

# Low-cost semi-automatic arrhythmia detector using ECG signal

Dair Alegría, Jorge Medina, Erika Uchuya, Renato Rojas and Fernando Puipulivia

**Abstract**—Cardiac arrhythmias are common heart rhythm disorders with significant global impact. They are a leading cause of death worldwide, accounting for 31 percentage of all deaths according to the World Health Organization. Atrial fibrillation, ventricular tachycardia, and sinus bradycardia are among the most frequent types of arrhythmias. Researchers have developed low-cost devices and machine learning algorithms to detect and classify arrhythmias using electrocardiogram signals. Improving access to specialized healthcare, particularly in rural areas, is crucial in addressing the disparities in diagnosing and managing cardiac arrhythmias. Early evaluation using cost-effective devices can be a valuable tool in identifying potential heart rhythm abnormalities and guiding patients towards appropriate care, preventing emergencies and long-term complications.

**Index Terms**—Cardiac arrhythmias, low-cost devices, electrocardiogram signals, diagnosing.

## I. INTRODUCCIÓN

Las arritmias cardíacas son trastornos del ritmo cardíaco que afectan a millones de personas en todo el mundo. Según la Organización Mundial de la Salud, las enfermedades cardiovasculares, incluyendo las arritmias, son la principal causa de muerte en todo el mundo, representando el 31% de todas las muertes [1]. Entre los tipos de arritmias más frecuentes se tiene la fibrilación auricular, la taquicardia ventricular y la bradicardia sinusal. La fibrilación auricular es el tipo de arritmia cardiaca más común a nivel mundial diagnosticada en el 1% a 2% de la población. Además, afecta a 1 de cada 25 adultos mayores a 60 años y a 1 de cada 10 adultos mayores de 80 años [2]. En el Perú, de acuerdo a un estudio observacional, el 17,49% de pacientes ingresados al servicio de cardiología de un hospital público del Perú padecen esta condición [3]. Respecto a la taquicardia ventricular, el 15% de pacientes con enfermedad coronaria tienen esta condición [4]. Además, es el causa más común de muerte cardiaca súbita en pacientes con cardiomiopatías isquémicas y no isquémicas [5], pues del 5% al 10% de pacientes con síndrome coronario agudo son diagnosticados con arritmias ventriculares, entre ellas, la taquicardia ventricular [6]. Por otro lado, la bradicardia sinusal es común en adultos mayores de 65 años y atletas jóvenes de ambos sexos. Uno de cada 600 adultos mayores de 65 años tiene disfunción del nodo sinusal; sin embargo, se requiere mayor investigación para recolectar data epidemiológica a nivel mundial específicamente en pacientes con bradicardia sinusal. [7]

Mañay, E. et al. (2022) desarrollaron un electrocardiógrafo de bajo costo que utiliza una red neuronal 1D entrenada con la base de datos MIT-BIH para detectar y clasificar arritmias en tiempo real. A pesar de detectar el 98,6% de arritmias,

se identificó la limitación de no poder clasificar los subtipos específicos de arritmias [8]. Dasgupta, A. & Ghosh, S. (2020) desarrollaron un sistema para la clasificación de arritmias en base a los datos espectrales de ECG a partir de tres algoritmos: Multi-Layered perceptron, Support Vector Machine y Random Forest. El sistema fue entrenado con las bases de datos Standford-Irhythmic y MIT-BIH. La mayor precisión fue de Random Forest con un 98.50% [9]

Así mismo, estudios sugieren que la clasificación basada en los síntomas del paciente como síntomas de arritmia pueden confundirse con otras enfermedades. Las posibilidades de error en el diagnóstico clínico pueden reducirse si los algoritmos de Machine Learning se entrena para reconocer un patrón específico de la señal de electrocardiograma (ECG).

Por ello, el objetivo del estudio es desarrollar un dispositivo para la detección automática de arritmias de bajo costo que por medio de adquisición de señales cardíacas pueda clasificar de manera preliminar si el paciente presenta o no presenta una arritmia cardíaca.

## II. PROBLEMÁTICA

En el Perú existe una brecha entre zonas rurales y urbanas en el acceso de servicios médicos y falta de personal médico especializado ya que de cada diez médicos colegiados, seis se concentran en Lima y Callao. Si bien Perú es un país en vías de desarrollo por ahora las desigualdades recaen en las personas que se encuentran alejadas de las grandes ciudades y sin la opción de acceder a un sistema de salud de calidad y sobre todo especializado. Los bajos salarios, las exigencias del trabajo o el difícil acceso a zonas rurales o alejadas traen como consecuencia una distribución desigual de especialistas en hospitales y centros de salud.

La evaluación del latido cardíaco mediante ECG como un primer paso de descarte se reconoce como una herramienta sólida para detectar ritmos cardíacos normales, ya que la detección de una arritmia indica una alteración del patrón común de la onda. A pesar de no brindar un diagnóstico específico de las arritmias, la evaluación preliminar del latido cardíaco mediante un dispositivo de bajo costo resulta útil en áreas remotas donde el acceso a cardiólogos es limitado. Esta evaluación permite identificar posibles anomalías y orientar a los pacientes hacia un examen de electrocardiograma con un especialista, lo cual es fundamental para obtener un diagnóstico más preciso y completo.

No detectar problemas cardíacos o no llevar un registro de la actividad cardíaca de pacientes que padecen de arritmias en zonas rurales puede derivar en situaciones de emergencia

que requieran de atención inmediata; debido a la brecha de infraestructura esto puede resultar fatal para el paciente lo que generaría problemas permanentes que puede significar un gasto económico en una familia ya de por sí con bajos recursos.

### III. PROPUESTA DE SOLUCIÓN

Se propone un dispositivo de bajo costo que haga uso de Machine Learning para el descarte preliminar de arritmias cardíacas, siendo el modelo entrenado con 4 tipos de arritmias cardiacas (contracción ventricular prematura (PVC), contracción supraventricular prematura, latido ectópico y patología desconocida). Para el desarrollo del dispositivo se implementó un modelo de redes neuronales haciendo uso de la técnica de Autoencoders para la clasificación de una señal ECG de derivación I, así como una Interfaz Gráfica de Usuario (GUI) para el funcionamiento del dispositivo, llamada Predictor ECG250.

El proceso para el desarrollo del modelo de clasificación contempla:

- Adquisición de señales: Se usó el dataset ECG5000 que cuenta con 7600 datos de entrenamiento incluyendo 1900 datos de prueba. Cada dato consta de la medición de un ciclo cardíaco interpolado a 140 muestras. Además, los datos están clasificados en 5 categorías: ritmo normal, contracción ventricular prematura (PVC), contracción supraventricular prematura, latido ectópico y patología desconocida

- Pre-procesamiento (filtrado y normalización): Se aplicó inicialmente un filtro Notch con frecuencia rechazada en 60 Hz y, posteriormente un filtro pasabajo con frecuencia de corte de 35 Hz. Luego, se realizó la normalización de la señal hacia un rango de 0 a 1.

- Clasificación: Debido a la diferencia en la cantidad de datos disponibles entre señales ECG de pacientes sanos y señales arrítmicas, se optó por utilizar Autoencoders, que son redes neuronales artificiales, entrenadas de manera no supervisada. El modelo cuenta con 3 capas de codificación y un proceso de optimización tipo “Adam”.

El proceso para el desarrollo de la interfaz Predictor ECG250 contempla:

- Adquisición de señales: Se obtienen 5 segundos de medición de señal ECG provenientes de la derivación I, a partir de comunicación serial entre la laptop y el Arduino UNO, el cual está conectado al módulo AD8232.

- Pre-procesamiento (filtrado y normalización): Se aplica el mismo pre-procesamiento realizado para las señales con las que fue entrado el modelo de redes neuronales. Es decir, un filtro Notch con frecuencia rechazada en 60 Hz y un filtro pasabajo con frecuencia de corte de 35 Hz. Luego, se realiza la normalización hacia un rango de 0 a 1.

- Segmentación e Interpolación: Se realiza una segmentación entre picos R-R de los 5 segundos de señal preprocesada, a fin de obtener cada latido por separado. Después, se interpola la cantidad de muestras a 140 para cada latido, a fin de tener la misma cantidad de muestras que los datos con los que fue entrenado el modelo de clasificación.

- Clasificación: Se aplica el modelo en Formato de Datos Jerárquicos HDF creado anteriormente para la clasificación de los latidos procesados.

- Visualización: Se generó una GUI desarrollada en el lenguaje Python 3.7, que cuenta con un botón digital para iniciar el pre-procesamiento, la segmentación e interpolación de la señal. Además, se cuenta con un botón para realizar el descarte de arritmias, en el que se indicará si se trata de un ritmo normal o de una señal arrítmica.

Con respecto al Hardware, se utilizaron componentes para la adquisición de señales y transmisión de datos de manera serial a la GUI como: un módulo AD8232 ECG pulso cardíaco, un Arduino UNO, cables de conexión a electrodos junto con sus respectivos electrodos, un pulsador el cual tiene la función de permitir al usuario que manipule el dispositivo controlar el momento del inicio de la recolección de datos, un LED verde el cual indica que el dispositivo se encuentra encendido y finalmente un LED azul el cual se enciende cuando se inicia el proceso de recolección de datos y se apaga cuando termina este último proceso. Posteriormente se hizo la implementación y programación necesaria en Arduino IDE para la adquisición de señales del cuerpo humano. Debido a que es un dispositivo de bajo costo se buscó reducir el presupuesto por lo que la carcasa se realizó con material MDF y corte láser.

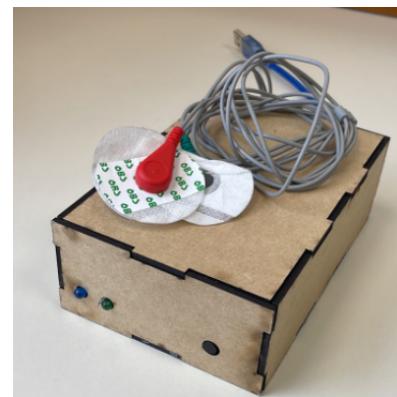


Fig. 1: Resultado final del proyecto

## IV. RESULTADOS

### A. Resultados del modelo

Antes de explicar los resultados del proyecto se debe explicar el funcionamiento del modelo a usar. El modelo que se implementó hace uso de los autoencoders y de forma complementaria se usa el dataset ECG500. Cada dato contiene un ciclo cardíaco con 140 muestras, y que puede pertenecer a una de 5 categorías:

1. Normal
2. Anormal: contracción ventricular prematura
3. Anormal: contracción supra-ventricular prematura
4. Anormal: latido ectópico
5. Anormal pero patología desconocida

Los datos con los que ha sido entrenado el modelo es separando los latidos en sus picos R-R. En las siguientes imágenes se puede observar la comparación entre un latido normal comparado con cada categoría

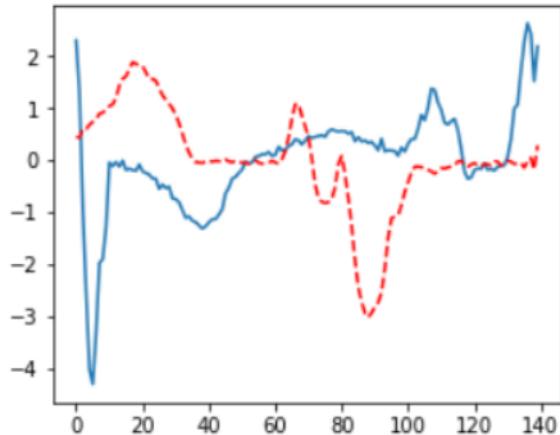


Fig. 2: Comparación de un dato normal (azul) y contracción ventricular prematura (rojo)

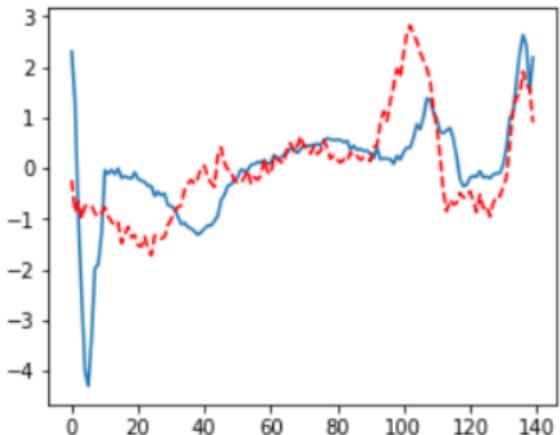


Fig. 3: Comparación de un dato normal (azul) y contracción supra-ventricular prematura (rojo)

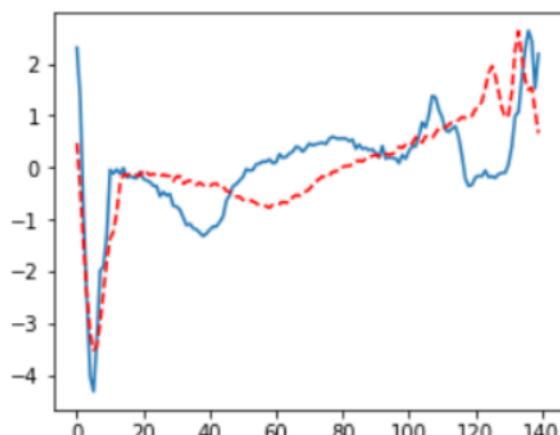


Fig. 4: Comparación de un dato normal (azul) y latido ectópico(rojo)

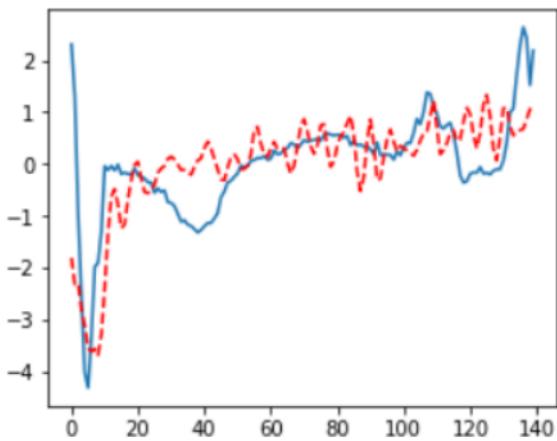


Fig. 5: Comparación de un dato normal (azul) y una patología extraña (rojo)

Luego el modelo hace uso de un autoencoder el cual posee tres capas de codificación y se entrena únicamente con datos de latidos normales, debido a la falta de datos de las otras arritmias. Así, al reconstruir un dato anormal el error será alto, por eso mismo se realiza el autoencoder con el optimizador “Adam” y usando como pérdida el MAE (Mean Absolute Error)

$$MAE = \frac{\sum_{i=1}^{140} |x_i - \hat{x}_i|}{140} \quad (1)$$

Como se mencionó anteriormente, en las imágenes se observa que la señal anormal reconstruida posee más error que la reconstrucción de la señal normal.

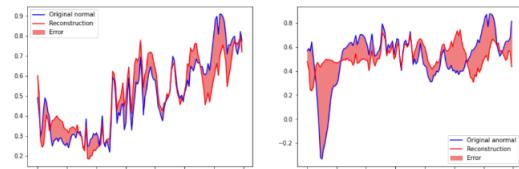


Fig. 6: Error entre señal obtenida y reconstruida

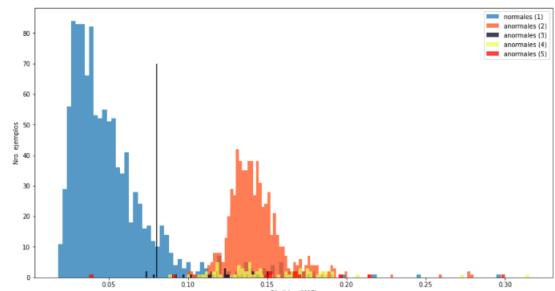


Fig. 7: Distribuciones de los errores de construcción por categoría

Finalmente el modelo clasifica las señales haciendo uso de la especificidad, el cual mide la proporción de normales que fueron detectados correctamente como normales. Una especificidad del 100 por ciento detectará a todos los pacientes sanos.

Mientras que también se calcula la sensibilidad para medir la proporción de anormales que fueron detectados correctamente como anormales. Una sensibilidad del 100 por ciento detectará a todos los pacientes enfermos.

TABLE I: Cálculo de especificidad y sensibilidad para cada categoría

	cat. 1	cat. 2	cat. 3	cat. 4	cat. 5
especificidad	88.4	-	-	-	-
sensibilidad	100	90.9	100.0	90.9	

### B. Resultados del proyecto

Primero se debe de obtener la señal fisiológica de tres sujetos de prueba, los cuales se sabe de antemano de que no poseen ningún tipo de arritmia cardíaca, esto va a ser relevante más adelante al momento de calcular la especificidad ya se realizarán 10 mediciones para cada sujeto de prueba donde se evaluará el funcionamiento del dispositivo. Una vez que la señal ha sido adquirida ,se procede a realizar un procesamiento digital mediante Python para eliminar cualquier tipo de ruido, especialmente la interferencia eléctrica el cual se encuentra generalmente a 60 Hz en el dominio de frecuencia compleja. Para ello se decidió usar un filtro pasa baja con frecuencia de corte en 35 Hz para asegurarnos que la señal esté correctamente filtrada y posteriormente se realiza la segmentación entre los picos R-R.Después, se interpola la cantidad de muestras a 140 para cada latido, a fin de tener la misma cantidad de muestras que los datos con los que fue entrenado el modelo de clasificación. En las siguientes figuras se puede observar paso a paso como se realizó lo anteriormente expuesto.

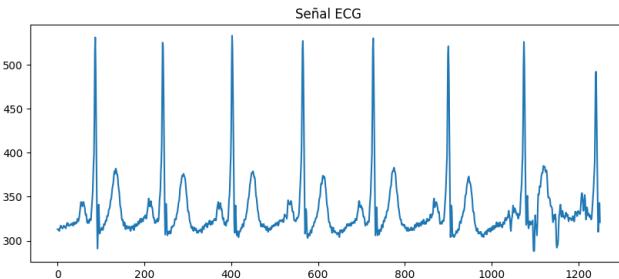


Fig. 8: Señal ECG adquirida del sujeto de prueba

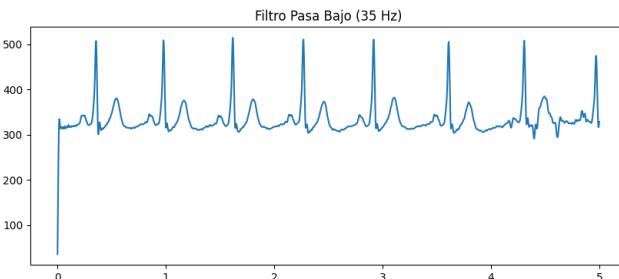


Fig. 9: Señal filtrada con filtro pasa baja

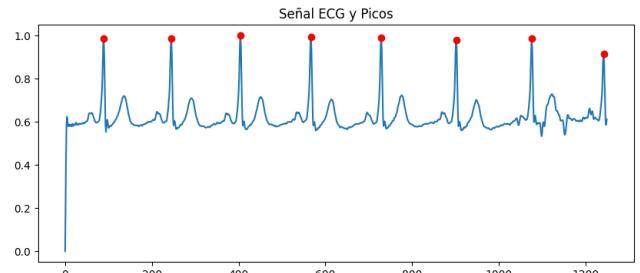


Fig. 10: Segmentación de la señal ECG entre picos R-R

Después de realizar la segmentación se procede a realizar una interpolación. Luego se implementa la señal interpolada en el modelo para obtener el diagnóstico del paciente en base a la cantidad de latidos normales frente a los arrítmicos. Es decir, en caso 10 latidos 6 son clasificados como normales y 4 son clasificados como arrítmicos entonces se concluye que la señal es no arrítmica.

Finalmente se debe calcular la especificidad, esto debido al propósito del proyecto el cual es priorizar el descarte de una arritmia mas no diagnósticas el tipo de ésta. Para ello se realizó 10 mediciones a 3 sujetos de prueba, los cuales se sabe de antemano que no tienen ningún tipo de arritmia con el objetivo de poder medir las veces que el Predictor ECG 250 es capaz de distinguir correctamente entre una señal arrítmica o no arrítmica. Por ello se opta por medir la especificidad y no selectividad. En la siguiente tabla se puede ver los resultados de las mediciones para cada sujeto de prueba y el cálculo de la especificidad del Predictor ECG 250.

TABLE II: Resultados de especificidad del Predictor ECG 250

	medición correcta (normal)	medición incorrecta (arritmia)	especificidad
sujeto 1	8	2	80%
sujeto 2	7	3	70%
sujeto 3	9	1	90%

Con los datos obtenidos de forma experimental se calcula la especificidad del Predictor ECG 250.

$$\text{especificidad} = \frac{80\% + 70\% + 90\%}{3} = 80\% \quad (2)$$

### V. DISCUSIÓN

Como se puede observar la especificidad es alta para poder evaluar si una persona tiene un problema cardíaco, lo cual es útil para el contexto socio-económico en el que se decidió enfocar el dispositivo. Aún así, como propuesta de mejora a futuro se propone realizar un aumento de la data disponible para otros tipos de arritmias y de esta forma poder tener la opción que el dispositivo pueda detectar tipos de arritmias de manera específica.

### VI. CONCLUSIONES

Se ha logrado desarrollar un dispositivo de detección semi-automática de arritmias de bajo costo basado en la adquisición de señales cardíacas a través de un módulo AD8232 vinculado a un algoritmo de Machine Learning realizado en Python. Los resultados obtenidos demuestran una alta especificidad

en la toma de decisión del algoritmo acerca de la presencia o ausencia de arritmia cardíaca en los sujetos de prueba. El dispositivo ofrece una solución eficiente y asequible para la detección temprana de las arritmias cardíacas, lo que puede contribuir a un diagnóstico preciso y oportuno. Su bajo costo y accesibilidad se presentan como una opción factible para mejorar el acceso a la detección de arritmias en zonas rurales, permitiendo así una alerta de prevención rápida y económica. Por último, las mejoras a futuro del Predictor ECG 250 radican en el aumento de la cantidad de datos para el entrenamiento del algoritmo de Machine Learning al mismo tiempo de poder hacer capaz al modelo basado en Autoencoders de poder realizar la detección específica de uno o varios tipos de arritmias cardíacas.

## VII. REFERENCIAS

1. “Cardiovascular diseases (CVDs).” [https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-\(cvds\)](https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)) (accessed Jul. 07, 2023).
2. A. S. Go et al., “Prevalence of Diagnosed Atrial Fibrillation in Adults,” *Jama*, vol. 285, no. 18, p. 2370, 2001, doi: 10.1001/jama.285.18.2370.
3. D. Chambergo-Michilot, B. Velit-Rios, and A. Cuevas-Parra, “Prevalence of cardiovascular diseases in the Dos de Mayo National Hospital in Peru,” *Rev. Mex. Angiol.*, vol. 48, no. 3, pp. 84–89, 2020, doi: 10.24875/RMA.20000012.
4. “Ventricular Tachycardia - StatPearls - NCBI Bookshelf.” <https://www.ncbi.nlm.nih.gov/books/NBK532954/> (accessed Jul. 07, 2023).
5. C. Pandozi et al., “Mapping of ventricular tachycardia in patients with ischemic cardiomyopathy: Current approaches and future perspectives,” *Clin. Cardiol.*, vol. 42, no. 10, pp. 1041–1050, 2019, doi: 10.1002/clc.23245.
6. S. Gupta, G. S. Pressman, and V. M. Figueredo, “Incidence of, predictors for, and mortality associated with malignant ventricular arrhythmias in non-ST elevation myocardial infarction patients,” *Coron. Artery Dis.*, vol. 21, no. 8, pp. 460–465, 2010, doi: 10.1097/MCA.0b013e32834022fa.
7. M. Brodsky, D. Wu, P. Denes, C. Kanakis, and K. M. Rosen, “Arrhythmias documented by 24 hour continuous electrocardiographic monitoring in 50 male medical students without apparent heart disease,” *Am. J. Cardiol.*, vol. 39, no. 3, pp. 390–395, 1977, doi: 10.1016/S0002-9149(77)80094-5.
8. E. D. Mañay, D. Martínez, M. D. Chiliquinga, V. D. Criollo, E. F. Rivera, and R. M. Toasa, “Low-cost deep learning-based architecture for detecting cardiac arrhythmias in ECG signals,” *Iber. Conf. Inf. Syst. Technol. Cist.*, vol. 2022-June, no. June, pp. 22–25, 2022, doi:0.23919/CISTI54924.2022.9820243.
9. A. Dasgupta and S. Ghosh, “Detection of Cardiac Arrhythmia using Spectral features,” *Proc. 2020 IEEE Int. Conf. Mach. Learn. Appl. Netw. Technol. ICMLANT 2020*, 2020, doi: 10.1109/ICMLANT50963.2020.9355982.



**Héctor Dair Alegría Cortez** Ninth semester student of the PUCP-UPCH Biomedical Engineering degree with interest in Project Management in health. Junior Consultant in Integrated Management Systems, certified in Interpretation of ISO 9001:2015 and ISO 45001:2018. Currently working as an Execution Assistant in the development of projects related to Medical Signals and Images in the PUCP Bioengineering Section.



**Jorge Eduardo Medina Celiz** Seventh semester student of the PUCP-UPCH Biomedical Engineering degree with interest in digital electronics and analog electronics to be applied in the health sector. He plans to specialize in the area of Biomechanics and Rehabilitation.



**Erika Uchuya Trocones** Eighth semester student of the PUCP-UPCH Biomedical Engineering degree interested in the area of clinical engineering, especially in medical devices. Currently working on the continuous development of the acquisition flow of a medical equipment in a private company.



**Felix Renato Rojas Arellanos** Seventh semester student of the PUCP-UPCH Biomedical Engineering degree with interest in the area of Tissue Engineering and Biomaterials.



**Fernando Eduardo Puipulivia Zárate** Seventh semester student of the PUCP-UPCH Biomedical Engineering degree with an interest in solving the structural problems that exist in the Peruvian health system. Plans to specialize in the area of Clinical Engineering.