21].

3.7. Data Analysis Methodology

Specific hypotheses were formulated for two indicators related to the dependent variable. For the indicator Number of New Cases, the alternative hypothesis (Ha) proposed that the use of a mobile application based on Machine Learning leads to an increase in the number of new cases during the post-test phase of the Experimental Group (NCNGE) compared to the post-test results of the Control Group (NCNCG). Similarly, for the Detection Time indicator, the alternative hypothesis (Ha) stated that the use of the Machine Learning-based mobile application reduces detection time in the post-test of the Experimental Group (DTGE) compared to that of the Control Group (DTCG). Lastly, for the Detection Accuracy indicator, the goal was to directly compare the actual clinical cases with the predictions generated by the application.

Descriptive Analysis: For the indicators Number of New Cases and Detection Time, the average values were calculated for both the experimental and control groups. A reference value was also established based on prior studies, and the number of observations exceeding both the mean and the reference value was recorded. Percentages were then calculated, and a descriptive table was created with both the mean and median using the Jamovi statistical software. Additionally, a detailed analysis of the collected data was conducted.

For the Detection Accuracy indicator, two tables were created with two columns each, registering whether the mobile application correctly identified the likelihood of an asthma attack based on 30 clinical histories. In each table, correct predictions by the app were marked with an "X," and the results were visually represented using a bar chart.

Inferential Analysis: The normality of the data was assessed. Given that the Number of New Cases and Detection Time indicators had fewer than 50 records, the Shapiro-Wilk test was applied. Results showed that none of the variables followed a normal distribution. Consequently, the non-parametric Mann-Whitney U test was used to determine whether the independent variable had a statistically significant effect on the dependent variable.

4. Results

This section presents the results obtained from the evaluation of the proposed mobile application based on Machine Learning, focusing on three key performance indicators: Number of New Cases (NCN), Detection Time (DT), and Detection Accuracy (DA). The analysis is divided into two main parts: a descriptive analysis, which outlines summary statistics and observed trends within the experimental and control groups; and an inferential analysis, which applies appropriate statistical tests to determine the significance of observed differences. Together, these findings aim to validate the impact of the mobile application on improving early asthma attack detection in a practical setting.

4.1. Descriptive Analysis

Figure 8 shows how the mobile application successfully processed 30 clinical records. Of these, 24 cases were accurately identified, while the remaining 6 resulted in false negatives. These results suggest that the machine learning–based tool demonstrates strong effectiveness in identifying genuine cases of patients diagnosed with asthma attacks.

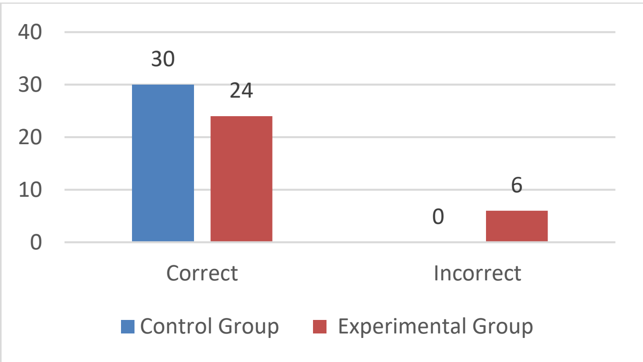

Figure 8. Mobile app effectiveness

The trained machine learning model demonstrates excellent performance, facilitating the accurate identification of individuals who may belong to previously undetected at-risk groups. Moreover, a precise predictive model not only contributes to the early detection of the disease but also enables the recognition of cases where patients appear healthy, even though the illness may be developing while undetected.

It is important to note that for the Detection Accuracy (DA) indicator, only a descriptive analysis was conducted. This is because the aim was to directly compare actual disease cases with the results provided by the application, assessing its effectiveness in a practical setting. Since the objective was not to make generalizable statistical inferences, but rather to validate the mobile application using specific data, the descriptive analysis proved to be the most appropriate and relevant method.

Table 1 shows that 44.83% of the values for the Number of New Cases (NCN) in the post-test of the Experimental Group were above its own group average. Similarly, 82.76% of the cases exceeded the predetermined target. Finally, it was confirmed that 100% of the NCN values in the Experimental Group's post-test surpassed the average recorded in the Control Group's post-test.

Table 1. Number of New Cases descriptive analysis

	CG Post-test	EG Post-test	
Average	1.48	3.38	
Target		2	
No higher than average	13	24	29
% Higher than average	44.83	82.76	100

Table 2 shows that 86.67% of the detection time (DT) values recorded during the post-test of the Experimental Group were below their own group average. Additionally, 33.33% of the cases fell below the predefined benchmark. Finally, it was found that 100% of the DT values in the Experimental Group's post-test were lower than the average recorded in the Control Group's post-test.

Table 2. Detection Time descriptive analysis

	CG Post-test	EG Post-test	
Average	23967	17	
Target		8	
No higher than average	26	10	30
% higher than average	86.07	33.33	100

4.2. Inferential Analysis

According to the data presented in Table 3, the indicator Number of New Cases (NCN) in the post-test for both the Control Group (CG) and the Experimental Group (EG) had a sample size of fewer than 50 observations. Therefore, the Shapiro–Wilk test was used to assess the normality of the data. In the CG, the p-value was 0.001, which is below the 0.05 threshold, indicating a non-normal distribution. Similarly, the EG yielded a p-value of 0.004, also confirming that the data is not normally distributed.

Table 3. Shaprio-Wilk normality test

Indicators	Statistics		p	
	CG	EG	CG	EG
NCN	0.710	0.882	0.001	0.004
DT	0.943	0.466	0.107	0.001

Regarding the Detection Time (DT) indicator in the post-test phase for both the Control Group (CG) and the Experimental Group (EG), the sample size was fewer than 50 records Therefore, the Shapiro–Wilk test was applied to assess the normality of the data. In the CG, a p-value of 0.107 was obtained, which suggests a normal distribution, as it exceeds the 0.05 threshold. In contrast, the EG yielded a p-value of 0.001, indicating a non-normal distribution.

For the Detection Accuracy indicator, only a descriptive analysis was conducted, as the primary objective was to directly compare the actual clinical cases with the outcomes generated by the mobile application in a practical setting. Since the aim was not to make generalizable statistical inferences, but rather to validate the application's performance using specific data, a descriptive approach was deemed the most appropriate and relevant.

It is important to underscore that in the post-test evaluation, the Detection Time values followed a normal distribution only in the Control Group, while the Experimental Group did not show this characteristic. Meanwhile, the Number of New Cases (NCN) data did not follow a normal distribution in either group. Therefore, the non-parametric Mann–Whitney U test was applied to compare both independent groups, as shown in Table 4.

Table 4. Mann–Whitney U test results per indicator

Indicators	Statistics	p
NCN	46	0.001
DT	0.00	0.001

The results from the statistical test presented in Table 4 reveal p-values of 0.001 for both indicators, well below the 0.05 threshold. This provides sufficient evidence to reject the null hypotheses and accept the alternative hypotheses. Thus, regarding the Number of New Cases, it is confirmed that the use of a mobile application based on Machine Learning significantly increases case detection. Likewise, for Detection Time, the findings validate that the use of such an application effectively reduces the time required to identify potential asthma attacks.

5. Discussion

This section examines the impact of using a mobile application based on Machine Learning to improve the process of predicting asthma attacks. It assesses the influence of the proposed method on three key indicators number of new cases detected, detection time, and detection accuracy by comparing the results of the experimental and control groups. The analysis is based on data collected during the study and contextualized with prior research on the subject.

Regarding the first indicator, Number of New Cases (NCN), 43 cases were recorded in the control group, while 98 were identified in the experimental group. Notably, all values in the experimental group exceeded the average recorded in the control group. These findings align with those reported in [12], where a predictive tool was developed to anticipate asthma attacks, identifying 78 new cases over a 12-week period. This highlights the effectiveness of technological solutions as essential tools in the healthcare field. Mobile applications enhanced with machine learning functionalities enable healthcare professionals to conduct faster and more accurate risk assessments, thereby increasing case detection and reducing the number of undiagnosed episodes.

With respect to the second indicator, Detection Time (DT), the control group recorded an average detection time of 23,967 milliseconds, while the experimental group averaged only 17 milliseconds. This reduction amounting to 6,967 milliseconds demonstrates that patients receiving preliminary results via the mobile application benefit from faster and more efficient evaluations. These findings are consistent with the study by [13], which proposed a general framework for standardizing diagnostic data entry, reducing the average detection time to 6.3 seconds and enhancing the ability to identify patients with a high likelihood of disease development.

Finally, for the third indicator, Detection Accuracy, 30 clinical records of patients diagnosed with asthma were analyzed.

The machine learning–based mobile application correctly evaluated 24 of these cases, resulting in an accuracy rate above 80%. This suggests that the trained model offers strong performance, accurately identifying individuals who may belong to previously unrecognized risk groups. Moreover, the results are comparable to those reported in [14], where 470 clinical records were correctly classified with an accuracy of 93.8%. This underscores the importance of developing highly precise machine learning models, as they enhance real-world case assessment and enable the identification of hidden patterns in patients who may appear healthy.

6. Conclusion

This section presents a summary of the study's findings, highlighting the notable improvements in asthma attack prediction achieved through the use of a mobile application based on machine learning. The statistical analyses revealed an increase in the number of newly detected cases, a reduction in detection time, and an improvement in detection accuracy.

The number of new cases recorded in the post-test of the Experimental Group (EG) doubled compared to the post-test results of the Control Group (CG). Moreover, the Mann-Whitney U test ($p < 0.001$) provided strong statistical evidence supporting the conclusion that the use of a Machine Learning–based mobile application contributed to the increase in new case detection. The detection time in the post-test of the Experimental Group was significantly lower than that of the Control Group. Similarly, the Mann-Whitney U test ($p < 0.001$) offered solid evidence confirming that the implementation of the mobile application effectively reduced detection time.

The detection accuracy reached an 80% improvement when evaluating actual cases of asthma attacks, representing a significant enhancement compared to previous methods and existing standards.

For future work, it is recommended to increase the sample size and expand the study's geographic coverage to validate the mobile application in more diverse settings. Additionally, integrating the tool with real-time healthcare systems could enhance clinical monitoring. Further improvements may include refining the machine learning model using more advanced techniques, incorporating telemedicine features to support patient follow-up, and embedding educational content to aid in early symptom identification. It is also suggested to assess the application's long-term clinical impact and to strengthen data security and privacy measures, ensuring an ethical and secure implementation.

References:

[1]. Arias Gonzáles, J. L. (2021). *Diseño y metodología de la investigación*. Enfoques Consulting EIRL.

[2]. Bendezú, R. (2022). *Asma: una enfermedad crónica por combatir en el Perú.* La República. Retrieved from: https://especial.larepublica.pe/asmazero/2022/08/25/asma-una-enfermedad-cronica-por-combatir-en-el-peru-259 [accessed: 15 June 2025]

[3]. Balasubramanian, S., & Rajadurai, P. (2024). Machine Learning-Based Classification of Pulmonary Diseases through Real-Time Lung Sounds. *International Journal of Engineering & Technology Innovation*, *14*(1). Doi: 10.46604/ijeti.2023.12294

[4]. Casusol-Campos, M. E., Cieza-Mostacero, S. E., & Luque-Sanchez, V. F. (2023). Influencia de Aplicación Móvil con Machine Learning en el Sector Salud: Una Revisión Sistemática de la Literatura. *Revista Ibérica de Sistemas e Tecnologias de Informação*, (E64), 366-377.

[5]. Cilluffo, G., et al. (2022). Machine learning: A modern approach to pediatric asthma. *Pediatric Allergy and Immunology*, *33*, 34-37. Doi: 10.1111/pai.13624

[6]. Estrada, L. (2025). *L. Executable software code*. Google Drive. Retrieved from: https://drive.google.com/drive/u/1/folders/141ctUNWzqnGvpi4JvCh6jz0CSRHpDhSb [accessed: 17 June 2025]

[7]. Howell, I., et al. (2024). How have we measured trial outcomes of asthma attack treatment? A systematic review. *ERJ Open Research*, *10*(1). Doi: 10.1183/23120541.00660-2023

[8]. Huang, A. A., & Huang, S. Y. (2023). Use of feature importance statistics to accurately predict asthma attacks using machine learning: a cross-sectional cohort study of the US population. *Plos one*, *18*(11), e0288903. Doi: 10.1371/journal.pone.0288903

[9]. Hussain, Z., Shah, S. A., Mukherjee, M., & Sheikh, A. (2020). Predicting the risk of asthma attacks in children, adolescents and adults: protocol for a machine learning algorithm derived from a primary care-based retrospective cohort. *BMJ open*, *10*(7), e036099. Doi: 10.1136/bmjopen-2019-036099

[10]. Leyva-Vásquez, L. M., et al. (2022). Multiplatform Mobile Application to Improve the Process of Tasks in an Agroexport Company. *Proceedings of the LACCEI International Multi-conference for Engineering, Education and Technology*. Latin American and Caribbean Consortium of Engineering Institutions. Doi: 10.18687/LACCEI2022.1.1.188

[11]. Lugogo, N. L., et al. (2022). A predictive machine learning tool for asthma exacerbations: results from a 12-week, open-label study using an electronic multi-dose dry powder inhaler with integrated sensors. *Journal of Asthma and Allergy*, 1623-1637. Doi: 10.2147/JAA.S377631

[12]. Lutimath, N. M., Sharma, N., & Byregowda, B. K. (2021). Prediction of heart disease using biomedical data through machine learning techniques. *EAI Endorsed Trans. Pervasive Health Technol.*, *7*(29). Doi: 10.4108/eai.30-8-2021.170881

[13]. Maini, E., et al. (2021). Machine learning–based heart disease prediction system for Indian population: An exploratory study done in South India. *Medical Journal Armed Forces India*, *77*(3), 302-311. Doi: 10.1016/j.mjafi.2020.10.013

[14]. Malik, S., et al. (2019). Data driven approach for eye disease classification with machine learning. *Applied Sciences*, *9*(14), 2789. Doi: 10.3390/app9142789

[15]. Molina Ríos, J. R., Honores-Tapia, J. A., Pedreira-Souto, N., & Pardo-León, H. P. (2021). Comparativa de metodologías de desarrollo de aplicaciones móviles. *3C Tecnología: glosas de innovación aplicadas a la pyme, 10*(2), 73–93. Doi: 10.17993/3ctecno/2021.v10n2e38.73-93

[16]. Moslemi, A., et al. (2022). Differentiating COPD and asthma using quantitative CT imaging and machine learning. *European Respiratory Journal*, *60*(3). Doi: 10.1183/13993003.03078-2021

[17]. IBM. (n.d.). *¿Qué es el análisis predictivo?*. IBM. Retrieved from: https://www.ibm.com/mx-es/topics/predictive-analytics [accessed: 20 January 2026]

[18]. Sabry, A. H., et al. (2024). Lung disease recognition methods using audio-based analysis with machine learning. *Heliyon*, *10*(4). Doi: 10.1016/j.heliyon.2024.e26218

[19]. Tibble, H., et al. (2019). Predicting asthma attacks in primary care: protocol for developing a machine learning-based prediction model. *BMJ open*, *9*(7), e028375. Doi: 10.1136/bmjopen-2018-028375

[20]. Tsang, K. C. H., et al. (2022). Predicting asthma attacks using connected mobile devices and machine learning: the AAMOS-00 observational study protocol. *BMJ open*, *12*(10), e064166. Doi: 10.1136/bmjopen-2022-064166

[21]. Cordero, Z. R. V. (2009). La investigación aplicada: una forma de conocer las realidades con evidencia científica. *Revista educación*, *33*(1), 155-165. Doi: 10.15517/revedu.v33i1.538