

# Benemérita Universidad Autónoma de Puebla

Facultad de Ciencias Físico Matemáticas

Diseño e instrumentación de un sistema de análisis de  
señales biológicas para diagnóstico de COVID-19.

Tesis presentada al  
Colegio de Física

Como requisito parcial para la obtención del grado de

LICENCIADO EN FÍSICA

Por

Guadalupe Sandoval Moreno

Asesorada por

Dr. Eduardo Moreno Barbosa & Dr. Benito de Celis Alonso

Puebla Pue. 21 de diciembre de 2020

**Título:** Diseño e instrumentación de un sistema autónomo de señales biológicas para la mitigación en la propagación del COVID-19.

**Estudiante:** GUADALUPE SANDOVAL MORENO

COMITÉ

---

Dr. Javier Miguel Hernández López  
Presidente

---

Dr. Ricardo Agustín Serrano  
Secretario

---

M.C. Margarita Amaro Aranda  
Vocal

---

Dr. Eduardo Moreno Barbosa  
Asesor

---

Dr. Benito de Celis Alonso  
Asesor

En memoria del amor más grande de mi vida y mi mayor tesoro, mamá...

**Agradecimientos:**

*Agradezco infinitamente a mi familia por su paciencia, comprensión, amor. Especialmente a mi mamá, que se desveló muchas veces acompañándome a lo largo de toda la carrera, que me escuchó innumerables veces, me motivó cuando quise rendirme y me alentó siempre; a Dulce que muchas veces sin tener idea de lo que leía, trataba de ayudarme, por escucharme, corregirme e incluso regañarme, a César por darme motivos para seguir cuando perdí el rumbo.*

*Gracias a Manuel, por todo el tiempo, comprensión, apoyo, sugerencias, aliento, compañía a lo largo de la carrera, cariño. Gracias por la guía, sin ti no hubiese llegado a este momento.*

*A todos mis amigos por hacer más liviano el peso del día a día, por tanto cariño, risas y pláticas larguísimas que nunca acaban, en especial a Mario, Andrea y Jorge. A la FCFM por cambiar mi perspectiva acerca de la vida y al Cuerpo Académico de Física Médica por aceptarme tan gratamente y por darme la oportunidad de colaborar con ustedes.*

**Índice general**

**Lista de abreviaturas**

**Lista de figuras**

**Lista de tablas**

## **1. INTRODUCCIÓN**

**1.1 Justificación**

**1.2 Hipótesis**

**1.3 Objetivo principal**

**1.4 Objetivos secundarios**

**1.5 Resumen**

## **2. MARCO TEÓRICO**

**2.1. Pandemias**

**2.2. Protocolos de uso ante una pandemia**

**2.3. Salud digital**

**2.4. COVID-19**

**2.4.1. Propagación del SARS-Cov-2**

**2.4.2. Sintomatología del COVID-19**

**2.4.3. Diagnóstico de COVID-19**

**2.4.4. Inaccesibilidad de la población mexicana a las pruebas diagnósticas de SARS-Cov-2**

**2.5. Interpretación de los sonidos**

**2.5.1. Sonograma**

**2.5.2. Calibración**

**2.6. Aprendizaje Automático**

**2.6.1. Procedimiento de clasificación.**

## **3. MATERIALES Y MÉTODOS**

**3.1. Características y costos de los materiales**

**3.2. Arreglo experimental**

**3.3. Bases de datos de tos**

### **3.4. Metodología para la interpretación de señales de temperatura y tos**

#### **3.4.1. Sensor de temperatura**

#### **3.4.2. Estructura funcional del sensor de tos**

### **3.5. Aplicación móvil**

#### **3.5.1. Funcionamiento y análisis de datos**

#### **3.5.2. Diseño y nombre de la aplicación**

### **3.6. Software**

### **3.7. Protocolo de medida**

### **3.8. Análisis de datos**

### **3.9. Estadísticas**

#### **3.9.1. Validez de una prueba diagnóstica**

## **4. RESULTADOS**

### **4.1. Sensibilidad y Especificidad en pruebas diagnósticas de COVID-19**

### **4.2. Otras probabilidades de la presente prueba diagnóstica**

### **4.3. Diferenciación entre modelos**

## **5. DISCUSIÓN**

### **5.1. Limitaciones**

## **6. CONCLUSIONES**

## **7. APÉNDICES**

### **7.1. Apéndice A Programación del modelo 505**

### **7.2. Apéndice B Programación de la APP móvil**

### **7.3. Apéndice C Programación de tablas**

### **7.4. Apéndice D Programación de gráficas**

## **8. ANEXOS**

### **8.1. Anexo A. Carta de consentimiento informado y motivos.**

### **8.2. Anexo B. Banco de tos**

### **8.3. Anexo C. Portada de APP móvil**

### **8.4. Anexo D. Fondo de despedida APP móvil**

### **8.5. Anexo E. Logotipo NanNa (APP)**

## 9. BIBLIOGRAFÍA

## **Lista de abreviaturas.**

- 1 **a-Cov.** a-Coronavirus.
- 2 **ASTM.** American Society for Testing and Materials.
- 3 **b-Cov.** b-Coronavirus.
- 4 **BLE.** Bluetooth Low Energy.
- 5 **CONEVAL.** Consejo Nacional de Evaluación de la Política de Desarrollo Social.
- 6 **c-Cov.** c-Coronavirus.
- 7 **d-Cov.** d-Coronavirus.
- 8 **DFT.** Transformada Discreta de Fourier.
- 9 **E-Learning.** Learning electrónico o Learning digital.
- 10 **ER.** Error Relativo
- 11 **EX.** Exactitud
- 12 **E-Salud.** Salud electrónica o Salud digital.
- 13 **E.** Especificidad
- 14 **FN.** Falso Negativo
- 15 **FP.** Falso Positivo.
- 16 **FT.** Fragmentos de Tos.
- 17 **FFT.** Transformada Rápida de Fourier.
- 18 **IDE.** Integrated Development Environment.
- 19 **InDRE.** Instituto de diagnóstico y Referencia Epidemiológicos.
- 20 **I2C.** Circuito inter-integrado.
- 21 **LRHN.** Likelihood Ratio Negativo.
- 22 **LRHP.** Likelihood Ratio Positivo.
- 23 **ML.** Machine Learning.
- 24 **M-Salud.** Salud por dispositivos móviles.
- 25 **NanNa.** Nantzin Nahualli.
- 26 **OCDE.** Organización para la Cooperación y Desarrollo Económico.
- 27 **OMS.** Organización Mundial de la Salud.
- 28 **OPS.** Organización Panamericana de la Salud.



- 29 PROFECO.** Procuraduría Federal del Consumidor.
- 30 VSR.** Virus Respiratorio Sincitial.
- 31 RT-RCP.** Reacción en cadena de la polimerasa con transcriptasa inversa.
- 32 RCP.** Reacción en cadena de la polimerasa.
- 33 S.** Sensibilidad
- 34 UIDD.** Universally Unique Identifier.
- 35 VN.** Verdadero Negativo.
- 36 VP.** Verdadero Positivo.
- 37 VAPN.** Valor Predictivo Negativo.
- 38 VAPP.** Valor Predictivo Positivo
- 39 WHO.** World Health Organization.

## Lista de figuras

|   |    |
|---|----|
| Ilustración 1. Distribución de la aparición de síntomas para pacientes con COVID-19 confirmado mediante pruebas de laboratorio: casos importados (naranja), casos de contacto importados (gris), casos de contacto local (amarillo). La línea de la ilustración muestra los casos ..... | 12 |
| Ilustración 2 Estructura de una red neuronal. ....  | 13 |
| Ilustración 3 Componentes electrónicos utilizados .....   | 13 |
| Ilustración 4 Diseño del circuito impreso funcional para los tres modelos .....   | 13 |
| Ilustración 5 Distribución de FT del modelo 505 .....   | 13 |
| Ilustración 6. Histograma 1 .....   | 13 |
| Ilustración 7. Histograma 2 .....   | 13 |
| Ilustración 8. Histograma 3 .....   | 13 |
| Ilustración 9 Histograma 4. ....  | 13 |
| Ilustración 10. Histograma 5 .....  | 13 |
| Ilustración 11. Histograma 6.....   | 13 |
| Ilustración 12. Histograma 7. ....  | 13 |
| Ilustración 13. Histograma 8 .....  | 13 |
| Ilustración 14. Histograma 9. ....  | 13 |
| Ilustración 15. Histograma 10. ....   | 13 |
| Ilustración 16. Histograma 11.....  | 13 |
| Ilustración 17. Histograma 12 .....   | 13 |
| Ilustración 18. Histograma 13 .....   | 13 |
| Ilustración 19. Histograma 14 .....   | 13 |
| Ilustración 20. Histograma 15 .....   | 13 |
| Ilustración 21. Histograma 16 .....   | 13 |
| Ilustración 22. Histograma 17 .....   | 13 |
| Ilustración 23. Gráfica 18 Distribución helicoidal 1 .....  | 13 |
| Ilustración 24. Gráfica 19. Distribución helicoidal 2. ....   | 13 |

## Lista de tablas

|   |    |
|---|----|
| Tabla 1. Costo total del material. ....                           | 13 |
| Tabla 2. Tiempo total de datos ingresados a la red neuronal. .... | 13 |
| Tabla 3. Parámetros del espectrograma. ....                       | 13 |
| Tabla 4. Resultados de la matriz de confusión. ....               | 13 |
| Tabla 5. Resultados de S. y E. ....                               | 13 |
| Tabla 6. Probabilidades de las otras pruebas de calidad. ....     | 13 |
| Tabla 7. Rangos de LHR y su utilidad. ....                        | 13 |
| Tabla 8. Comparación de medias y desviaciones estándar. ....      | 13 |

# 1. INTRODUCCIÓN

## 1.1. Justificación.

En los últimos tres años la humanidad se ha visto afectada terriblemente por el azote de una nueva pandemia que ha dejado millones de muertes. La mayoría de los seres humanos habían creído que ninguna nueva enfermedad surgiría y tendría un efecto tan rapaz en su cotidianidad, sin embargo, una vez más como a lo largo de la historia se ha demostrado que todo tipo de eventos catastróficos se van a producir de forma regular. Ya sean los cambios climáticos para muchas especies extintas, meteoros cayendo sobre la tierra, erupciones volcánicas para los dinosaurios o múltiples enfermedades causadas por virus o bacterias que han ocasionado estragos en la población mundial.

Un ejemplo de estas últimas son las bacterias que ocasionaron la peste negra durante el siglo XIV, la cual es considerada la pandemia más devastadora de la historia. La bacteria responsable es identificada como *Yersinia pestis*, que según predicciones actuales mató a aproximadamente al 60% de la población europea y se extendió hasta Escandinavia y Moscú; Según registros, esta bacteria se ha presentado en tres ocasiones en la historia humana con otros nombres entre ellos destacan el de peste bubónica o plaga de Justiniano. Los primeros registros se presentaron durante el siglo VI y se extendieron hasta el siglo VIII; se cree que esta mató a un total de entre 30 y 50 millones de personas afectando mayormente a la zona circundante al mar Mediterráneo. El segundo registro fue el de la peste negra ya mencionada y por último la peste de Hong Kong causando aproximadamente 12 millones de muertes en todo el mundo excluyendo Australia y los polos.

La viruela también es considerada como una de las enfermedades más devastadoras de la antigüedad, es causada por el virus *variola* y se cree que se originó en la India o Egipto hace 3000 años; el rastro más antiguo de viruela corresponde al faraón egipcio Ramsés V (1157 a.c). Según historiadores la viruela se extendió a lo largo de las rutas comerciales de África, Asia y Europa, después con la llegada de los colonizadores europeos a las Américas aumentó su expansión

y con ella su letalidad en humanos, matando y debilitando también a la población indígena. Esta enfermedad era mayormente letal entre niños y lactantes.

La “gripe española” también requiere de mención; esta enfermedad es causada por un virus llamado *influenzavirus A subtipo H1N1* y se cree que mató a más de 40 millones de personas en todo el mundo durante los años 1918-1920.

Contrario a lo que se piensa en la actualidad y en el pasado más próximo, las pandemias no se erradicaron. El SIDA es un claro ejemplo de ello, esta enfermedad ha matado aproximadamente a un total de 25-35 millones de personas desde 1981 hasta la actualidad y en última instancia la enfermedad más estudiada de los últimos tiempos llamada COVID-19 ocasionada por un virus llamado *SARS-Cov-2*.

La COVID-19 ha generado un sin fin de estudios para su erradicación debido a que, según datos de la OMS, se han registrado 6,137,553 millones de muertes hasta el 1 de abril del 2022 en un periodo no mayor a tres años. Aún con los avances tecnológicos con los que cuenta la humanidad actualmente no se ha descubierto una cura completamente efectiva para la COVID-19, pero el avance médico si ha logrado desarrollar vacunas que reducen el riesgo de hospitalización y en consecuencia el riesgo de muertes; se han buscado métodos eficaces para la detección de la misma que sean de fácil acceso para la población en general. Pensando en ello surge la idea de crear por medio de medicina digital, un método de fácil detección de COVID-19 sin necesidad de usar pruebas de laboratorio

Durante los últimos años la tecnología como apoyo en distintas áreas, entre ellas la medicina va en aumento. Usar las herramientas disponibles para el desarrollo de un método de detección de COVID-19 surge como una idea ambiciosa pero posible gracias al desarrollo del aprendizaje automático o aprendizaje de máquinas, quienes en principio pueden ser capaces de proporcionar modelos de categorización o clasificación de datos de acuerdo con el entrenamiento proporcionado.

En términos de COVID-19 se pueden resaltar las principales sintomatologías medibles, como lo son la fiebre y los periodos de tos.

Según la hipótesis principal de este trabajo es posible diferenciar una muestra de tos de una persona con resfriado común o cualquier otro padecimiento de una persona con COVID-19; para ello debemos desarrollar una aplicación que base sus principios en la creación de redes neuronales que aprendan el espectro de sonido de toses con tres posibles clasificaciones.

- La primera: *Tos sin covid*.  
Se toman muestras aleatorias de tos de personas “*sanas*”<sup>1</sup>.
- La segunda: *Tos con covid*.  
Se toman muestras aleatorias de tos de personas que hayan sido diagnosticadas con COVID-19 siempre y cuando esto sea confirmado con una prueba de laboratorio.
- La tercera: *Ruido*.  
Se toman muestras aleatorias de distintos ruidos de fondo como voces, música, sonidos de autobuses, ladridos de perros, risas, llanto, etc.

Las redes neuronales deberán aprender el espectro de sonido<sup>2</sup> del fragmento de tos en cuestión y recabar las características necesarias para su clasificación, como variaciones de frecuencia e intensidad de sonido. O dicho de otro modo recabar las variaciones de los ciclos ondulatorios y la cantidad de energía emitida por la fuente sonora por unidad de tiempo en ambos.

Después se debe aplicar un conjunto de filtros necesarios para poder separar ambas muestras y ensayar al modelo para que sea capaz de identificar una tos y acomodarla en el espectro que mas se asemeje a sus características, donde los modelos sean distinguibles el uno del otro. Para el filtro de clasificación, el modelo se basará en un método de aprendizaje automático llamado *Classification (keras)*, cuyo objetivo es etiquetar objetos de consulta nunca antes vistos mientras distingue dos o más clases de destino.

- 1 A partir de ahora llamaremos personas “sanas” a pacientes que hayan presentado sintomatología asociada a la COVID-19 pero se hayan hecho pruebas de laboratorio descartando dicha enfermedad o hayan sido diagnosticados con algún otro padecimiento por algún médico.
- 2 Se le llama espectro de sonido a la representación gráfica de una señal en función del tiempo.

Además, se agregará un detector de temperatura que indicará si el paciente presenta fiebre; para complementar el análisis también se añadirá un mapa con registros de casos COVID-19 por región y de esta forma determinar si esta aplicación puede arrojarlos correctamente o no si una persona padece COVID-19.

El sensor de temperatura no debe estar íntimamente relacionado con las redes neuronales creadas para la clasificación de toses, debido a que este parámetro es completamente ajeno al área de sonido. Sin embargo, la temperatura es parte fundamental en la detección de COVID-19 por lo que se debe crear una forma de detectar temperaturas superiores a 38°C, ya que este síntoma es un fuerte indicador de que la persona padezca COVID-19 al menos en sus primeras variantes. Esto se puede lograr simplemente con un sensor de temperatura y un programa que indique cuando el usuario de la aplicación tenga una temperatura superior a la normal, es decir superior a 36° C. En conjunto el detector de tos y el monitoreo de la temperatura cubren dos de los más importantes y menos variados síntomas que se presentan en personas que padecen COVID-19 que se completará con el registro que podrá encontrarse en el mapa.

## 1.2. Hipótesis

**Las variaciones de sonido en toses con COVID-19 y en toses sin COVID-19 son distinguibles entre sí y, además, ambas son distinguibles de ruidos externos.**

Algunas aportaciones proporcionadas por Birring et.al. (2008) en su artículo titulado "*The Leicester Cough Monitor: preliminary validation of an automated cough detection system in chronic cough*", presentan, la validación de un monitor de tos ambulatorio automatizado basado en sonidos. Los autores han demostrado que existen claras diferencias en la frecuencia de tos entre pacientes con tos crónica y pacientes sanos; lo más destacable es que sugiere que la medición de la frecuencia de la tos es potencialmente útil en la evaluación de pacientes que padecen tos crónica.

Otros autores también sugieren que es posible las diferenciaciones entre tos y otros sonidos, entre ellos se encuentran Barry, Dane, Morice y Walmsley (2006) que en su artículo *“The automatic recognition and counting of cough”* plasman el desarrollo de un sistema automatizado capaz de analizar archivos de sonido que contienen tos y otros eventos distintos, con buena precisión y alta robustez al discernir toses reales en las grabaciones de audio. Este sistema fue creado a partir del uso de una red neuronal probabilística cuya capacidad comprende el cálculo de frecuencias de tos en función del tiempo y el número total de toses a partir de las grabaciones.

### **1.3. Objetivo principal.**

**El presente trabajo tiene como objetivo proporcionar una aplicación móvil basada en una red neuronal para la detección de COVID-19 como coadyuvante a las pruebas diagnósticas comerciales.**

### **1.4. Objetivos secundarios**

- Desarrollar una aplicación móvil capaz de detectar y monitorear COVID-19.
- Implementar una forma de que dicha aplicación sea lo menos costosa y pueda ser accesible para la población general, principalmente a las personas de bajos recursos.
- Refutar o incentivar la eficacia de dicha aplicación de acuerdo con los resultados experimentales obtenidos.
- Averiguar la sensibilidad y especificidad de un modelo de detección automática de sonido como prueba diagnóstica.
- Crear un antecedente de detección de sonido como prueba diagnóstica para COVID-19.



## **1.5. Resumen**

En la presente crisis sanitaria surgió la necesidad de creación de métodos de detección de COVID-19, la realidad de la mayoría de la población mexicana respecto a su capacidad de costear una prueba de detección de COVID-19 es desalentadora. Por ello se desarrolló un método de detección coadyuvante, de bajo costo y reutilizable, a partir de una aplicación que pudiera distinguir toses positivas a COVID-19. Complementariamente, el sistema debería ser capaz de monitorear la temperatura y tener acceso a un mapa de casos COVID-19 en México. Toda esta información se combinaría para producir un diagnóstico.

## **2. MARCO TEÓRICO**

### **2.1. Pandemias**

A lo largo del tiempo la humanidad ha enfrentado diferentes eventos virológicos que han puesto en riesgo su supervivencia; entre los más conocidos se encuentra la gripe española de 1918 que se propagó mundialmente infectando a más de un tercio de la población mundial y causando la muerte de aproximadamente 50 millones de personas, la voracidad y carácter masivo de este evento virológico hace referencia a un término designado como pandemia, es decir, pandemia se define como un brote infeccioso que se propaga a nivel mundial.

### **2.2. Protocolos de uso ante una pandemia**

La organización mundial de la salud (OMS o WHO por sus siglas en inglés) se encargó de dividir el desarrollo de una pandemia en 1999 y lo rectificó en 2005 y 2009 después de la pandemia de gripe H1N1, dando como resultado su división en seis fases que proporcionen un marco global para la preparación, planificación y respuesta mundial ante una pandemia.

Fase 1 Periodo prepandémico. En este periodo no se ha presentado algún virus que circule entre animales cause infección en humanos.

Fase 2 Fase de amenaza pandémica. En esta fase se ha presentado un virus animal (que circula en animales domésticos o salvajes) que causa infección en humanos.

Fase 3. Al llegar a esta fase un virus animal o humano-animal ha causado casos esporádicos en personas, pero no ha provocado transmisiones de persona a persona suficientes para sostener un brote a nivel comunitario.

Fase 4. En este punto se ha verificado la transmisión de persona a persona de un virus animal o humano-animal, que es además capaz de sostener brotes a nivel comunitario.

Fase 5. El virus identificado ha causado brotes sostenidos a nivel comunitario en dos o más países en una región de la OMS.

Fase 6. El mismo virus identificado ha causado brotes sostenidos al menos en otra región de la OMS.

Sin embargo, es de importancia resaltar otros tres periodos importantes.

1 Periodo posterior al pico. En este momento los niveles de actividad pandémica en la mayoría de los países con vigilancia adecuada han caído por debajo de los niveles máximos.

2 Posible nueva ola. Al llegar a este lapso temporal, el nivel de actividad pandémica en la mayoría de los países con una vigilancia adecuada comienza a aumentar nuevamente.

3 Periodo pospandémico. Finalmente, los niveles de actividad de la enfermedad vuelven a los niveles estacionarios en la mayoría de países.

En los últimos años la salud debe ir de la mano con la era digital debido a que el entorno se desarrolla de esta forma desde hace bastante tiempo atrás; gracias a la cantidad de información generada es posible analizar datos con nuevas tecnologías, crear opciones para el manejo de enfermedades y proporcionar soluciones sobre todo para casos de emergencia sanitaria como lo es una pandemia.

### **2.3. Salud digital**

La última pandemia que ha surgido recientemente en el mundo ha ocurrido en este contexto tecnológico, de modo que la forma de abordar el problema y dar una solución debe tener un enfoque del mismo tipo, surgiendo de este modo la Salud Digital (Digital Health). La Organización Mundial de la Salud en conjunto con la Organización Panamericana de la Salud (OPS) se dieron a la tarea de definir la Salud Digital o Salud electrónica (eSalud) como: *“La eSalud consiste en el apoyo que la utilización costo-eficaz y segura de las tecnologías de la información y las comunicaciones ofrece a la salud y a los ámbitos relacionados con ella, con inclusión de los servicios de atención de salud, la vigilancia y la documentación sanitaria, así como la educación, los conocimientos y las investigaciones en materia de salud”* (citado por Fernández S. M. 2013, p.42)

La salud digital comprende un amplio rango de opciones de uso, como el de tecnologías de información y comunicación para difundir o acceder a datos masivamente para la creación de algoritmos (Big Data y Machine Learning) , el uso de robots quirúrgicos, la mejora en la producción de imágenes de diagnóstico, terapias de realidad virtual, ciencias ómicas y en general cualquier tipo de herramienta tecnológica en pro de la salud, inclusive la difusión de infomerciales preventivos de fuentes confiables en redes sociales e internet. La salud digital ofrece innovación y avances tecnológicos y científicos para los países usuarios a nivel poblacional y organizacional, facilitando un buen control pandémico si es usado adecuadamente, además de ser un aliado en la mitigación de la misma como estrategia de apoyo.

Entre las técnicas que se usan hoy en día en términos de salud digital se encuentran:

1 Registro médico electrónico o historia clínica electrónica: en ella se registran los datos médicos de cada paciente con el objetivo de poder acceder a ella en cualquier momento. Además, con esta información es posible crear un registro pandémico ya que se puede saber si el paciente es portador o no de la patología en cuestión.

2 Bases de datos: gracias a la información contenida en internet se pueden crear bases de datos en línea y con ayuda de algoritmos predecir ciertos brotes o comportamientos de algunas enfermedades.

3 Telesalud: consiste en la prestación de servicios de salud a distancia por medio de un medio digital.

4 Mejora a equipos ya existentes: como su nombre bien lo indica se hace uso de nuevas tecnologías para el mejoramiento de equipos ya existentes por ejemplo para la fabricación de medicamentos, prótesis o incluso en la realización de cirugías.

5 Salud por dispositivos móviles (*mSalud*): es el ejercicio de la medicina y salud pública a partir de dispositivos móviles

como teléfonos celulares, tabletas, dispositivos de monitoreo de pacientes, como los liberadores de insulina, recordatorios de salud, relojes inteligentes, etcétera.

6 eLearning: consiste en el aprendizaje enfocado a la salud en blogs, conferencias, sitios web y otros sitios donde se creen estrategias de enseñanza médica, incluyendo las redes sociales.

7 Estandarización e interoperabilidad: la estandarización hace referencia a la conjunción entre las tecnologías y las aplicaciones de software siguiendo estándares de regularización, mientras que la interoperabilidad permite el uso de cualquier dispositivo o recurso disponible.

*El resultado de la salud digital recae por ejemplo en el uso de redes sociales para el entendimiento de como los patrones de contacto humano favorecen o no a la propagación de enfermedades como el VIH en una población. (Fernández S. M. 2013, 42.)*

#### **2.4. COVID-19**

En el último año se ha presentado una alerta sanitaria a nivel mundial causada por un agente infeccioso inicialmente desconocido, los primeros registros se presentaron en diciembre del 2019 en Wuhan China como casos de neumonía causada por un nuevo coronavirus inicialmente llamado 2019-novel coronavirus (2019-n-Cov), sin embargo, tiempo después fue designado con el nombre de síndrome respiratorio agudo severo Coronavirus 2 (SARS-Cov-2). El 12 de enero del 2020 la OMS nombró oficialmente a la enfermedad que causa el SARS-Cov-2 como enfermedad por coronavirus 2019 (COVID-19).

Los coronavirus se dividen en cuatro géneros:

- a-Coronavirus (a-Cov)
- b-Coronavirus (b-Cov)
- c-Coronavirus (c-Cov)
- d-Coronavirus (d-Cov)

Nota: para mejorar el entendimiento en el texto solamente se usarán los nombres cortos.

Los  $\alpha$  y  $\beta$  Cov pueden infectar mamíferos, mientras que los  $\gamma$  y  $\delta$  Cov suelen infectar a las aves. Dentro de los géneros  $\alpha$  y  $\beta$  Cov se han identificado sólo seis susceptibles a los humanos:  $\alpha$ -Cov sHCoV-229E y  $\alpha$ -Cov HCoV-NL63 (pertenecientes a los  $\alpha$ -Cov),  $\beta$ -CoVs HCoV-HKU1  $\beta$ - HCoV-OC43 (pertenecientes a los  $\beta$ -Cov) cuyo impacto es de baja patogenicidad y causan síntomas respiratorios de baja intensidad similares a los de un resfriado común *“aunque algunas de ellas pueden causar infecciones graves en bebés, niños pequeños y personas de edad avanzada”* según Cui J., Li F. y Shi Z-L. (2019, p.181)

Los dos coronavirus restantes son también pertenecientes a los  $\beta$ -Cov conocidos como Síndrome Respiratorio de Oriente Medio coronavirus (MERS-Cov) y el SARS-Cov-2 que en palabras de Guo Y-R. Et al. (2019, p. 4) *“cuya afección provoca grandes infecciones en el tracto respiratorio potencialmente mortales”*. Tras los primeros estudios se ubicó al coronavirus como un  $\beta$ -Cov y se descubrió que los murciélagos son reservorios naturales de una amplia variedad de Coronavirus incluyendo los similares al SARS-Cov-2 y al MERS, se secuenció un genoma del COVID-19 y se mostró un 96.2% de general identidad con un genoma del murciélago, de modo que se intuye que este puede ser el origen del COVID-19.

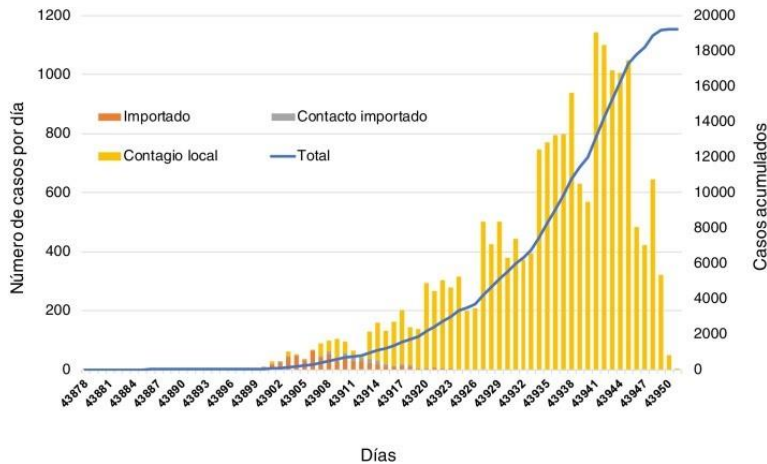
El primer caso confirmado de COVID-19 en México fue detectado el 27 de febrero del 2020 en la Ciudad de México, posteriormente el 11 de marzo la OMS declaró oficialmente al COVID-19 como una pandemia; solamente al 17 de Junio del 2020 se declararon a más de 8.2 millones de personas infectadas con variedad en la gravedad de sus síntomas, lo que dificulta la evaluación de su tasa de infección. Figura 1.

*Desde el primer caso el 27 de febrero de 2020, hasta el 30 de abril de 2020, el COVID-19 se ha extendido en todo el territorio mexicano. En ese periodo se llevaron a cabo 87,372 pruebas diagnósticas en todo el país. Resultando*

positivas, 19,224. El número total de defunciones en ese período de estudio fue de 1,859. (Suárez, et.al. 2020, pág. 470).

Los casos registrados en todo el mundo al 4 de septiembre del 2020 fueron más de 26,495,880 concluyendo con 873,618 muertes en este mismo periodo.

Ilustración 1. Distribución de la aparición de síntomas para pacientes con COVID-19 confirmado mediante pruebas de laboratorio: casos importados (naranja), casos de contacto importados (gris), casos de contagio local (amarillo). La línea de la ilustración muestra los casos



Nota: Adaptado de “Epidemiology of COVID-19 in Mexico: From the 27th of February to the 30th of April 2020” (pág.467), por V. Suárez, 2020, Revista Clínica Española ,220(8).

#### 2.4.1. Propagación del SARS-Cov-2

El COVID-19 se propaga principalmente a través del tracto respiratorio por gotas de saliva o secreciones, además se ha encontrado que el virus es capaz de transmitirse a través de superficies contaminadas o fómites y por contacto directo, “sin embargo se han encontrado también presencia de SARS-Cov-2 en hisopos fecales y sangre, lo que indica la posibilidad de transmisión de múltiples rutas”. (Guo, et.al. 2020, pág.4)

El SARS-COV-2 es extremadamente contagioso y especialmente peligroso en personas de la tercera edad, personas con enfermedades subyacentes o colectivos vulnerables (así como personas inmunocomprometidas: gestantes, personas discapacitadas, etc.), sin embargo, no deja de ser perjudicial y algunas veces letal para la población no catalogada como vulnerable; su periodo de incubación es de 1 a 14 días.

#### **2.4.2. Sintomatología del COVID-19**

En su mayoría los pacientes con COVID-19 presentan síntomas como fiebre, mialgia (fatiga), tos, dificultad para respirar, dolor de garganta, malestar general y cefalea; algunos casos menos frecuentes presentan esputo, diarrea y vómitos. Además, es posible encontrar personas asintomáticas lo cual representa un riesgo latente porque a pesar de no padecer síntomas esas personas si son capaces de transmitir el virus.

La gravedad e intensidad de los síntomas varía de persona en persona, sin embargo, según registros de Wang et.al. (2020) “los casos leves solo tienen fiebre baja y fatiga leve, sin neumonía. Los casos graves y moderados tenían manifestaciones clínicas de disnea, linfopenia e hipoalbuminemia, que se producían principalmente en pacientes de edad avanzada” (pág. 1631).

Es conveniente definir los siguientes términos:

La temperatura normal del cuerpo humano, según Guyton y Hall (2011)

*La temperatura de los tejidos profundos del organismo o temperatura central, permanece muy constante, con un margen de  $\pm 0.6$  °C, salvo cuando sobreviene una enfermedad febril. No existe una sola temperatura central que pueda considerarse normal, porque las mediciones efectuadas en muchas personas sanas revelan un intervalo normal de temperaturas bucales, desde 36 °C hasta más de 37 °C. En conclusión la temperatura central normal suele variar, por término medio, entre 36.5 °C y 37 °C si se mide en la boca y resulta 0.6 °C más alta si se mide en el recto. (pág.867).*



**La fiebre** es una respuesta cerebral a algún agente externo que provoca un foco infeccioso como las enfermedades bacterianas, tumores o incluso otro tipo de condiciones como las ambientales. Se considera que una persona tiene fiebre al alcanzar una temperatura de 38°C.

**La dificultad respiratoria** es el ineficaz intercambio gaseoso, la persona tiene una respiración incómoda, difícil y con sensación de no recibir el suficiente aire haciendo que los episodios de respiración sean más constantes y agotadores.

Finalmente, según Tortora y Anagnostakos (1981) **la tos** se define como *la inspiración prolongada y profunda, seguida por cierre total de la glotis lo que da por resultado espiración forzada que entreabre súbitamente la glotis y envía un chorro de aire por las vías respiratorias superiores. El estímulo para este acto reflejo puede ser un cuerpo extraño en la laringe, tráquea o epiglotis. (pág.730).*

Entonces, es notorio que el monitoreo de la temperatura y tos es crucial, por ello se requiere un sensor de temperatura sensible cuyo algoritmo de detección sea eficaz al detectar temperaturas altas para el cuerpo humano; existen varios sensores de temperatura, sin embargo se busca un sensor de temperatura que sea costeable para una buena cantidad de la población, que sea inalámbrico, discreto por cuestión estética y práctica y sobre todo cuya sensibilidad sea de un buen rango si se habla en términos de salud. El algoritmo del sensor debe indicar cuando se detecte una temperatura cuyo rango esté por encima del rango normal de la temperatura humana, es decir cuando el usuario registre temperaturas cercanas o incluso superiores a 38°C.

Hasta el momento no existen métodos digitales que ayuden a identificar enfermedades respiratorias.

Debido a que la tos puede representarse como una combinación de funciones de distribución de probabilidades, representada mediante algunos parámetros médicos, cuyo análisis indica que la dirección del chorro del episodio y el área de apertura de la boca no dependen de los parámetros fisiológicos de los humanos en

cuestión y combinando las características del flujo con información de distribución de gotas y virus es posible evaluar la fuerza de una fuente infecciosa al toser.

Por último, la ubicación geográfica; existen sistemas de información geográfica que proporcionan un mapeo en tiempo real o muy aproximado a él de enfermedades y su propagación, ya que se genera información necesaria para un mapeo predictivo de riesgos para la comprensión epidemiológica.

Es muy importante señalar el papel fundamental de la instrumentación científica enfocado a la medicina, ya que todo este tipo de avances no serían posibles sin ella.

### **2.4.3. Diagnóstico**

Un problema relevante en esta situación pandémica es la falta de recursos para el correcto diagnóstico del COVID-19 ya que los síntomas pueden ser muy parecidos a los que provocan otras enfermedades respiratorias como la influenza e incluso la gripe común, la alta demanda (y como consecuencia el desabasto) y la situación económica de cada paciente influye directamente en la búsqueda de un método diagnóstico efectivo.

Hasta el momento se dispone de dos tipos de técnicas de laboratorio para el diagnóstico de COVID-2019:

- 1 Diagnóstico agudo del virus SARS- CoV-2 en enfermedad aguda 3 a 7 días: la prueba de referencia es la técnica de reacción en cadena de polimerasa transcriptasa reversa (RT-PCR), que es una prueba de biología molecular en la que se detecta y amplifica una o varias regiones específicas del virus.*
- 2 Diagnóstico de formación de anticuerpos contra el virus SARS-CoV-2 después de 7 días: la prueba indicada es por medio de ELISA o inmunocromatografía para la detección de anticuerpos IgM e IgG para el virus SARS-CoV-2. (Díaz, 2020, pág. S-53.)*

El uso de RT-PCR para la detección de ARN viral solamente puede alcanzar una sensibilidad del 30 a 60% y su efectividad depende directamente de la condición del paciente, el número de muestras clínicas y el protocolo utilizado en cada muestra.

Hasta el momento se han implementado distintos equipos y plataformas moleculares para la identificación del SARS-Cov-2, entre ellas se encuentran:

- 1 Plataforma Gene Xpert con su prueba Xpert Xpress SARS-Cov-2, que se encarga de amplificar los genes de las proteínas E y la N2. Esta plataforma era utilizada para la detección de la tuberculosis y la resistencia a la rifampicina, la forma de acción del método consiste en la purificación, concentración, amplificación e identificación de las secuencias de los ácidos nucleicos específicos de la tuberculosis (mediante una prueba de RCP en tiempo real).
- 2 El equipo Film Array Biofire, agregó el Panel Respiratorio Biofire 2.1 cuyo trabajo consiste en detectar dos regiones genómicas del SARS-Cov-2 que permite además de la detección de este virus, la identificación de otros 19 virus respiratorios, lo que permite discernir entre una infección causada por influenza A o B frente al COVID-19.

En muchos casos los pacientes graves pueden ser diagnosticados mediante radiografías de tórax o tomografías computarizadas de tórax, sin embargo, en algunos casos las imágenes obtenidas a partir de pacientes con virus respiratorio sincitial (VSR), micoplasma, virus de la parainfluenza y COVID-19 comparten características según Wang et. Al. (2020) “lo que sugiere que las tomografías computarizadas de tórax no pueden identificar a los pacientes con COVID-19 y a los pacientes sin COVID-19 en algunos casos.” (pág.1633)

La mayoría de los sistemas de salud a nivel mundial experimentan tensiones bastante fuertes durante la pandemia, solo algunos pocos sistemas se han mantenido estables frente al SARS-Cov-2. Según Suárez et. Al. (2020).

*En México el número de test diagnósticos que se aplicó es de 0,69 pruebas por cada 1,000 habitantes, mientras que en China se estima que realizan tres pruebas por cada 1,000 habitantes. México ocupa el último lugar en el número de pruebas aplicadas para identificar la COVID- 19 entre los 36 países miembros de la Organización para la Cooperación y el Desarrollo Económico (OCDE), una limitación para poder identificar casos asintomáticos e incluso poder estimar la cantidad de infectados con precisión.* (pág.469)

Por tales motivos es de suma importancia aumentar el número de test diagnósticos en México.

#### **2.4.4. Inaccessibilidad de la población mexicana a las pruebas diagnósticas de SARS-Cov2**

La inversión que se ha destinado al diagnóstico del COVID-19 ha sido bastante significativa a nivel mundial, sin embargo, los métodos de detección son demasiado caros y muchas veces no son 100% seguros. Para poder dar un diagnóstico se han desarrollado distintos métodos de detección, pero estos no pueden ser ocupados por una gran cantidad de la población mexicana debido a sus altos costos; además no toda la población mexicana tiene acceso a zonas urbanas dónde se establecen los laboratorios clínicos en que se realizan las pruebas.

Según datos del CONEVAL basados en la información del INEGI, hasta febrero del 2022 las líneas de pobreza extrema por ingresos <sup>3</sup> y las líneas de pobreza por ingresos <sup>4</sup> fueron de \$1,496.32 en zonas rurales y de \$1950.26 en urbanas además de \$2,837.57 en las zonas rurales y de \$3,997.24 en las zonas respectivamente para cada línea. (Consejo Nacional de Evaluación de la Política de Desarrollo Social [CONEVAL], 5 de abril del 2022). También para el CONEVAL una familia conformada por cuatro personas se encuentra en una situación de pobreza si su ingreso mensual es menor de \$11,290.80 es decir que se puede destinar

3 La Línea de Pobreza Extrema por Ingresos hace referencia a los gastos invertidos en canasta alimentaria.

4 La Línea de Pobreza por Ingresos hace referencia a los gastos invertidos en canasta alimentaria más no alimentaria.

\$2822.70 para cada integrante, también es importante considerar que el salario mínimo en México es de \$2401.

Los niveles de pobreza en los Estados Unidos Mexicanos son muy altos, según (CONEVAL, 2022) a nivel país el porcentaje de los niveles de pobreza de la sociedad mexicana en el año 2020 fueron de la siguiente manera:

- Pobreza (P): 43.9 %
- *Pobreza Extrema (PE)*: 8.5 %

de acuerdo a las estadísticas los cambios porcentuales en los 4 años anteriores a este año respecto a estas cifras no son mayores a 1.3 % en PE y 0.7% en P; el total de la población en México según el (Instituto Nacional de Estadística y Geografía [INEGI], 5 de abril del 2022) hasta el 2020 fue de 126,014,024 personas, es decir que 55,320,156.54 personas son pobres en México, estas personas ganaron entre \$2401 a \$2822.70 aproximadamente.

Según (Procuraduría Federal del Consumidor [PROFECO], 5 de abril del 2022) el costo de las pruebas *PCR* está entre \$950 a \$3,299 si el paciente se traslada al laboratorio a hacerse la prueba, si no le es posible entonces hay servicio de toma de muestras a domicilio en algunos laboratorios y el precio por este servicio va desde los \$0 hasta los \$400 y el tiempo de entrega de la prueba es de 24 -72 horas. Por otro lado, la prueba de *Antígenos* tiene un costo de entre \$250 a \$999 y si el laboratorio tiene toma de muestras a domicilio el costo se incrementa \$150 extras y el tiempo de entrega es de 30 minutos - 3 horas.

Considerando los datos anteriores una persona que vive en una zona rural y en condiciones de pobreza extrema, tiene un salario mensual aproximado de \$2401 (salario mínimo) si esta persona invierte \$1496.32 pesos en su canasta alimenticia básica le queda un total de \$904.68 suponiendo que no hace ningún otro gasto más y se enferma, va al médico e invierte \$60 y este le sugiere hacerse una prueba para descartar COVID-19 pero en su comunidad no hay laboratorios ( debido a que en zonas marginadas es poco probable encontrar muchos servicios básicos de salud) entonces debe trasladarse hasta el centro urbano más cercano a su comunidad e invierte \$100 más en su viaje de ida y

regreso, también considerando que debe regresar días después a recoger sus resultados se gastaría \$100 extras.

Esta persona se queda con una cantidad total de \$644.68 después de gastos básicos, si el médico le recomendó hacerse la prueba de PCR ni siquiera tiene la oportunidad de poder pagar la más barata en el mercado según los precios de PROFECO, si el paciente está inmovilizado y no puede trasladarse a otro lugar y pide que le tomen la muestra a domicilio consideremos que le cobran un precio de \$200 (el promedio del costo más alto de este servicio) el costo de la prueba más barata le excede un total de \$245.32, es claro que el paciente no tiene la oportunidad de realizarse esta prueba, sin considerar las citas subsecuentes al médico y el gasto en medicamentos. También es probable que si el paciente se realiza una prueba distinta a la que la orden médica indica es muy probable que su resultado sea erróneo.

Por otro lado, si el médico recomienda que se haga una prueba de antígenos y el paciente se hace la prueba más barata es decir la de \$250 más los \$100 de transporte, su ingreso restante mensual se reduciría a \$554.68 para los gastos necesarios que tenga que hacer el resto del mes.

Para una persona que vive en una zona urbana que se encuentra dentro de la categoría de pobreza extrema, es decir que gana poco más del salario mínimo o sea aproximadamente \$2822.70 al mes la situación planteada con anterioridad es similar, usaría \$1950.26 en su canasta básica y le resta un total de \$872.44, menos el gasto de traslado al lugar de las pruebas que consideramos se reduce en \$50 (ida) y \$50 (regreso) por la “facilidad” de movimiento y transporte en las zonas urbanas, menos \$50 de la consulta médica, tendría un total final de \$722.44 nuevamente ni siquiera tiene la oportunidad de pagar la prueba de PCR más barata del mercado, el costo se excede por \$227.56, ahora si la prueba es a domicilio tendría un excedente de \$327.56.

En el caso de una prueba de antígenos sus ingresos totales mensuales se reducirían a \$472.44 considerando el costo de la prueba más barata, los gastos en transporte y consulta médica. Es claro que gran parte de la población

mexicana no es capaz de asimilar los gastos en pruebas diagnósticas de SARS-Cov2, en cifras específicas 55,320156.54 de la población en México (considerando que todas las personas contenidas en esta cifra se encuentran económicamente activas) no es capaz de costear pruebas de detección de SARS-Cov2 y mucho menos para más de un integrante en una familia, lo que representa un problema de salud.

## **2.5. Interpretación de los sonidos.**

El sonido puede definirse como un conjunto de ondas mecánicas de presión longitudinal que viajan por un medio y producen una perturbación audible.

La interpretación de los sonidos como medio de detección de enfermedades surgió hace aproximadamente 20 años, sin embargo, en los últimos años estas técnicas se han ido acompañando de Inteligencia Artificial (IA).

Generalmente, en acústica se trabaja con tres variables:

- El tiempo (t)
- La frecuencia ( $\nu$ )
- La intensidad (I) o nivel sonoro ( $N.S.=10\log I/I_0$ )

Esto sugiere que una representación gráfica tridimensional en la que un sonido aparece como superficie tridimensional llamada “objeto sonoro”. (Gérard, 2011, pág. 37)

Según Cortes O. J. A. et al. (2012) “la intensidad viene definida por la potencia, que a su vez está determinada por la amplitud de la onda sonora, permitiendo distinguir si el sonido es fuerte o débil. Esta cualidad se mide con el sonómetro.”

Dentro de las señales acústicas podemos encontrar ondas sinusoidales las cuales poseen algunas propiedades, entre ellas se encuentran las siguientes:

- Si dos señales sinusoidales tienen la misma frecuencia la suma es otra sinusoidal de la misma frecuencia, aunque cambia la amplitud.
- Si las sinusoides que se suman no poseen la misma frecuencia, la señal resultante ya no será sinusoidal y hasta puede no ser periódica
- Si se superponen señales sinusoidales de frecuencias cercanas, aparecen variaciones en la amplitud llamadas pulsaciones.

Las ondas sinusoidales puras son llamados armónicos, que al mezclarse o combinarse pueden regresar al sonido original y sus frecuencias son múltiplos del sonido fundamental.

El teorema de Fourier establece que toda función periódica de frecuencia  $\omega_0$  pueden descomponerse en una suma de funciones seno o coseno, que son múltiplos enteros de  $\omega_0$ , donde  $\omega_0$  es la frecuencia fundamental y a cada término de la función en particular, ya sean senos o cosenos, se le conoce como armónica.

Así mismo la transformada discreta de Fourier (DFT) nos dice que es posible descomponer una señal periódica en armónicos de frecuencias y amplitudes distintas, “en las aplicaciones de ingeniería y tratamiento de señales resulta más práctico considerar el proceso de manera discreta y no continua, ya que los sistemas de adquisición de datos no pueden obtener ni analizar la totalidad de la información” Cortes et al. (2012)

Para este tipo de análisis es posible usar también la transformada rápida de Fourier (FFT) que “realiza los mismos cálculos que la DFT, pero de manera mucho más rápida gracias a su recursividad.” Cortes et al. (2012).

### **2.5.1. Sonograma o espectrograma.**

Las interpretaciones de los objetos sonoros se dividen en tres planos distintos: 1. Plano dinámico, 2. Plano espectral y 3. Plano melódico; los cuales hacen referencia al nivel sonoro en función del tiempo para el primer caso, el nivel sonoro en función de la frecuencia para el segundo caso y finalmente la frecuencia en función del tiempo para el tercero.



Los planos melódicos o diagramas de frecuencia en función de tiempo que realiza un análisis de Fourier de cada sonido reciben el nombre de sonogramas o espectrogramas, (Leipp, 1984 así como cita Gérard, 2011) “Esta gráfica es un documento en el que es posible medir con precisión la frecuencia y la duración de cada nota, así como la evolución dinámica.”

### **2.5.2. Calibración.**

Para poder realizar un análisis correcto es de suma importancia contar con los parámetros necesarios para que este sea el correcto.

Una de las principales consideraciones es la configuración de las escalas (en todas las gráficas) tanto del eje horizontal como del vertical. Muchas veces los autores eligen escalas con rangos muy cortos dando lugar a que sólo se analiza una parte muy reducida del sonido, a veces tan pequeña que el análisis no muestra casi nada. (Gérard, 2011, pág. 46)

Según Gérard (2011), para realizar el análisis inicialmente se elige el rango de frecuencias a la que se trabajará de acuerdo a las necesidades del proyecto en cuestión, si se elige por ejemplo un rango de 0-2 Khz es posible que se pierda demasiada información, en consecuencia, conforme el rango de frecuencias sea mayor la pérdida de información será menor. En la resolución de sonograma en específico para el análisis FFT existen distintas resoluciones y dependiendo del programa de análisis se encontrarán una amplia gama, sin embargo, los principales son llamados banda ancha o *wide band* y banda delgada o *narrowband*. Cuando lo que se requiere es precisión en las frecuencias de banda delgada se pierde precisión en los niveles sonoros, la diferencia entre bandas depende básicamente del filtro aplicado y el número de puntos FFT; por lo tanto un filtro ancho es menos preciso en frecuencia y en consecuencia no muestra todos los detalles de su variación, solamente las generales pero si muestra variaciones de amplitud y el delgado proporciona mejor detalle de las frecuencias pero se pierde detalle en la

amplitud. También existe relación entre el número de puntos FFT y el ancho del filtro elegido, pues a menor ancho del filtro se requieren más puntos y viceversa.

“Los anchos de filtro estándares son: 45, 59, 150, 184, 300, 450, 600 Hz y los números de puntos de FFT son: 64, 128, 256, 512, 1024, 2048, 4096, 8192 y 16384”.

## **2.6. Aprendizaje Automático.**

El uso de la inteligencia artificial en el ámbito médico se ha usado desde los inicios de esta, proporcionando así un amplio aporte a la humanidad. En la actualidad la relación intrínseca entre la disciplina médica y la informática van de la mano al punto que es impensable imaginar a la ciencia médica independiente a la informática.

El deep learning junto a los especialistas del sector pueden hacer del diagnóstico médico una tarea más rápida y precisa (inmensas posibilidades para mejorar los diagnósticos, la creación de vías de atención y la reproducibilidad en los procedimientos quirúrgicos para, en última instancia, lograr mejores resultados clínicos. Arias et. al. (2019)

(Goodfellow, et. al. 2015, como se citó en Lee & Song, 2019) El deep learning es un enfoque combinado del aprendizaje automático que se basa en el conocimiento del cerebro humano, las estadísticas y las matemáticas aplicadas.

“El software aprende, en un sentido muy real, a reconocer patrones en las representaciones digitales de sonidos, imágenes y otros datos” Arias et. al. (2019).

### **2.6.1. Procedimiento de clasificación.**

Inicialmente se debe considerar que el deep learning basa su estructura en una red neuronal. “Una red neuronal se puede expresar como un conjunto de pesos (y sesgo) donde cada peso y cada sesgo se utilizan para el cálculo entre capas. El

conjunto de parámetros de una red se compone de todos los pesos y sesgos de un modelo.” Lee & Song (2019)

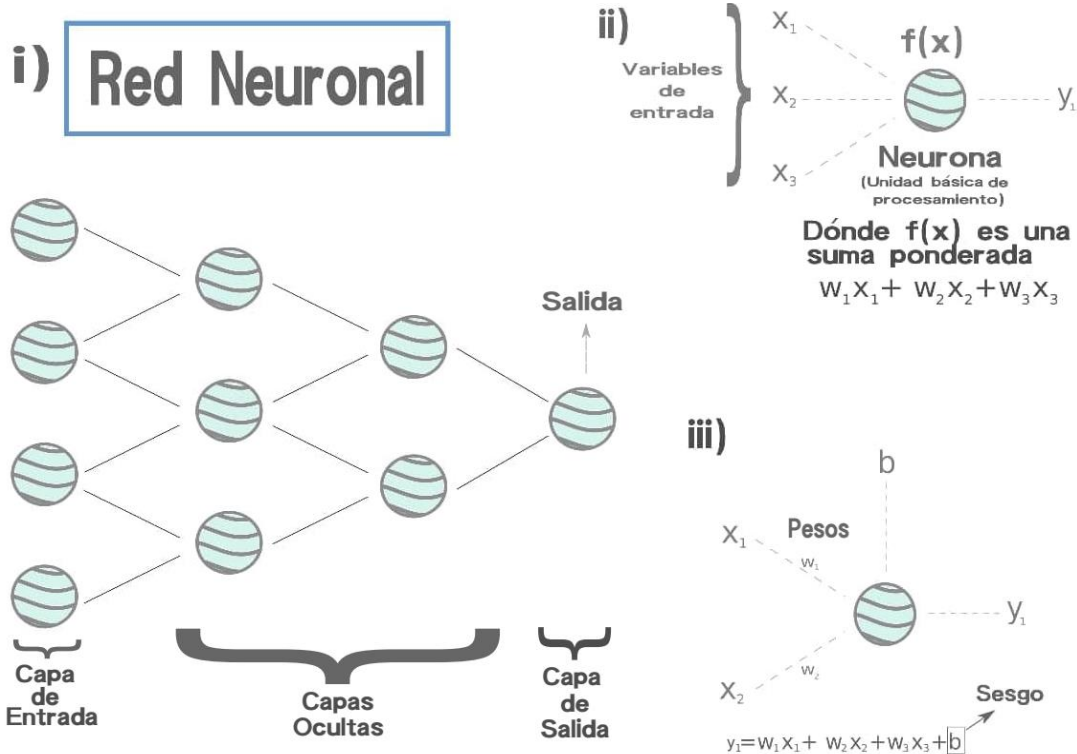
Es decir, que en el deep learning, una red neuronal está compuesta por neuronas, las que se encuentran en la capa de entrada corresponden a los datos de una observación y se otorga un peso en cada entrada lo que permite realizar una fase de propagación usando una función de activación.

Primero se deben proporcionar los datos característicos de entrada y los valores que deben llegar a la salida, el modelo aprende a extraer las características que pedimos y nos arroja un resultado más cercano posible.

La estructura de una red neuronal es de la siguiente manera:

En i) se observa la estructura de una red neuronal dónde podemos observar que en la primera capa de izquierda a derecha hay un mayor número de neuronas, esta capa es designada como capa de entrada y en ella ingresa la primera información adquirida, subsecuentemente la información viaja a las dos siguientes capas llamadas capas ocultas y finalmente la información más filtrada pasa a la capa de salida la cual arroja el modelo resultante; en la sección ii) se aprecia una neurona que forma parte de la red neuronal, la neurona es la unidad básica de procesamiento y cada neurona tiene conexiones de entrada por donde ingresan valores de entrada ( $x_1, x_2, x_3, \dots$ ) o datos que provienen del datasheet que queremos procesar.

Ilustración 2 Estructura de una red neuronal.



Nota: Imagen basada en distintos artículos científicos sobre deep learning, entre ellos Wang & Gao. (2015), Arias et. al. (2019), Sarmiento (2020).

Ya dentro se efectúa una función  $f(x)$  sobre los datos, donde  $f(x)$  es una suma ponderada de cada valor entrada multiplicado por un peso ( $w_1, w_2, w_3, \dots$ ) estos pesos serán los parámetros asociados a cada valor de entrada cuyo objetivo es modificar el tamaño de cambio que tendrá cada dato ( $x_1 w_1 + x_2 w_2 + x_3 w_3 + \dots$ ) y de esta forma generar un valor de salida  $y_1$ , sin embargo, aún no se cumple el objetivo, el cual es filtrar la información recibida en esa neurona para poder así clasificar la información útil que pasará a la siguiente neurona, de tal forma que se le agrega otra conexión a la neurona la cual es denominada  $b$  y su valor siempre es uno. El valor de  $b$  se modifica controlando el valor del parámetro de sesgo. Finalmente, el valor resultante  $y_1 = x_1 w_1 + x_2 w_2 + x_3 w_3 + \dots + b$  se compara con un umbral y entonces si  $y_1$  es mayor que ese umbral se le asigna una clasificación u otra.

Además cuando este proceso ha ocurrido cada neurona es capaz de “recordar” los pesos de ponderación (se crea un algoritmo), que permite que cuando otro conjunto de datos se procese su salida sea comparada con la salida primaria de modo que gracias a la comparación sean actualizados los pesos para que la siguiente vez que se vea los mismos valores de entrada, la salida se acerque lo más posible a la salida primaria, este proceso es iterativo hasta que los pesos se ajusten de tal forma que el modelo sea capaz de distinguir un tipo de valores de entrada y predecir su clasificación.

El método de trabajo de una red neuronal se basa en la regresión lineal; La organización neuronal por capas de comunicación repetitivas hace que la estructura se simplifique hasta llegar a una forma primaria de una neurona debido a que se concatenan distintas operaciones de regresión lineal, en embargo eso es un problema porque sumar varias operaciones de regresión lineal se puede resumir como hacer una sola, ya que la estructura neuronal compuesta por  $n$  neuronas se simplificaría a tener solamente una neurona; esto se resuelve usando funciones de activación las cuales generan manipulaciones no lineales que las modifica.

### 3. MATERIALES Y MÉTODOS.

Para realizar este proyecto se utilizaron los siguientes materiales:

- Sensor de temperatura MLX90614ESF-DCC (1)
- Placa de desarrollo Arduino Nano 33 BLE Sense (1)
- Switch Push Buton de 4 Pines (2)
- Resistencia de 10k  $\Omega$  (2)
- Resistencia de 4.7k  $\Omega$  (2)
- Protoboard 830 puntos (1)
- Alambre estañado (1)
- Cables Dupont 20cm (H-H), (M-M) (M-H) (15)
- Cautín (1)
- Pasta para soldar (1)
- Soldadura de estaño (1)
- Cable USB macho A mini USB (1)
- Placa de cobre para circuitos impresos de 50x50 mm (1)

#### 3.1. Características y costos de los materiales.

- **Sensor de temperatura MLX90614ESF-DCC.**

Este es un sensor infrarrojo que permite medir la temperatura de un objeto sin tener contacto con este, es fabricado por la empresa Melexis y tiene un costo aproximado de \$590. La gama de sensores de modelo MLX90614 poseen un chip de silicio con una membrana hecha para poder detectar la radiación infrarroja emitida por un objeto a distancia.

Se eligió este modelo de sensor por varias razones, entre ellas porque es un sensor de grado médico, de acuerdo con el datasheet del sensor este cumple con las normas de la American Society for Testing and Materials, ASTM por sus siglas en inglés o Sociedad Americana para Pruebas y Materiales (2003) en la sección 5.4 la cual indica textualmente que:

“Maximum Permissible Laboratory Error (for a Skin IR Thermometer):

Within the manufacturer's specified operating ambient conditions (see 5.6) over the display temperature range as specified in 5.2.2:

A skin IR thermometer shall display a subject's temperature over a minimum range of 22 to 40.0 °C (71.6 to 104.0 °F), laboratory error  $\delta$  as measured according to 6.1.5 shall be no greater than 0.3 °C (0.5 °F)".

Traducido al español:

[Error de laboratorio máximo permitido (para un termómetro IR de piel):

Dentro de las especificaciones de operación ambientales del fabricante (ver 5.6) sobre el rango de temperaturas visualizadas como se especifica en 5.2.2:

Un termómetro IR de piel no debe mostrar la temperatura del sujeto alrededor del rango mínimo de 22 a 40 °C (71.6 a 104 °F), la medida del error  $\delta$  del laboratorio de acuerdo a 6.1.5 no debe ser mayor que 0.3 °C (0.5 °F)] (p. 3)

Otra característica deseable es su disponibilidad en México, debido a que muchos componentes electrónicos con características similares solamente pueden ser encontrados principalmente en algunas regiones de Estados Unidos de América, Europa y Asia; por esta razón el costo se eleva considerablemente. Por ejemplo, para poder comprar un sensor de temperatura digital MAX30208 en la tienda Worldway Electronics (distribuidora de componentes electrónicos), se debe hacer una compra mínima de 100 USD más el costo en dólares del componente lo que se vuelve no viable debido a que uno de los objetivos es usar la menor inversión posible.

La última ventaja destacable de este sensor es que tiene un modo de programación amigable para el desarrollador, ya que se puede programar desde el entorno de desarrollo de Arduino y no requiere un arreglo electrónico complejo.

- **Arduino Nano 33 BLE Sense.**

El Arduino Nano 33 BLE Sense es una placa de desarrollo basada en el procesador nRF 52840. Es capaz de grabar instrucciones desde un entorno de

desarrollo amigable llamado IDE (Integrated Development Environment) de Arduino. Esta herramienta permite desarrollar y grabar el código programado de una manera sencilla. Arduino posee una interfaz de entrada y salida que se conecta a distintos periféricos y recaba la información que estos generan para el primer caso y análogamente distribuye información a los periféricos para el segundo caso; se escogió la placa Arduino Nano 33 Ble Sense debido a que a pesar de que no es la más económica en el mercado el costo-beneficio que ofrece es mejor que el que otras placas ofrecen, debido a que tiene incluido un sensor de audio o micrófono. Este es el micrófono digital MP34DT05 necesario para grabar las toses, necesita un voltaje de operación de 3.3 V menor a los 5-12 V que requieren otros arduinos; tiene un diseño pequeño y práctico, contrario al Arduino Uno, por ejemplo, aunque no es la más pequeña como la mini, pero la Sense requiere menor voltaje de operación que esta. El precio aproximado del Arduino Nano 33 Ble Sense es de \$1835.

También incluye Bluetooth y se puede alimentar con un cargador de celular de entrada A mini simplificando así el arreglo necesario para poder hacer nuestro proyecto, en caso de usar otra placa arduino también es posible hacer el arreglo, sin embargo, se tendrían que comprar los componentes independientes a la placa, el montaje sería más estorboso y muy probablemente el costo sería muy similar. Por ejemplo, un micrófono de las características que incluye el Nano 33 Ble Sense no es muy comercializado en México, habría que gastar en exportación, buscar existencias, compatibilidad con la placa, etc.

No fue posible usar el sensor de temperatura que incluía el arduino debido a que la exactitud es de  $\pm 5^{\circ}C$  y como se menciona con anterioridad se requiere una precisión menor a  $\pm 3^{\circ}C$ . Otros beneficios que ofrece una placa Arduino es que su hardware y software son libres y en particular la placa Nano 33 Ble Sense tiene compatibilidad con Edge Impulse, una plataforma de desarrollo para aprendizaje automático.



- **Costos totales.**

*Tabla 1. Costo total del material.*

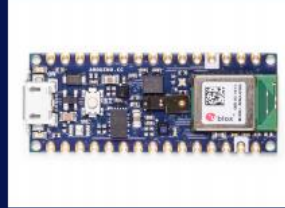
| Costo por material.        |                        |                  |                   |
|----------------------------|------------------------|------------------|-------------------|
| Material                   | Precio (\$) por unidad | Numero de piezas | Precio (\$) total |
| Sensor MLX90614ESF-DCC     | 590                    | 1                | 590               |
| Arduino Nano 33 BLE Sense  | 1835                   | 1                | 1835              |
| Switch Push Buton          | 1.50                   | 2                | 3                 |
| Resistencia 10k $\Omega$   | 1                      | 2                | 2                 |
| Resistencia 4.7k $\Omega$  | 1                      | 2                | 2                 |
| Protoboard                 | 55                     | 1                | 55                |
| Alambre estañado           | \$6/m.                 | 2 m.             | 12                |
| Cables Dupont              | 2                      | 6                | 12                |
| Cautín (kit)               | 150                    | 1                | 150               |
| Pasta para soldar (kit)    | (kit)                  | 1                | (kit)             |
| Soldadura de estaño (kit)  | (kit)                  | 1                | (kit)             |
| Cable USB macho A mini USB | 50                     | 1                | 50                |
| Placa de cobre             | 10                     | 1                | 10                |
| <b>Precio Total</b>        |                        |                  | <b>2721.00</b>    |

El costo total del material utilizado fue de \$2721.00 considerando que se compraran todos los materiales, en el caso del reciente proyecto la inversión fue menor debido a que se usaron materiales con los que ya se contaba por ejemplo el kit: cautín, soldadura de estaño y pasta para soldar, el cable USB macho A mini USB o cargador de entrada A mini para celular (cualquier marca), es decir la inversión total fue de: \$2521.00.

*Ilustración 3 Componentes electrónicos utilizados*



**Sensor  
MLX90614ESF-DCC**



**Arduino Nano  
33 BLE Sense**



**Switch Push  
Buton**



**Resistencia  
10 K**



**Resistencia  
4.7K**



**Protoboard**



**Alambre  
Estañado**



**Cables  
Dupont**



**Cautín**



**Pasta y estaño  
para soldar**



**Cable USB macho  
A mini USB**

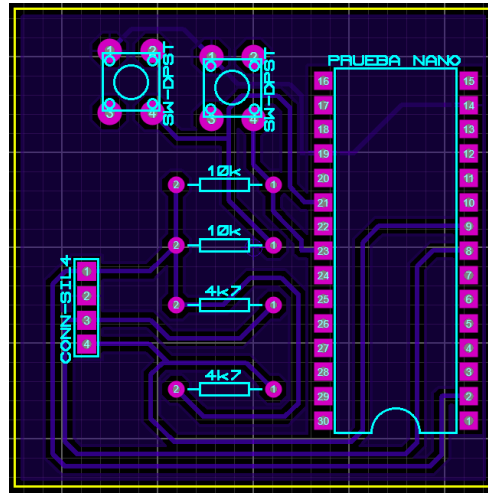


**Placa de  
cobre**

### 3.2. Arreglo experimental

Se diseñó el circuito impreso en el simulador Proteus del arreglo electrónico montado previamente en la protoboard ilustración 4.

*Ilustración 4 Diseño del circuito impreso funcional para los tres modelos*



- Se buscó que la placa fuera lo más pequeña posible, en este caso 50x50mm para que el diseño fuera cómodo y lo menos estorboso posible.
- Se hizo uso de botones físicos para las fases de prueba y para simplificar el trabajo al hacer la conexión con la aplicación móvil. Las resistencias de 10k Ohms están soldadas a los push buton actúan como Pull-Up, es decir mantiene un estado alto en el pin cuando el pulsador está en reposo y en bajo cuando el pulsador es presionado.
- El sensor de temperatura requería dos resistencias de 4.7k Ohms debido a que ese valor requiere la comunicación I2C que es el tipo de comunicación con la que trabaja el sensor para mantener en estado alto los pines en el Arduino.
- Todos ellos conectados a diferentes pines del Arduino en el que se cargaron los modelos.
- La alimentación y tierra van a los pines 23 y 14 respectivamente.

- La alimentación es de 3.3 volts y puede ser alimentado con un cargador de teléfono de entrada A mini USB macho.

### **3.3. Bases de datos de tos.**

Se buscaron bases de datos de tos de personas que habían sido diagnosticadas con COVID-19 y se consiguió una de uso libre que fue creada con fines similares al de la presente tesis, su objetivo también es crear una prueba de Covid-19 a partir de patrones de tos.

El porqué de su base de datos abierta se debe a que ellos consideran que la sociedad en conjunto debe trabajar para superar la pandemia de COVID-19 ya que es en extremo perjudicial para todos, además de que creen que dicha base estimulará la innovación y colaboración de la academia e industria en todo el mundo. (Virufy AI Research Group. [Virufy], 2020). Sus datos clínicos son confiables, ya que fueron recabados en un hospital bajo supervisión médica que seguía los procedimientos operativos estándar y con consentimiento informado del paciente; además de que tomaron los datos demográficos del paciente; el resultado final de toses fue de 16 pacientes, su clasificación se divide en Covid positivos y Covid negativos, además de que proporcionan dos carpetas de datos, la primera con los audios completos y la segunda con audios fragmentados. Virufy-data. (2020).

Debido a que el muestreo total obtenido de la base de datos de Virufy era considerablemente pequeño, para la realización del presente proyecto se creó un “Banco de tos” virtual exclusivamente para completar la base de datos obtenida previamente.

Este banco de datos se pensó regido bajo un método exclusivo para la toma de muestras de tos, dicho método fue revisado por personal médico (un médico y una enfermera) y además cuenta con el consentimiento informado de los pacientes

que participaron, el método es sencillo y se describe en el documento de permiso informado, este se adjuntará en la sección de **Anexos**.

Para crear el *Banco de tos* se pensó en un método de toma de muestra para que dicho ejercicio fuera realizado de la mejor manera y el audio fuera útil para el análisis.

Fue necesario crear dicho método debido a que en los “Lineamientos para la toma, manejo y envío de muestras para diagnóstico a la red nacional de laboratorios de salud pública” del Instituto de Diagnóstico y Referencia Epidemiológicos “Dr. Manuel Martínez Báez” perteneciente a la Subsecretaría de Prevención y Promoción de la Salud, unidad Dirección General de Epidemiología perteneciente a la Secretaría de Salud no se encontró lineamiento alguno que rigiera la naturaleza de las tomas de muestras necesarias para hacer la base de datos. (Instituto de Diagnóstico y Referencia Epidemiológicos. [InDRE], 2020). En otras palabras, no existía un método de toma de muestra de tos.

Se registró la historia clínica<sup>5</sup> de los pacientes que contribuyeron en la donación de toses para el *Banco de tos*, como información complementaria. La edad de los participantes en la base de datos de Virufy se encontraba en un rango de los 18 a los 65 años, por tal motivo el *Banco de tos* admitió participantes en ese mismo rango de edad, las toses de los pacientes se clasificaban como *toses positivas* y *negativas*, si se consideraban COVID positivos esto debía ser corroborado por una prueba diagnóstica (preferentemente por la prueba PCR).

De la base de Virufy se consiguieron un total de 16 audios de tos de duración variable, de los cuales 9 eran de personas sanas y 7 de personas diagnosticadas con COVID-19, el *Banco de tos* consiguió un total de 9 audios de tos de duración variable, de los cuales 3 eran de personas sanas y 6 de personas COVID-19 positivas, en conjunto se obtuvo una cantidad total de 25 toses de las cuales 12 eran “COVID negativas” y 13 “COVID positivas”, al terminar de reunir todas las toses se

<sup>5</sup> Edad, sexo biológico, padecimientos previos, etc.

recortaron en fragmentos de tos (FT)<sup>6</sup> útiles de 3 segundos en *Audacity* (editor de audio) dónde solamente se detectaban toses con mayor nitidez y se evitan las combinaciones de audios de toses con voces, silencios o sonidos que perturbaban la tos, sin embargo no se pudieron eliminar del todo los ruidos externos.

### **3.4. Metodología para la interpretación de las señales de temperatura y tos**

#### **3.4.1. Sensor de temperatura.**

Para la toma de temperatura se tomó en cuenta la consideración de que una persona tiene fiebre cuando su temperatura corporal rebasa los 38°C.

El sensor de temperatura infrarrojo MLX90614ESF-DCC detecta temperaturas del cuerpo en cuestión al que se le apunte, es decir, detecta la radiación emitida por este cuerpo.

El voltaje de alimentación del sensor es de 3 a 5 V (esta viene indicada en el datasheet), para que el sensor (conocido como esclavo) sea capaz de hacer una lectura de temperatura necesita que otro dispositivo conocido como “maestro” (en nuestro caso el arduino Nano 33 BLE Sense) le dé la indicación de que debe hacer una lectura de datos, esto se conoce como protocolo de comunicación; entonces nuestro sistema de comunicación está basado en el protocolo I2C.

La creación del código dónde se indican las acciones que deben tomar tanto el esclavo como el maestro se realizó en la IDE de arduino y su estructura es la siguiente:

- 1 Se creó un arreglo electrónico sencillo para poder conectar el sensor infrarrojo MLX90614ESF-DCC, el cuál constaba únicamente de dos resistencias de 4.7k Ohms conectada al sensor, un botón con una resistencia de 10k Ohms y el arduino, todas las partes montadas en una protoboard.
- 2 Se estableció la temperatura considerada como *fiebre* (38°C).
- 3 Si el botón era pulsado entonces el arduino daba la señal al sensor de temperatura de realizar una medición.

6 Fragmentos de tos

- 4 Al soltar el botón se generaba la medición y el sensor enviaba el dato medido hacia el arduino.
- 5 El arduino decodificaba el dato e imprimía la lectura de éste. La lectura se podía observar en un monitor serial.
- 6 Sí la lectura del dato leído por el arduino era menor a 38°C entonces se imprimía la temperatura detectada con un mensaje que decía “tu temperatura es normal”; caso contrario se imprimía la temperatura con un mensaje que decía “tienes fiebre, tu temperatura es riesgosa monitorea”.
- 7 Por último, se estableció una medida inferior y una medida superior límite debido a que las mediciones debían detectar temperaturas en el rango de 35°C a 41°C, las temperaturas más altas y más bajas no tenía sentido detectarlas debido a que en ese rango se encuentra la temperatura de los seres humanos.

#### **3.4.2. Estructura funcional del sensor de tos**

Para la creación del sensor de tos se consideró que éste debía distinguir 3 tipos de sonido: 1) toses positivas, 2) toses negativas y 3) ruidos.

Debido a que ya se contaba con las toses positivas y negativas consiguientemente se obtuvieron los sonidos faltantes o sea los ruidos, para ello se grabaron audios aleatorios (de la misma medida) de ruidos diversos entre ellos ladridos de perros, sonidos de automóviles, voces de personas, música, llantos de bebés, risas etc. La duración de estos audios fue de 1 segundo.

#### **3.5. Aplicación móvil.**

Se debe señalar que la aplicación se desarrolló para controlar los sensores. Finalmente se trabajó en la aplicación móvil en la plataforma APP Inventor, aquí se crearon los bloques necesarios para la integración y combinación del código final que incluía el funcionamiento de ambos sensores y se sustituyeron los botones físicos que desencadenaban la toma de muestras por botones digitales en esta misma, se integró también el mapa que brinda información sobre la situación

geográfica de COVID-19, su desarrollo ya no fue necesario debido a que Google Maps lo implementó en el 2020 así que solamente se adjuntó el mapa de la página oficial creada por Google.

La aplicación consiste en un modelo amigable con el usuario, inicialmente se pide el nombre y sexo del usuario. Posteriormente se pide que el usuario se tome la temperatura y grabe su tos; se procesan los datos y se arroja un resultado, "*si la temperatura sobrepasa los 38°C se considera factor de enfermedad febril*"(Guyton y Hall, 2011). Al final la aplicación da un mensaje de despedida y se cierra.

(Ver código de bloques en el Apéndice B)

### **3.5.1. Funcionamiento del análisis de datos.**

Como se ha mencionado con anterioridad en la sección de *Características y costos de los materiales*, el arduino Nano 33-BLE Sense tiene integrado un bluetooth Low Energy. Entonces la aplicación llama al BLE para establecer conexión con ella mediante su UUID, después de la vinculación, la aplicación manda un string, el arduino lo detecta y dependiendo de cuál sea (se declararon dos strings en la programación, uno asociado al botón digital "temperatura" y otro a botón digital "tos"), el arduino ordena al sensor de temperatura y al micrófono se generen los datos. Al contar con la información el arduino envía el dato que se pide (tos o temperatura) a la aplicación, posterior a eso se pide el segundo dato (tos o temperatura) y se genera un diagnóstico basado en los resultados de sensibilidad y especificidad obtenidos por el sensor de tos; el funcionamiento del diagnóstico es idéntico al descrito en la *Estructura funcional del sensor de tos*, sin embargo, ahora los mensajes se imprimen en la pantalla del teléfono.



### 3.5.2. Diseño y nombre.

Se realizó el diseño de la imagen principal de la app en el editor de gráficos vectoriales Inkscape; la imagen es un lobo como condecoración a la mascota de la universidad, el lobo BUAP.

Y el nombre se eligió para la aplicación fue **NanNa** como abreviación, el nombre completo es Nantzin Nahualli en honor a la lengua de los Pobladores Nahuas establecidos en Puebla, cuya traducción al español sería “Mamita bruja” ya que la aplicación es encargada de monitorear y “predecir” un diagnóstico para cuidar de la salud del usuario; es bien sabido que las madres son las mejores haciendo esta labor y la predicción se considera virtud de las brujas.

## 3.6. Software

- Arduino IDE.  
Integrated development environment (IDE) es un entorno de desarrollo multiplataforma compuesto por un conjunto de herramientas, consiste en un editor de código, compilador, constructor de interfaz gráfica y depurador. Su función es desarrollar, cargar y ejecutar programas en placas Arduino. Los programas de arduino son llamados sketch, están compuestos por un fichero con extensión “ino”.
- Se usó para hacer el código del sensor de temperatura y unirlo con el sensor de tos.
  - (i) Wire
  - (ii) Adafruit\_MLX90614
  - (iii) PDM
  - (iv) a505\_interferencing
  - (v) GSM0903\_inferencing
  - (vi) a1310\_inferencing

- Edge Impulse:

Edge impulse es una plataforma de desarrollo especializado en machine learning para microcontroladores o microprocesadores. Se especializa en aplicaciones para la industria y profesionales dónde se puede proporcionar mantenimiento predictivo, seguimiento y monitoreo de activos, además detección de actividades humanas y de animales. Proporciona herramientas para el soporte al usuario en el proceso de desarrollo, de ML acelerando el tiempo de desarrollo de aplicaciones, habilitando modelos altamente optimizados que se pueden implementar en una amplia gama de hardware. Sus herramientas son capaces de manipular conjuntos de datos: adquisición, visualización, análisis, procesamiento, filtrado, creación de datasheet, etc. Su biblioteca contiene algoritmos extensibles, personalizables y aplicación de datos de forma automatizada. Posee también la infraestructura necesaria para la creación y entrenamiento de redes neuronales, permitiendo siempre al usuario tener el control de cada parte del proceso. Permite crear librerías en C++ sin dependencias externas, librerías para Arduino, firmware para gran variedad de placas de desarrollo, etc. Gutiérrez (2021)
- Se desarrolló ahí la red neuronal.
  - i) Edge Impulse CLI
    - a. Dependencias:
      - Python 3
      - Node.js
  - ii) Arduino CLI
  - iii) Terminal Ubuntu
- App Inventor.

App Inventor es un entorno de programación por bloques creada por el Instituto Tecnológico de Massachusetts y Google Education que permite crear aplicaciones móviles de manera amigable. Proporciona una librería con

miles de bloques asociados a un código de programación, los bloques pueden agruparse y combinarse en cadena para poder crear la aplicación. Es una herramienta gratuita y se pueden crear aplicaciones para Android como para iOS.

- Aquí se hizo la aplicación móvil.
  - (i) MIT AI2 Companion
  - (ii) Bluetooth Low Energy library
- Inkscape  
Editor de gráficos vectoriales de código abierto. Los gráficos vectoriales son imágenes descritas por funciones matemáticas que permiten redimensionar imágenes sin pérdida de calidad. El formato principal que usa Inkscape es Scalable Vector Graphics (SVG), pero es posible exportar e importar a distintos tipos de archivo.
- Se crearon los fondos de pantalla de la aplicación.
- Audacity  
Programa de grabación y edición de audio multiplataforma de uso libre y código abierto.
- Aquí se dividieron los audios en fragmentos de tos.
- LibreOffice.  
Paquete de software de oficina de uso libre y código abierto.
- Se realizó el registro de datos en Libre office calc.
- LaTeX  
Sistema de composición de textos orientado a la creación de documentos científicos y técnicos, libros etc. que contengan el uso de fórmulas matemáticas.
- En este programa se realizaron todas las tablas.
  - i) Texmaker

- ROOT (Data Analysis Framework).  
Framework para el desarrollo de aplicaciones de análisis de datos científicos a gran escala desarrollado por el CERN, está orientado a objetos y su lenguaje es C++.
- En esta plataforma se graficaron todos los histogramas.
  - i) dpkg-dev
  - ii) cmake
  - iii) g++
  - iv) gcc
  - v) binutils
  - vi) libx11-dev
  - vii) libxpm-dev
  - viii) libxft-dev
  - ix) libxext-dev
  - x) libpng
  - xi) libjpeg
  - xii) Python
  - xiii) libssl-dev

### **3.7 Protocolo de medida.**

#### *Sensor de temperatura.*

El proceso de toma de temperatura consistía en poner el sensor cerca de la frente del paciente (máximo 5 cm) y pulsar el botón físico llamado “Temperatura” en la placa del circuito.

### *Sensor de tos.*

Para poder tomar la muestra de las toses (al contar únicamente con grabaciones) se colocaba el arduino a una distancia de 2-5 cm de la bocina de la computadora que emitía el sonido de las toses. Se reproducía el FT en reproducción continua y automática y se pulsaba el botón físico de “Tos”

Cada modelo de tos se sometió a un periodo de pruebas de calidad para conocer la exactitud de diagnóstico de cada modelo. Un periodo o ciclo constaba de 5 pruebas.

## **3.8 Análisis de datos**

- Sensor de tos.

**NOTA:** Para aumentar la exactitud del sensor de tos se crearon varios modelos de detección distintos, sin embargo, los que dieron mejores resultados fueron los modelos llamados: **modelo 505**, **modelo 1310** y **modelo 903**, la diferencia entre los tres modelos fueron los parámetros.

Ya con todos los datos necesarios se comenzó a trabajar en el código. Este se trabajó con deep learning, es decir con redes neuronales, inicialmente yo ingresé los FT con cada una de sus etiquetas de clasificación, por ejemplo ingresé ***n*** fragmentos de tos COVID-19 positiva con la etiqueta **1** y esa etiqueta la asocié a las “toses covid positivas”, también ingresé ***m*** FT COVID-19 negativos con la etiqueta **2** asociada a las “toses covid negativas” e ingresé ***p*** FR<sup>7</sup> con la etiqueta **3** asociados a fragmentos de “ruido”, las redes neuronales detectaron las características de cada una de las clasificaciones de las señales de audio.

7 Fragmentos de ruido

Estas características las obtuvieron gracias al análisis de Fourier (FFT) midiendo con precisión la frecuencia y el tiempo de duración de cada sonido, en otras palabras, con el espectrograma. Se eligió el rango de frecuencias y el filtro que se le asoció a este. Después las neuronas aprendieron a distinguir de entre las características, para cuando se ingresaba un dato nuevo entonces la red neuronal había aprendido a clasificar el audio en cuestión y en consecuencia a asociarlo a una de las tres etiquetas posibles, pero además también detectaba anomalías o inconsistencias, por ejemplo si algún FT o FR tenía asociado un 50% de sus características a la etiqueta 1 y el otro 50% a la etiqueta 2 o 50% asociado a la etiqueta 2 y el otro 50% a la 3 o cualquiera de las combinaciones posibles, también si le asociaba una 33.33% de características a cada una de las etiquetas o simplemente no podía realizar la clasificación, entonces se asociaba a una nueva etiqueta denominada “ambiguo”.

Todo esto se desarrolló en la plataforma *Edge Impulse* dónde es posible exportar el código de la red neuronal como una librería para la IDE de arduino.

Durante la edición del código en la IDE se asociaron las características para que el programa pudiera dar un diagnóstico, es decir:

- Sí el resultado final del sensor de temperatura era *fiebre* y el del sensor de tos era *Tos Covid positiva*, entonces el diagnóstico final era que el paciente en cuestión era “COVID-19 Positivo” en cierto porcentaje de exactitud (el cuál fue determinado de acuerdo a las pruebas experimentales).
- Sí el resultado final del sensor de temperatura era *temperatura normal* y el del sensor de tos era *Tos Covid positiva*, entonces el diagnóstico final era que el paciente en cuestión era “COVID-19 Positivo” en cierto porcentaje de exactitud menor al primer caso.
- Sí el resultado final del sensor de temperatura era *fiebre* y el del sensor de tos era *Tos Covid negativa*, entonces el diagnóstico final era que el paciente en cuestión era “COVID-19 Negativo” con cierto porcentaje de exactitud.
- Sí el resultado final del sensor de temperatura era *temperatura normal* y el del sensor de tos era *Tos Covid negativa*, entonces el diagnóstico final era

que el paciente en cuestión era “COVID-19 Negativo” con menor porcentaje de exactitud al caso anterior.

- Sí el resultado final del sensor de temperatura era *fiebre o temperatura normal* y el del sensor de tos era *Ruido o ambiguo*, entonces se pedía que se repitiera la prueba.
- Posteriormente a la creación del código del sensor de tos, se unieron ambos códigos en la IDE de arduino y se implementó un botón y una resistencia más al diseño electrónico ya montado en la protoboard para poder disparar la orden por medio del botón de que el micrófono debía comenzar a grabar y de esta forma se pudieran detectar los datos asociados al sensor de tos, o dicho de otro modo detectar la tos de los pacientes. (Véase el apéndice A)

Respecto al espectrograma y el procesamiento neuronal, se trabaja con frecuencias de 16k Hz. La cantidad total de muestras de FT obtenidas de *Virufy* y del *Banco de tos* fueron 98 para la **etiqueta 1**<sup>8</sup>, 101 para la **etiqueta 2**<sup>9</sup> y 70 para la **etiqueta 3**<sup>10</sup>, es importante mencionar que la ausencia de sonido (silencio) se consideró dentro de la etiqueta 3.

Es decir, un total de 199 FT y 70 FR, dónde cada uno de los audios ya fueran de FT o FR medían 3 segundos.

Para cada modelo se realizaba una selección, la cual consistía en realizar diversas pruebas, la primera de ellas era una prueba general la cual se basaba en el análisis de 10 toses, si el modelo en turno detectaba bien 6 o más toses entonces se realizaban las siguientes pruebas, si no se descartaba y se pasaba a un nuevo modelo con nuevos parámetros.

8 FT asociados a la etiqueta de “toses Covid positivas” (**etiqueta 1**)

9 FT asociados a la etiqueta de “toses Covid negativas” (**etiqueta 2**)

10 FR asociados a la etiqueta de “Ruido” (**etiqueta 3**)

Cuando el modelo era viable se sometía a un proceso cíclico que constaba de 5 pruebas, al terminarse el proceso se repetía nuevamente, es decir, cada ciclo de 5 pruebas se repetía 5 veces.

En la primera prueba de los ciclos se elegían las toses que pudieran clasificarse con el modelo con el que se estuviera trabajando sin importar si los FT eran correctamente clasificados, es decir solamente importaba que el modelo fuera capaz de distinguir toses; en otras palabras, se buscaba únicamente que cualquier FT se pudiera clasificar en alguna de las tres posibles opciones. El objetivo de ello era calcular el número de FT distinguibles.

Se hicieron diversas pruebas con distintas cantidades de muestras o tiempos de procesamiento para crear los modelos, siendo la más eficiente la prueba del modelo 505 donde los tiempos de procesamiento (en segundos) se pueden apreciar a continuación en la tabla 2.

*Tabla 2. Tiempo total de datos ingresados a la red neuronal.*

| Modelo 505, Tiempo de recolección. |                        |              |
|------------------------------------|------------------------|--------------|
| Covid Positivos<br>(s)             | Covid Negativos<br>(s) | Ruido<br>(s) |
| 82                                 | 82                     | 80           |

### ***Espectrograma:***

Los valores que se muestran en la tabla 3, son los parámetros asociados al espectrograma para los tres modelos del sensor de tos.

*Tabla 3. Parámetros del espectrograma.*

| Espectrograma      |                             |                              |                              |                               |                                |                          |
|--------------------|-----------------------------|------------------------------|------------------------------|-------------------------------|--------------------------------|--------------------------|
| Frecuencia<br>(Hz) | Ventana<br>de<br>tiempo(ms) | Ventana<br>de<br>aumento(ms) | Longitud<br>de<br>cuadro(ms) | Paso<br>por<br>fotograma(f/s) | Frecuencia<br>de<br>bandas(Hz) | Ruido de<br>piso<br>(dB) |
| 16k                | 1000                        | 500                          | 200                          | 100                           | 128                            | -52                      |



El funcionamiento del espectrograma es el siguiente: primero se divide la ventana en distintos marcos superpuestos, el tamaño del fotograma se determina con la *longitud de cuadro (frame length)* y el número de fotogramas con el *paso por fotograma (frame stride)*. Luego cada marco se divide en intervalos de frecuencia utilizando FFT y se calcula su espectro de potencia. El número de intervalos de frecuencia es igual al parámetro de bandas de frecuencia dividido por dos más uno. Finalmente, el ruido de piso se aplica al espectro de potencia.

Las características generadas por el bloque de espectrograma son iguales al producto del número de marcos de tiempos generados por el número de intervalos de frecuencia.

Existe una relación entre el parámetro de las bandas de frecuencia (tamaño del FFT) y la longitud del cuadro. La longitud del cuadro se recorta o se llena con el valor de las bandas de frecuencia mientras se aplica la FFT. Edge Impulse (2022).

### **Red neuronal:**

La arquitectura de la red se basó en un modelo secuencial<sup>11</sup>, es una red convolucional (1D), es decir que aprende directamente de los datos sin necesidad de extraer características manualmente. Tiene un número de 100 ciclos de entrenamiento y un learning rate de 0.005 segundos. Los FR se proporcionaron únicamente para entrenar la red neuronal.

Edge impulse es capaz de proporcionar gráficas de acuerdo a los parámetros ingresados, dichas gráficas son representaciones en 3D de la distribución de FT y FR, sus ejes coordenados estaban conformados por: tamaño de la ventana de tiempo, frecuencia y tamaño de una ventana de aumento de tiempo (sirve para determinar el corrimiento de una ventana de tiempo cuando la muestra a analizar es más grande que la ventana de tiempo).

11 La API de Keras tiene dos modos de construcción de redes neuronales: Modelo secuencial (permite que las capas se agreguen en secuencia) y Modelo funcional (permite la construcción de gráficos en capas, es decir puede usar un solo grafo de capas para generar múltiples modelos)

### 3.9 Estadísticas.

*Periodo de prueba para el sensor de tos.*

#### **I Clasificación FT preliminar (prueba 1)**

Durante la presente prueba se clasificaban los 199 FT, cada uno se asociaba a una etiqueta.

#### **II Depuración (prueba 2)**

El total de FT a depurar en esta prueba dependía del número de fragmentos etiquetados como “Ruido” o “Ambiguo”, a cada uno de los FT clasificados con ambas etiquetas se les detectaba nuevamente con el sensor de tos un total de 3 veces por FT, si obtenían una clasificación  $x \geq 2$  ya sea “tos cov positiva” o “tos cov negativa”, entonces la nueva clasificación era la determinada por el mayor número de etiquetas obtenidas en esta parte, en el caso contrario se tomaban dos consideraciones para su clasificación:

i Si el sonido de los FT clasificado con la **etiqueta 3** sonaba no distinguible (no sonaba como una tos) entonces obtenía nuevamente la clasificación de “Ruido” o “Ambiguo” y se anulaba.

ii En caso contrario (si el FT era distinguible, es decir que si sonaba como tos) se le hacían nuevamente 10 detecciones con el sensor de tos; la etiqueta asignada a cada detección, con mayor número de repeticiones era la clasificación asignada al FT en cuestión.

#### **III Clasificación FT filtrados (prueba 3)**

La cantidad de FT resultantes de la prueba “*Depuración*” que tenían la nueva clasificación asignada ya sea a la **etiqueta 1** o a la **etiqueta 2** se recontaban.

#### **IV Conteo FT negativos (prueba 4)**

Se separaban los FT clasificados con la **etiqueta 1** y se contaba el total.

#### **V Conteo FT positivos (prueba 5)**

Se separaban los FT clasificados con la **etiqueta 2** y se contaba el total.

Este ciclo de 5 pasos se repetía 5 veces. Es decir que para la primera fase se analizaron un total de 995 FT solamente para un modelo, para los 3 modelos se

analizaron 2,985 FT para la primera prueba; para la segunda prueba del modelo 505 se analizaron en promedio 537 (Pruebas con etiquetas 1 y 2) más 3750 (para la etiqueta 3), o sea 4287 en total. En la tercera, cuarta y quinta parte, solamente se reorganizaron y recontaron los FT. Finalmente se promediaba todo.

(Se menciona solamente los números para el modelo 505 porque fue el de mejores resultados)

### **3.9.1 Validez de una prueba diagnóstica**

Para someter a revisión los datos arrojados en el presente proyecto debemos tener claro el concepto de prueba diagnóstica, “Una prueba diagnóstica se refiere a cualquier método para obtener información adicional del estado de salud del paciente” (Bravo & Cruz, 2015, pág. 158).

La evaluación de una prueba diagnóstica arroja cierta incertidumbre bajo la cual se decide, con base en ciertos parámetros, qué tanto esa incertidumbre valida o invalida la prueba diagnóstica (exactitud diagnóstica).

Los resultados obtenidos con la prueba diagnóstica en evaluación se comparan con los de un estándar de referencia en un mismo grupo de pacientes. El estándar de referencia, también llamado Gold standard (estándar de oro) corresponde, en resumen, a la mejor manera disponible y ampliamente aceptada para establecer la presencia o ausencia de determinada condición. (Bravo y Cruz, 2015, pág. 159)

Usualmente las pruebas diagnósticas se cuantifican según los parámetros:

- Sensibilidad (S)  
Este parámetro se refiere a la cantidad de pacientes con la enfermedad que tendrán un resultado positivo en la prueba diagnóstica.
- Especificidad (E)  
Se refiere a la cantidad de pacientes sin la enfermedad que obtendrán un resultado negativo en la prueba diagnóstica.
- Valor Predictivo Positivo (VAPP)

La probabilidad de que una persona que se realizó la prueba y obtuvo resultado positivo tenga la enfermedad.

- Valor Predictivo Negativo (VAPN)

La probabilidad de que una persona que se realizó la prueba diagnóstica y obtuvo resultado negativo no tenga la enfermedad.

- Exactitud (EX)

La probabilidad de que la prueba diagnóstica prediga correctamente si el paciente tiene o no la enfermedad.

Este análisis se hace en función de cuatro clasificaciones posibles de acuerdo a los resultados experimentales: *verdadero positivo*, *falso positivo*, *falso negativo* y *verdadero negativo*.

- *Verdadero positivo (VP)*: el paciente tiene la enfermedad y la prueba diagnóstica es positiva.
- *Falso positivo (FP)*: el paciente no tiene la enfermedad y la prueba diagnóstica es positiva.
- *Verdadero negativo (VN)*: El paciente no tiene la enfermedad y la prueba diagnóstica es negativa.
- *Falso negativo (FN)*: El paciente tiene la enfermedad y la prueba diagnóstica es negativa.

El cálculo de parámetros se determina con base en los diagnósticos muy amigablemente con las siguientes fórmulas:

$$S = \frac{VP}{VP+FN}$$

$$E = \frac{VN}{FP+VN}$$

$$VAPP = \frac{VP}{VP+FP}$$

$$VAPN = \frac{VN}{FN+VN}$$

$$EX = \frac{VP+VN}{VP+FP+FN+VN}$$

#### 4. RESULTADOS

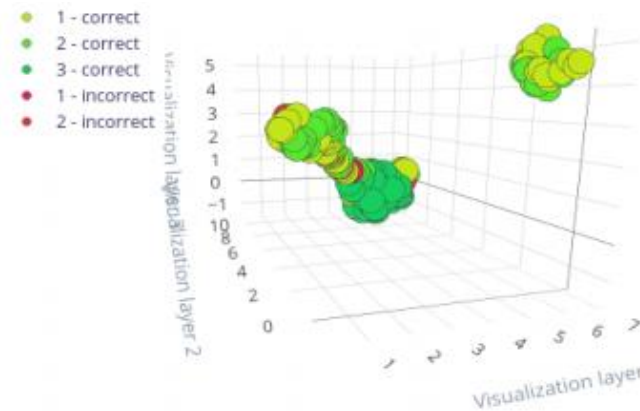
Rememorando, entre los principales síntomas del SARS-Cov-2 se encuentran fiebre, tos y dificultad para respirar; los cuales son un fuerte indicador de riesgo, por ello surge la necesidad de monitorear, registrar síntomas y crear un método de diagnóstico accesible para la mayor parte de la población, convirtiéndose en el objetivo principal del presente trabajo.

Bajo los objetivos de proporcionar un método de detección coadyuvante capaz de realizar monitoreo del COVID-19 se ha desarrollado una aplicación móvil pensada para cubrir las necesidades de una prueba diagnóstica preliminar para la población de bajos recursos; es decir que bajo sospecha de tener COVID-19; antes de hacer una inversión en una prueba diagnóstica comercial, pueden consultar el diagnóstico que proporcione NanNa, dado el caso que la probabilidad de tener COVID-19 se alta entonces tengan más seguridad en encontrar un resultado positivo en SARS-Cov-2 y cobre más sentido realizarse una prueba más exacta o puedan simplemente tomar medidas preventivas necesarias para evitar el esparcimiento del virus. NanNa se piensa accesible ya que la mayoría de la población mexicana dispone de un dispositivo móvil. Consta de un modelo sencillo y práctico formado por cuatro partes principales: 1. Un sensor de temperatura para vigilar las temperaturas del usuario, 2. Un sensor de tos, 3. La ubicación gps del teléfono móvil para determinar la zona geográfica en la que se localiza el usuario y poder obtener un registro de los casos positivos de COVID-19 alrededor de dichas zonas. 4. Una aplicación móvil para realizar las detecciones de tos y temperatura.

En la parte del desarrollo del sensor de tos posteriormente a la asignación de valores del espectrograma y al procesamiento de la red neuronal, la plataforma Edge Impulse proporcionó una imagen de la distribución en tercera dimensión de las toses en el espacio de análisis, los ejes de la gráfica son: tamaño de la ventana de tiempo, frecuencia y tamaño de una ventana de aumento de tiempo

(sirve para determinar el corrimiento de una ventana de tiempo cuando la muestra a analizar es más grande que la ventana de tiempo). Ver la ilustración 5.

Ilustración 5 Distribución de FT del modelo 505



Las siguientes tablas muestran los valores resultantes de los parámetros anteriormente definidos, la tabla 4 muestra las clasificaciones de los datos VP, VN, FP, FN, la tabla 5 muestra los valores resultantes de la sensibilidad y especificidad del datasheet según el número fragmentos de tos de pacientes con y sin COVID-19; por último en la tabla 6 se registraron los promedios de los valores predictivos positivos y negativos, los Likelihood ratios positivos tanto como negativos y finalmente la exactitud.

Tabla 4. Resultados de la matriz de confusión.

| Media de las clasificaciones de muestra de los tres modelos |                |                    |                |
|---|----------------|--------------------|----------------|
| Verdadero Negativo  | Falso Positivo | Verdadero Positivo | Falso Negativo |
| <b>Modelo 505</b>   |                |                    |                |
| 84.4  | 7.8            | 49.6               | 46.6           |
| <b>Modelo 1310</b>  |                |                    |                |
| 80  | 6.2            | 42.8               | 51.4           |
| <b>Modelo 903</b>   |                |                    |                |
| 0   | 72.6           | 87.8               | 0              |

En promedio la suma de los VN, FP, VP, FN debía aproximarse en la mayor medida posible al número total de fragmentos analizados, o sea  $VN + FP + VP + FN \approx FT$  debía aproximarse a 199.

Para el modelo 505 se obtuvo un total de  $FT = 188.8$  analizados, del 100% de datos ingresados, el modelo pudo detectar en promedio el 94.87% FT y tuvo un error relativo de detección de  $\pm 5.13\%$ .

En el caso del modelo 1310 se obtuvo  $FT = 180.4$ , este modelo pudo detectar el 90.65% de FT y obtuvo un  $ER = \pm 9.35\%$  no tan lejano al del modelo 505 y por último el modelo 903 obtuvo  $FT = 160.4$  y  $ER = \pm 19.40$ .

#### **4.1. Sensibilidad y Especificidad en pruebas diagnósticas de COVID-19**

*La Gold Standar* de las pruebas diagnósticas de SARS-Cov-2 es la RT-PCR (exudado nasofaríngeo/orofaríngeo) que da un resultado de sensibilidad de 84-90% de ahí sigue la RT-PCR (saliva) y RT-PCR en exudado nasofaríngeo (Multiplex), que tienen sensibilidad en el rango 5-91%; el Test antígenicos rápidos de última generación (exudado nasofaríngeo) tiene sensibilidad (en personas sintomáticas) mayor al 95% y no hay evidencia para los asintomáticos. Por último, el Test de determinación de anticuerpos depende enormemente del día de toma de la muestra respecto al día de inicio de los síntomas y su sensibilidad va en una escala del 50-90%.

***Para el caso de la prueba de detección de COVID-19 por sonido, el modelo 505 arrojó resultados de sensibilidad del 51%. Para el modelo 1310 la sensibilidad es de 45% en consecuencia queda por debajo del 50% del rango de sensibilidad. Ahora bien, el modelo 903 posee una sensibilidad del 100% es decir mayor al Gold standar.***

En cuanto a especificidad respecta el *Gold standar* tiene una especificidad del 99.5%, RT-PCR (saliva y multiplex) similar a la del *Gold standar*, la prueba de antígenos varía de un 95-99% y el test de anticuerpos varía en un rango de 90-99%. El modelo 505 posee una especificidad del 91%, es decir dentro del rango de la prueba de anticuerpos, el modelo 1310 tiene el mejor valor de especificidad de entre nuestros modelos con un 92% también dentro del rango de la prueba de anticuerpos.

Por último, se encuentra el modelo 903 con una especificidad de 0% lo cual contrarresta los beneficios que pueda aportar como prueba diagnóstica, ya que todo lo detecta como Verdadero Positivo. (Se puede consultar el código de las tablas en el Apéndice C)

*Tabla 5. Resultados de S. y E.*

| Sensibilidad y Especificidad<br>relación con el número de<br>pacientes con y sin COVID-19 |               |
|---|---------------|
| Sensibilidad  | Especificidad |
| Modelo 505  |               |
| 0.51  | 0.91          |
| Modelo 1310   |               |
| 0.45  | 0.92          |
| Modelo 903  |               |
| 1   | 0             |

#### **4.2. Otras probabilidades de la presente prueba diagnóstica**

Como se ha mencionado antes, el valor predictivo positivo de una prueba diagnóstica ayuda a conocer la probabilidad de que un paciente se encuentre enfermo si el resultado de la prueba es positivo. Dicho de otra forma, es la cantidad de personas resultantes con COVID-19 positivo. El modelo con mayor VAPP es el modelo 1310 con un 87% seguido inmediatamente por el 505 con un VAPP del 86%



y en último lugar el modelo 903 con solamente 54%. Por el contrario, el VAPN es mayor para el modelo 505 con el 64%.

*Tabla 6. Probabilidades de las otras pruebas de calidad.*

| Media de las probabilidades para VAPP, VAPN, LHRP, LHRN y EX |                           |                           |                           |           |
|--|---------------------------|---------------------------|---------------------------|-----------|
| Valor Predictivo Positivo                                    | Valor Predictivo Negativo | Likelihood ratio Positivo | Likelihood ratio Negativo | Exactitud |
| Modelo 505   |                           |                           |                           |           |
| 0.86   | 0.64                      | 6.12                      | 0.52                      | 0.71      |
| Modelo 1310  |                           |                           |                           |           |
| 0.87   | 0.60                      | 6.31                      | 0.58                      | 0.68      |
| Modelo 903   |                           |                           |                           |           |
| 0.54   | 0                         | 1                         | 0                         | 0.54      |

*Tabla 7. Rangos de LHR y su utilidad.*

| Rangos Likelihood Ratios y su utilidad |                           |                     |
|--|---------------------------|---------------------|
| Likelihood Ratio Positivo              | Likelihood Ratio Negativo | Clasificación       |
| 10                                     | <0.1                      | Altamente relevante |
| 5-10                                   | 0.1-0.2                   | Buena               |
| 2-5                                    | 0.5-0.2                   | Regular             |
| <2                                     | > 0.5                     | Mala                |

Nota: Alba & Villagra (2016) mencionan los anteriores rangos de valores de Likelihood de utilidad clínica.

Se puede observar entonces que el modelo 1310 y el 505 tienen una buena clasificación LHRP ya que sus valores se encuentran entre 5-10, mientras que el modelo 903 tiene un LHRP=1, por lo tanto, su clasificación es mala y respecto a los LRHN los valores de los tres modelos son malos.

Exactitud:

La exactitud de las pruebas varía dependiendo de algunos factores como la técnica de recogida, el tipo de muestra (nasofaríngea/orofaríngea, etc.), la prevalencia de la enfermedad en la población, sensibilidad de las personas asintomáticas, etc. La exactitud con mejor resultado se encuentra entre el rango de 90-99% siendo la más alta la *Gold standard*. Las exactitudes de los modelos realizados por nosotros varían de 54 a 71% en efectividad.

#### **4.3. Diferenciación entre modelos.**

Durante la segunda etapa del ciclo sí se consideran los resultados obtenidos en la clasificación de los datos preliminares (datos de la primera etapa), es decir, en la segunda etapa sí se analiza la efectividad (de todos los modelos) al clasificar los datos, las clasificaciones se diferencian en las siguientes tres clases:

- Bien: muestras bien clasificadas.
- Indeterminado: en esta clasificación se determinaron tres cosas, la primera: sonidos que no eran tos (ruidos), la segunda ausencia de ruidos (se clasificó dentro de “ruido”) y la tercera: si el modelo era incapaz de identificar lo que se le pedía analizar es decir una señal “ambigua”.
- Mal: muestras mal clasificadas

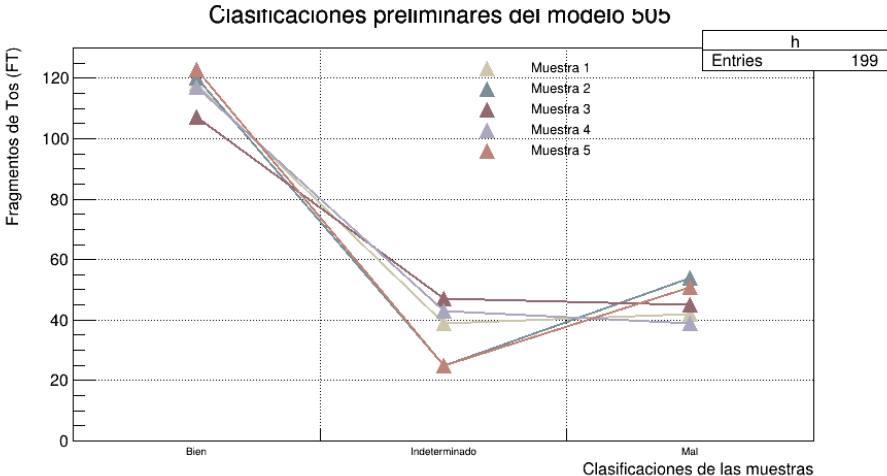
En los siguientes histogramas se muestran los valores obtenidos en cada categoría para cada uno de los modelos en los 5 ciclos de prueba correspondientes.

Hay que recordar que FT=199.

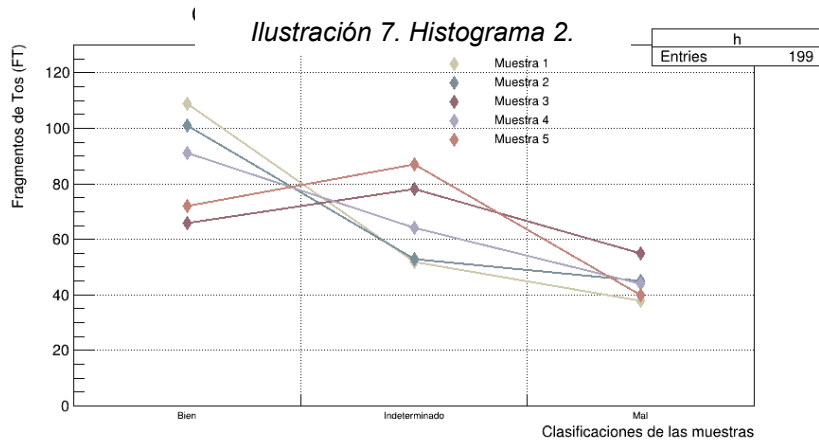
En el modelo 505 se puede observar que la mayoría de los ciclos coinciden en resultados de clasificación “Bien”, sin embargo, en el ciclo 3 (muestra 3) existe un salto notable en la disminución de resultados clasificados correctamente en dicha

categoría. La media de los valores “Bien” es de 115.5, su desviación estándar es de 5.23%, la de “Mal” es de 46.2, su desviación estándar de 13.46% y de “Indeterminado” 35.8 y su desviación estándar del 28.65%. (Ver el código de los histogramas en el Apéndice D)

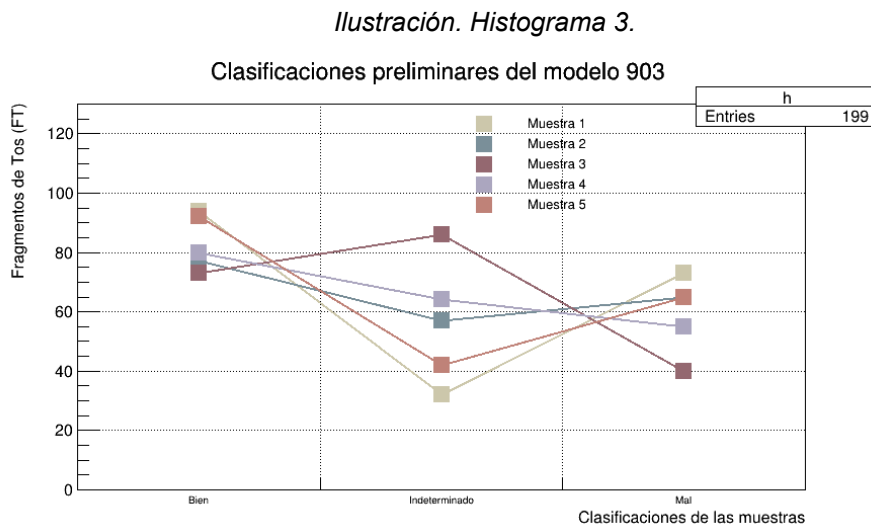
Ilustración 6. Histograma 1.



En el caso del modelo 1310 los valores para las tres categorías son muy variados, para la categoría “Bien” encontramos una media de 87.8, su desviación estándar de 20.99%, para “Mal” la media es de 44.4 su desviación estándar de 14.82% y para “Indeterminado” 66.8 y 23.08% para media y desviación estándar respectivamente



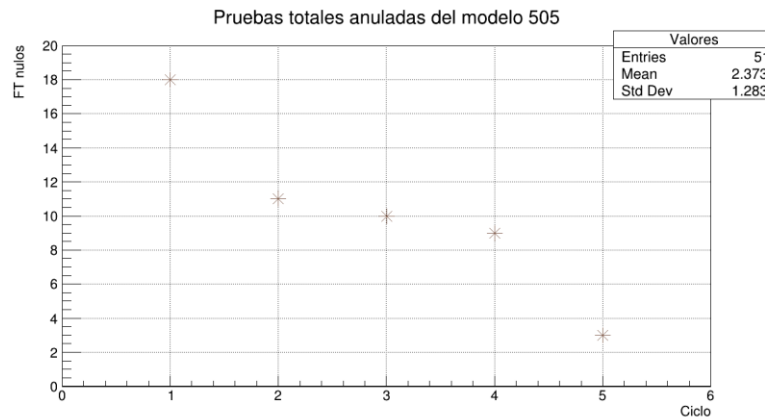
El modelo 903, obtuvo un total de 83.2 para la media y 11.19% en desviación estándar en “Bien”, “Mal” con 59.6 en media y 21.27% desviación estándar y finalmente 56.2 media y 37.07% desviación estándar para “Indeterminado”.



Notablemente para la primera etapa del ciclo, el mejor clasificador de pruebas “Bien” fue el modelo 505, con una menor desviación estándar, posteriormente el modelo 903 y por último el modelo 1013 que tuvo una mayor dispersión de datos. En “Mal” el modelo con menor dispersión fue nuevamente el 505, seguido del modelo 1310 y por último el 903. Finalmente, en “Indeterminado” el modelo con menor dispersión fue 1310 y nuevamente la mayor dispersión fue de 903.

El tercer paso del ciclo consistía en hacer la depuración o eliminación de ruido, el modelo 505 obtuvo los resultados mostrados en la gráfica 4.

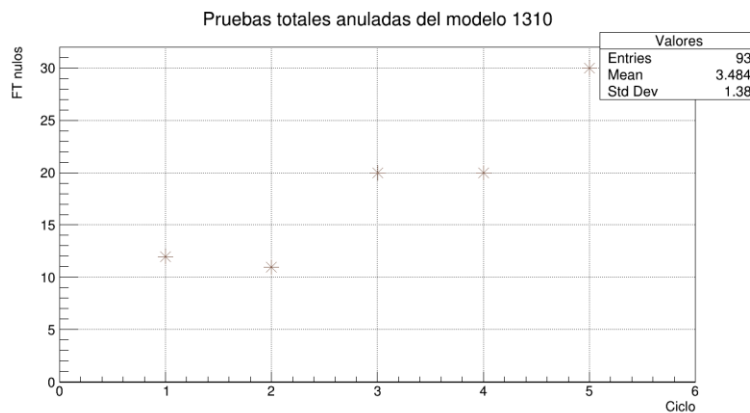
*Ilustración 8 Histograma 4.*



Nótese que el mayor número de FT eliminados fue en el primer ciclo de prueba y comenzó a disminuir en los ciclos posteriores.

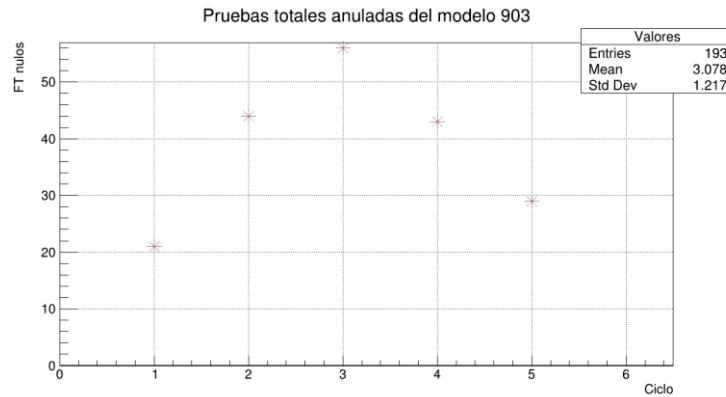
Caso contrario al modelo 1310, que conforme se hacía un nuevo ciclo de prueba se mantenía aproximado al dato anterior o aumentaba de manera muy considerable.

*Ilustración 9. Histograma 5.*



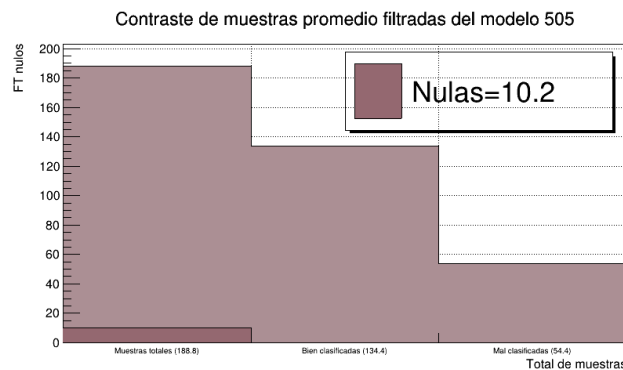
El modelo 903 obtuvo una distribución aproximada a una distribución normal. En los primeros tres ciclos aumentó gradualmente y en el tercero y cuarto disminuyó de la misma forma.

Ilustración 10. Histograma 6.



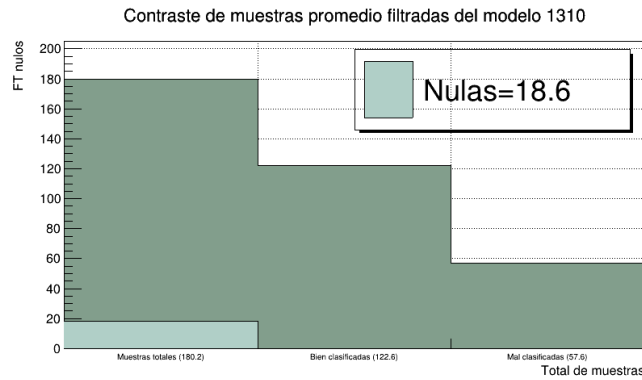
En la tercera parte del ciclo, cuando se llevó a cabo el reajuste de valores posterior al filtrado, es decir los valores obtenidos en las distintas categorías posterior a la eliminación de valores nulos (FT no distinguibles). De un total de 995 FT, se anularon 5.12% para el modelo 505 véase la gráfica 7, 9.34% para el modelo 1310 (gráfica 8) y 19.39% para el modelo 903 (gráfica 9).

Ilustración 11. Histograma 7.



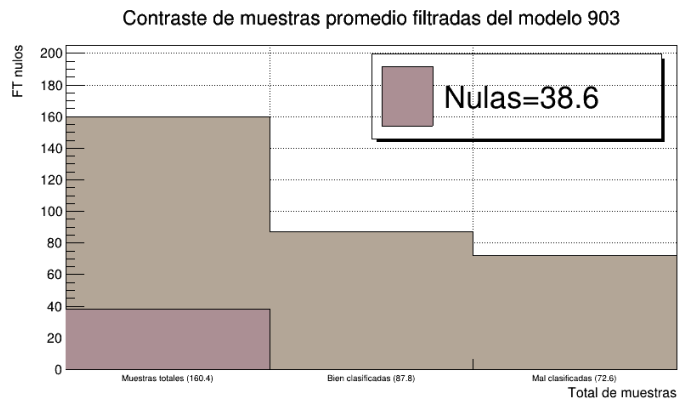
Note que el modelo 1310 tiene un porcentaje menor en efectividad para la detección correcta de muestras, respecto al modelo 505 la diferencia entre efectividad de detección es apenas del 4.56%

Ilustración 12. Histograma 8.



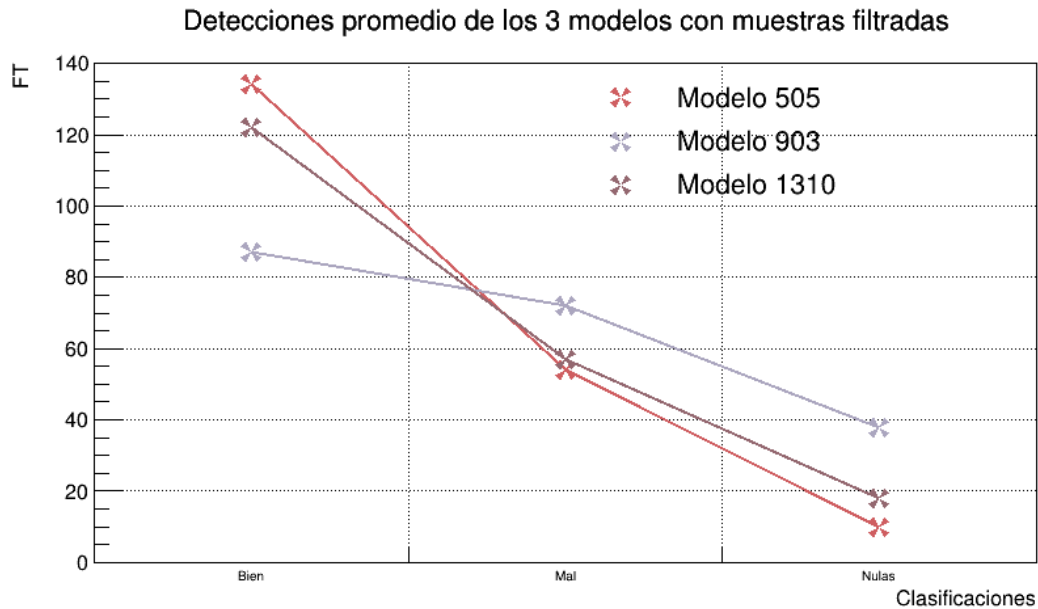
Por el contrario, el modelo 903 tiene una diferencia de 15.05% de efectividad respecto al modelo 505.

Ilustración 13. Histograma 9.



En términos generales de entre los tres modelos, el modelo 505 obtiene mejores resultados en las pruebas de filtrado con menor número de nulas y mayor número de FT clasificadas como “Bien”.

Ilustración 14. Histograma 10.



En el paso número 4 y 5 para cada ciclo se hacía la clasificación final de los FT de manera formal para hacer la evaluación a la prueba diagnóstica VN, FP, VP y FN, los resultados se pueden apreciar en la tabla 8.

*Tabla 8. Comparación de medias y desviaciones estándar.*

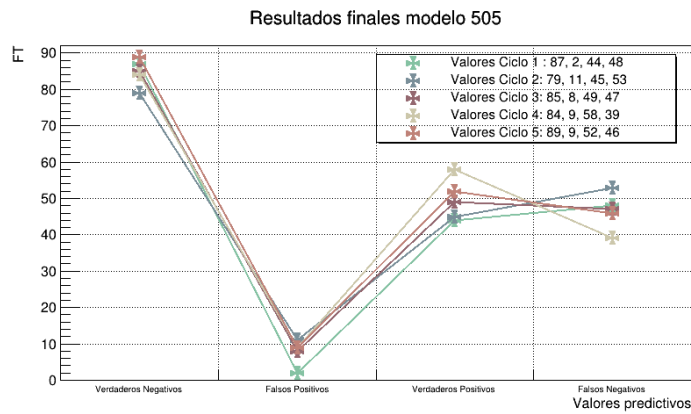
| Medias para VN, FP, VP y FN |      |      |      | Desviaciones estándar para VN, FP, VP y FN |      |       |       |
|-----------------------------|------|------|------|--|------|-------|-------|
| VN                          | FP   | VP   | FN   | VN   | FP   | VP    | FN    |
| <b>Modelo 505</b>           |      |      |      | <b>Modelo 505</b>                          |      |       |       |
| 84.8                        | 7.8  | 49.6 | 46.6 | 3.76                                       | 3.42 | 5.68  | 5.02  |
| <b>Modelo 1310</b>          |      |      |      | <b>Modelo 1310</b>                         |      |       |       |
| 80                          | 6.2  | 42.8 | 51.4 | 3.80                                       | 4.54 | 13.38 | 11.22 |
| <b>Modelo 903</b>           |      |      |      | <b>Modelo 903</b>                          |      |       |       |
| 0                           | 72.6 | 87.8 | 0    | 0  | 8.7  | 6.37  | 0     |



### Modelo 505.

Nótese la concordancia de resultados para cada uno de los ciclos en cada una de las clasificaciones, la dispersión entre ellos no es muy notoria.

Ilustración 15. Histograma 11.

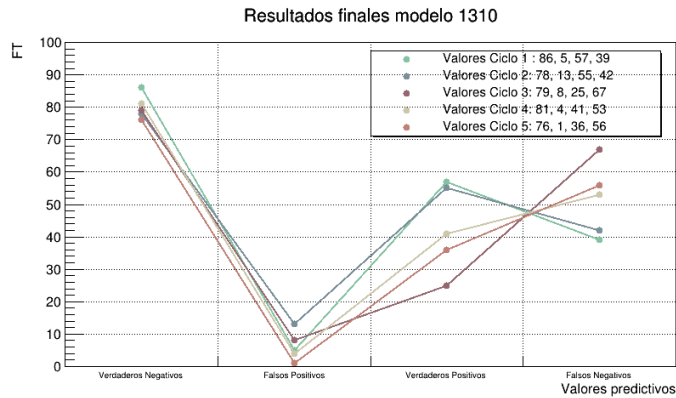


La desviación estándar para VN fue del 4.44%, para FP del 43.85%, en VP del 11.45% y en FN del 10.73%.

### Modelo 1310

En el presente modelo la desviación estándar para VN, FP, VP y FN fueron de 4.75%, 73.38%, 31.27% y 21.95% respectivamente.

Ilustración 16. Histograma 12.

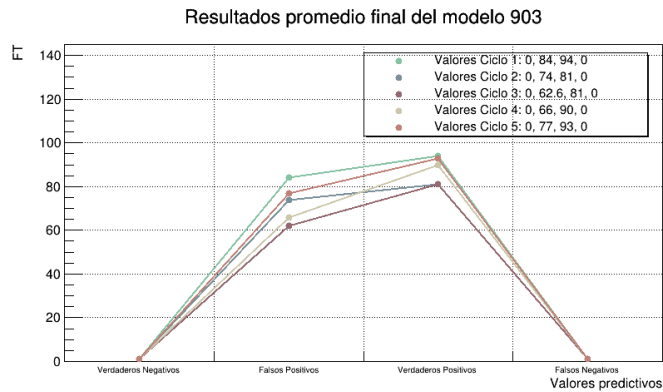


En la gráfica 12 se puede apreciar que hay saltos abruptos en los resultados de VP y FN.

### **Modelo 903.**

Para este modelo no hubo datos clasificados como VN y FN, en los FP y VP la desviación estándar fue de 12.07% para el primer caso y de 7.26% para el segundo.

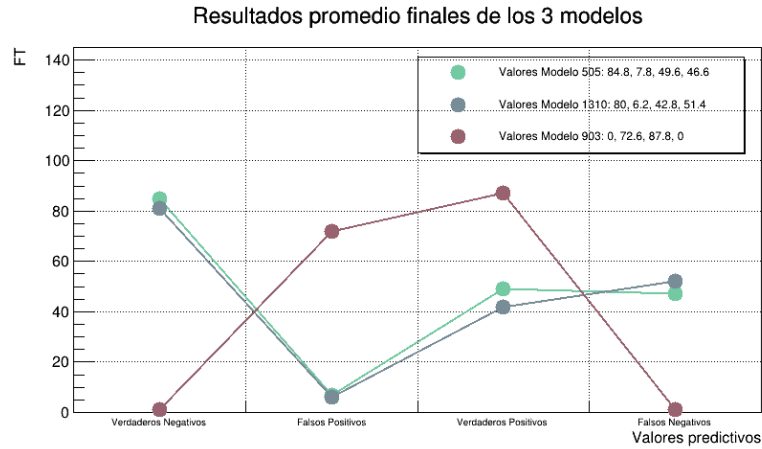
*Ilustración 17. Histograma 13.*



La dispersión de datos para FP fue mayor que para VP.

Observe en la gráfica 14 que, en promedio de los 5 ciclos correspondientes para cada modelo, el más eficaz, es decir, el que clasifica de mejor manera es el modelo 505, seguido ligeramente por debajo por el 1310. El modelo 903 tiene mejor detección de verdaderos positivos, sin embargo, en los otros tres campos su resultado deja bastante que desear.

Ilustración 18. Histograma 14.



Con la información de la gráfica 14 se obtuvieron los resultados que se muestran en los tres histogramas siguientes (histogramas 15, 16 y 17). En consecuencia, de los resultados obtenidos para las categorías FP, VP, FN y VN la comparación de sensibilidad, especificidad y valores predictivos positivos y negativos tienen mejor resultado en el modelo 505.

Ilustración 19. Histograma 15.

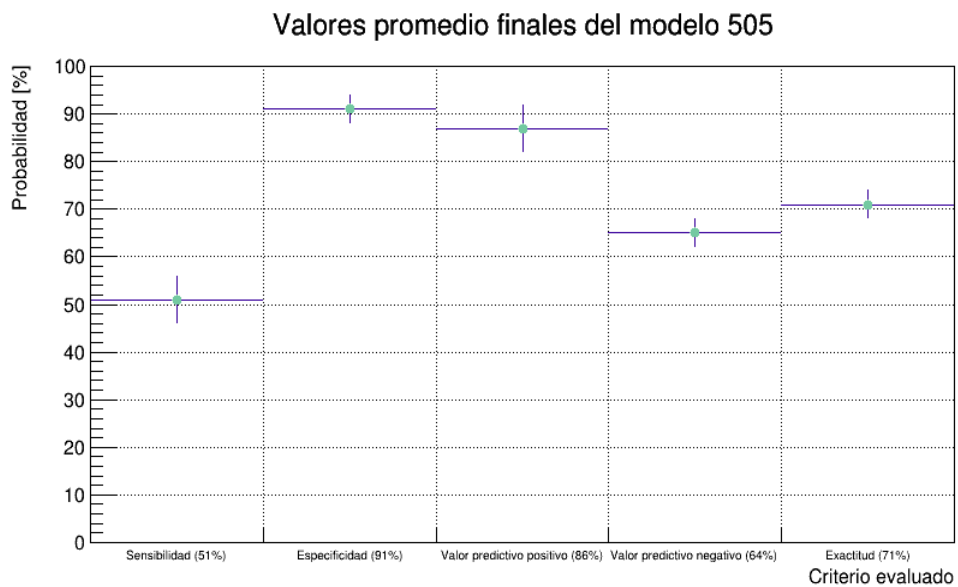
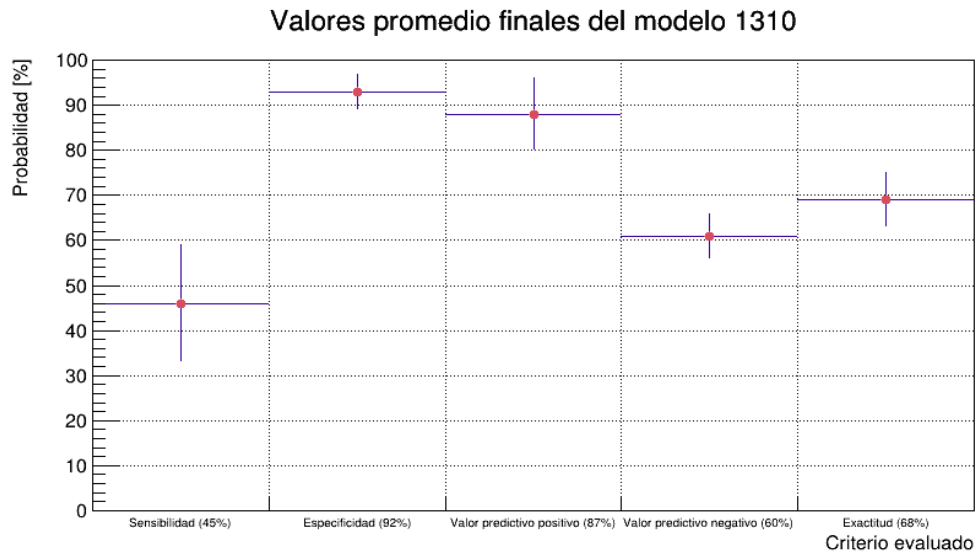
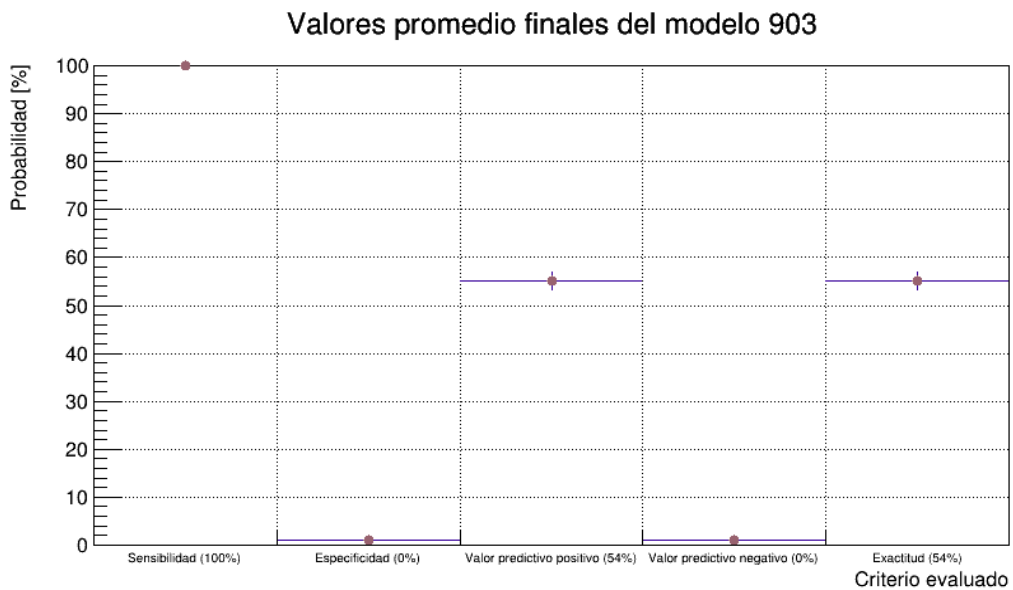


Ilustración 20. Histograma 16.



El modelo 903 no tiene valor alguno en especificidad y en valor predictivo negativo, esto se debe claramente a que el modelo obtuvo 0 clasificaciones en VN y FN.

Ilustración 21. Histograma 17.



## 5. DISCUSIÓN

Durante el proceso de clasificación de FT sin importar si era correcto o no el resultado, el modelo 505 fue el más aproximado al obtener 188.8 de los 199 FT posibles.

El error relativo (ER) aceptable para experimentos muy controlados y con variables bien definidas es de  $\pm 5\%$ . El modelo 505 tiene un error relativo muy cercano al error relativo aceptable, su valor es de  $\pm 5.33\%$ , es importante resaltar que las variables en esta prueba no están bien definidas, no hay investigaciones acerca de los rangos de frecuencias para toses, es decir ningún parámetro está establecido aún. Además, que esta prueba no es controlada.

Basándonos en los datos obtenidos es notable que el modelo más funcional de los aquí realizados, fue el modelo 505, seguido de cerca por el modelo 1310. La incertidumbre de detección, es decir las FT registrados como “Indeterminados” fue menor en el modelo 505, durante el filtrado se eliminó una menor cantidad de FT registradas como “nulas” además la desviación estándar de las muestras clasificadas como VP, FP, VN y FN fue la más cercana a su media respectiva.

En consecuencia, respecto a sensibilidad, especificidad y valores predictivos nuevamente obtuvo un mejor puntaje.

Respecto a las pruebas comerciales la presente prueba se encuentra en el rango más bajo en cuanto a sensibilidad, especificidad y valores predictivos, sin embargo, el modelo 505 sigue muy de cerca a los resultados de las pruebas de anticuerpos e incluso se obtiene un resultado ligeramente mayor en cuanto a sensibilidad si consideramos los primeros días de enfermedad con un porcentaje de 51% contra 50% en los primeros 5 días.

Para el caso de la prueba de detección de COVID-19 por sonido, el modelo 505 arrojó resultados de sensibilidad del 51%, es decir que se encuentra en un rango muy inferior respecto a la sensibilidad de la *Gold standar*, pero dentro del

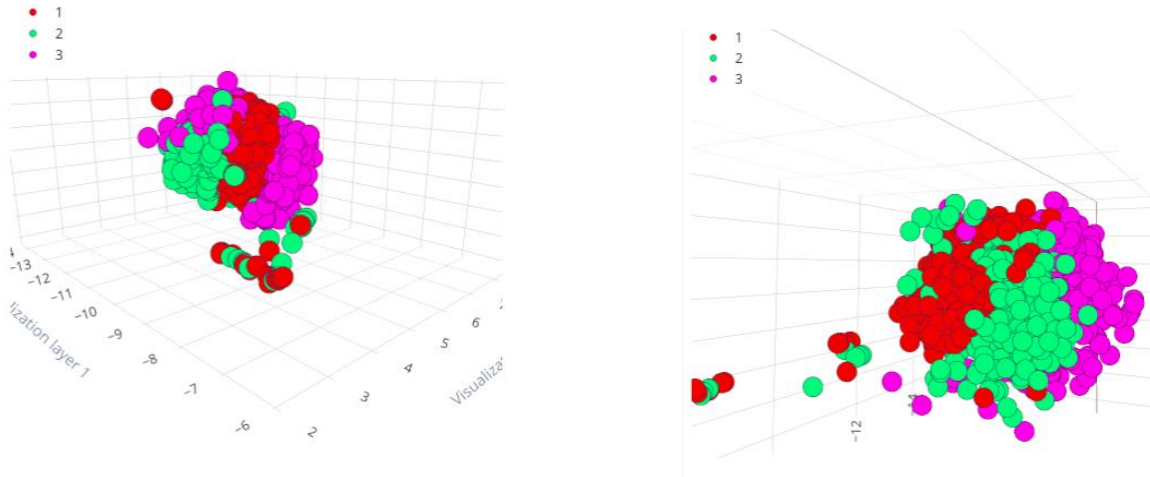
rango de sensibilidades para la prueba diagnóstica “*Test de determinación de anticuerpos*”. Para el modelo 1310 la sensibilidad es de 45%, queda por debajo del 50% del rango de sensibilidad.

Se observó que a mayor número de FT para el aprendizaje la detección de VP aumenta considerablemente sin embargo la detección de VN disminuye de la misma forma, análogamente si disminuimos el número de FT aumenta la detección de VN pero disminuye la de VP; es decir que debe existir una forma de determinar el correcto número de FT para encontrar un modelo que alcance mejores resultados, esta relación puede depender de la forma de agrupamiento de sonidos por el espectrograma (o sea depende del tiempo de análisis de cuadro o longitud del mismo y la frecuencia), debido a que también se observó que por ejemplo con Frame Stride variable y los demás parámetros fijos a potencias muy pequeñas o muy grandes tienden a desordenarse, es decir que no tienen buena agrupación entre las 3 categorías que se habían establecido (tos covid positiva, tos covid negativa y ruido) sin embargo no se pudo establecer ningún patrón debido a que Edge tiene un método establecido de análisis de datos aleatorio en cuanto a qué longitud de la ventana se comienza la medición y en cada prueba nueva los valores de dónde se tomaba el inicio de la ventana de tiempo variaba.

En algunas ocasiones dónde el agrupamiento fue bueno podía distinguirse que el comportamiento de la tos tenía forma helicoidal, lo cual podría arrojar un resultado importante sin embargo repetir el experimento resultaba ser difícil por lo antes mencionado.

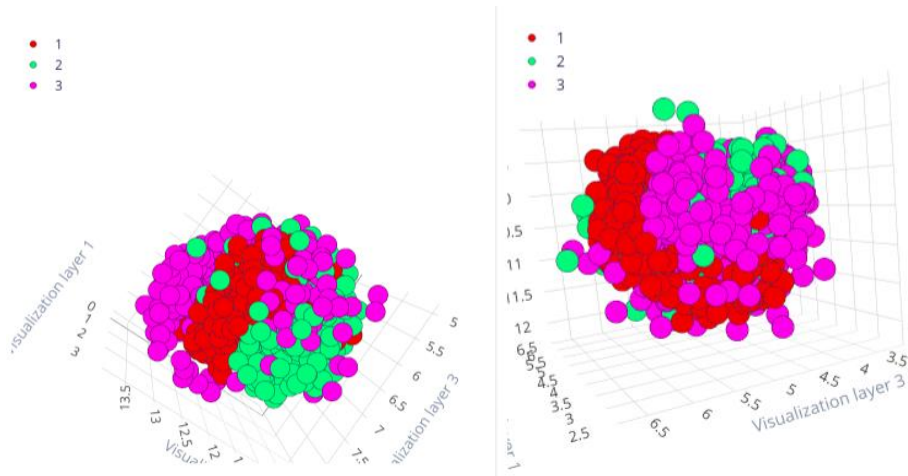
A continuación, se muestra un ejemplo de dos gráficas (gráfica 18 y 19) con vistas de dos ángulos diferentes:

Ilustración 22. Gráfica 18 Distribución helicoidal 1.



Variando otros parámetros igualmente se encontraron resultados similares durante la agrupación de datos cabe mencionar que la gráfica 18 y 19 está prediseñada por Edge Impulse de acuerdo a los datos que nosotros definimos previamente y nosotros no tenemos la capacidad de manipularla.

Ilustración 23. Gráfica 19. Distribución helicoidal 2.



Dicho lo anterior se puede concluir que los modelos son mejorables, es decir se puede alcanzar una mejor sensibilidad, mediando no descuidar el buen puntaje de la especificidad e incluso se puede mejorar.

Una posible ampliación del proyecto podría ser la implementación de un registro de temperatura y tos para una o más personas (dependiendo del número de usuarios que se requiera).

Una segunda mejora puede ser la implementación de órdenes y resultados por sonido para personas con discapacidad visual o problemas visuales e incluso para personas que no están relacionadas en gran medida con la tecnología como los adultos mayores.

Respecto a la inversión resulta ser una prueba “cara” para una primera adquisición debido a que el costo de los sensores es elevado, considerando que si se quiere reproducir por otras personas los gastos se disminuyen en cierta cantidad debido a que muchas de las cosas con las que se fabrica la prueba son cosas que regularmente la gente posee (cautín, cables, pasta para soldar, soldadura de estaño y cable USB de teléfono), los botones y el protoboard no se necesitan debido a que ese diseño solamente fue implementado para hacer el montaje durante el desarrollo y las resistencias son muy económicas. Si se piensa en una comunidad alejada en una zona rural, es posible que el número de habitantes sea relativamente pequeño, por lo tanto, se podría adquirir un sensor para toda la población reduciendo enormemente la inversión por poblador; si se habla de zonas urbanas los vecinos de determinado lugar podrían hacer la adquisición y elaboración colectiva del sistema de igual forma.

Si se piensa en el uso particular entonces la relación costo-beneficio también es alta ya que al adquirir un sistema de estos es reutilizable para una infinidad de veces. Si se adquiere por familia es el mismo caso.

Una incógnita importante de la prueba es que el periodo de prueba se realizó con grabaciones de tos de pacientes COVID-19 positivos y no con toses en tiempo real, por el grado de peligro que implicaba dicha situación, en principio no debería haber problema alguno, sin embargo, se debe hacer un análisis riguroso de ello. En el caso de los negativos y los ruidos o silencios funciona bien. Otra consideración



importante es mantener el sensor cercano a la boca al realizar la prueba de tos, no mayor a una distancia de 10cm.

### 5.1. Limitaciones

Las pruebas diagnósticas se usan como apoyo de la práctica clínica para sustentar las hipótesis diagnósticas; Alba & Villagra (2016) mencionan que las hipótesis diagnósticas que se originan de un proceso de análisis clínico llevan asociadas una probabilidad de ocurrencia denominada “*probabilidad pretest*”. Esta se logra a partir de un conjunto de elementos, entre ellos la experiencia del médico que hace el diagnóstico, el examen físico y elementos anamnésicos<sup>12</sup>. Si al juicio del médico la probabilidad de dicha prueba no supera el *umbral diagnóstico*<sup>13</sup>, se rechaza la hipótesis diagnóstica planteada inicialmente por el médico. Caso contrario, si la probabilidad supera un segundo umbral llamado *umbral terapéutico*<sup>14</sup>, el médico avala su hipótesis diagnóstica.

Según Bravo & Cruz (2015) los Likelihood ratio (razones de verosimilitud) positivos indican la probabilidad de un resultado específico en pacientes con la enfermedad versus los que no están enfermos; los valores que toman LHRP están en un rango  $(1, \infty)$ , mientras que los LHRN pertenecen a un rango de  $(0,1)$ . Si la razón de verosimilitud es igual a 1 la probabilidad del diagnóstico es la misma antes y después de aplicar la prueba y en este caso la prueba es inútil, porque no tiene capacidad discriminante. Cuanto más se aleje de 1 el valor de la razón de verosimilitud será menor la incertidumbre diagnóstica.

Ahora bien, el modelo 903 posee una sensibilidad del 100% es decir mayor al *Gold standar*.

12 Anamnesis: Historia clínica o conjunto de datos relativos a los antecedentes patológicos individuales y familiares del enfermo.

13 Aquel valor de probabilidad por debajo del cual podemos descartar la enfermedad de forma razonable, por lo que no tiene sentido realizar más pruebas.

14 Valor por encima del cual podemos confirmar razonablemente el diagnóstico e iniciar un tratamiento

Otro resultado importante es que NanNa no puede ser ocupada como prueba diagnóstica primaria para confirmar o descartar la presencia de COVID-19 debido a que en cuanto a Likelihood ratios negativos no pasó la prueba de calidad por  $-0.2$  y  $-0.8$  para el modelo 505 y 1310 respectivamente, contrariamente en Likelihood ratios positivos la clasificación de ambos modelos se posicionó como “Buena”, es decir que no puede ser usado después a la valoración médica ya que no arrojará un resultado confiable.

Según Kantor (2020) en una prueba diagnóstica con 90% de sensibilidad y 99% de especificidad el VAPP es de 95.7% y el VAPN de 97.5%. Estos datos se obtuvieron en el barrio Padre Mugica de Retiro en Argentina; para realizar dicha investigación se aplicó una prueba RT-PCR a un millón de habitantes; sin embargo, el resultado se relacionó con el grado de prevalencia de COVID-19 en la población.

Con base en estos resultados. Los modelos 1310, 505 y 903 poseen respectivamente 8.7%, 9.7 % y 41.7% menos exactitud que la *Gold standar* en VAPP y para el VAPN de 37.5, 33.5 y 97.5 %, en el mismo orden.

## 6. CONCLUSIONES

En conclusión, esta aplicación no puede ser una solución exacta al diagnóstico del COVID-19 debido a que los factores que influyen en la detección certera del virus sobrepasan la capacidad de dicha aplicación. Sin embargo, puede dar en buena medida una aproximación al correcto diagnóstico del COVID-19 e incluso ser una precursora de una prueba mejor elaborada, cumpliendo así con nuestros objetivos planteados; es una aplicación mejorable, lo que aumenta el grado de confiabilidad en el sistema. No se debe utilizar como método primario de detección debido a los resultados arrojados, quizá con la mejora pueda hallarse un sistema más confiable. Respecto al costo, en primera instancia es una prueba “cara”, sin embargo, sabiendo que es reutilizable una infinidad de veces se reduce el coste y aumenta el beneficio en gran medida, cumpliendo también con el objetivo de que pueda ser adquirida por la sociedad mexicana de bajos recursos.

El mayor indicador de COVID-19 positivo es el sensor de tos (siempre y cuando esta se presente en el posible portador) debido a que con la evolución de la enfermedad la situación respecto a sintomatología cambió. En personas asintomáticas el posible uso como método de diagnóstico coadyuvante pierde sentido, es decir que la prueba es completamente inválida para personas asintomáticas. Respecto a las probabilidades LHRN de acuerdo con los criterios descritos por Bravo & Cruz, el modelo 903 no tiene capacidad discriminante. Como resultado extra se pueden plantear la posible creación pruebas diagnósticas de enfermedades respiratorias (diversas) con base en algoritmos de Deep Learning mediante el análisis de tos.

**El presente trabajo logró el objetivo proporcionar una aplicación móvil basada en una red neuronal para la detección de COVID-19 como coadyuvante a las pruebas diagnósticas comerciales.**

Es decir

- Se desarrolló una aplicación móvil capaz de detectar y monitorear COVID-19.
- Se implementó una forma de que dicha aplicación sea lo menos costosa y pueda ser accesible para la población general, principalmente a las personas de bajos recursos.
- Se considera que la eficacia de dicha aplicación es aceptable de acuerdo con los resultados experimentales obtenidos.
- Se averiguó la sensibilidad y especificidad de un modelo de detección automática de sonido como prueba diagnóstica.
- Y sobre todo se creó un antecedente de detección de sonido como prueba diagnóstica para COVID-19 e incluso otras enfermedades del sistema respiratorio.

## 7. APÉNDICES

Nota: Se muestran solamente los códigos principales.

### 7.1. Apéndice A. Código del sensor de temperatura y tos

```
1 #include <ArduinoBLE.h>
2 #define EIDSP_QUANTIZE_FILTERBANK 0
3 #include <Wire.h>
4 #include <Adafruit_MLX90614.h>
5 #include <PDM.h>
6 #include <a505_inferencing.h>
7
8 BLEService lecturaService("19B10000-E8F2-537E-4F6C-D104768A1214");
9 //UUID del arduino
10
11 BLEStringCharacteristic readCharacteristic("19B10001-E8F2-537E-4F6C-
12 D104768A1214", BLERead | BLEWrite | BLENotify, 8); //Char UUID
13
14 float temperatura;
15 float temperatura_limite=38.0;
16 Adafruit_MLX90614 mlx = Adafruit_MLX90614();
17
18 float valor [3];
19
20 bool tos_covpos;
21 bool tos_covneg;
22 bool ruido;
23 bool ambiguo;
24 bool temp_covpos;
25 bool temp_covneg;
26 bool neg;
27 bool bandera1=0;
28 bool bandera2=0;
29 float oldtemp = 0;
30 String a;
31 String b,c,g;
32 typedef struct {
33     int16_t *buffer;
34     uint8_t buf_ready;
35     uint32_t buf_count;
36     uint32_t n_samples;
37 } inference_t;
38
39 static inference_t inference;
40 static signed short sampleBuffer[2048];
41 static bool debug_nn = false;
42
43
44 void setup()
```

```

45 {
46     Serial.begin(115200);
47     if (!BLE.begin()) {
48         Serial.println("BLE falló");
49     }
50 }
51
52 mlx.begin(); //se inicia el sensor
53
54     pinMode (5, INPUT);
55     pinMode (3, INPUT);
56
57 // Nombre del Blue
58     BLE.setLocalName("BLE GSM");
59     BLE.setAdvertisedService(lecturaService);
60 // Char del servicio
61     lecturaService.addCharacteristic(readCharacteristic);
62
63     // agregar servicio
64     BLE.addService(lecturaService);
65
66 //anuncio del bluetooth en otros dispositivos
67     BLE.advertise();
68     BLE.setConnectable(true);
69     //a=(String)(oldtemp);
70 //readCharacteristic.writeValue(a); // set initial value for this
71 characteristic
72
73     Serial.println("Edge Impulse Inferencing Demo");
74
75     ei_printf("Inferencing settings:\n");
76     ei_printf("\tInterval: %.2f ms.\n", (float)EI_CLASSIFIER_INTER-
77 VAL_MS);
78     ei_printf("\tFrame size: %d\n", EI_CLASSIFIER_DSP_IN-
79 PUT_FRAME_SIZE);
80     ei_printf("\tSample length: %d ms.\n", EI_CLASSIFIER_RAW_SAM-
81 PLE_COUNT / 16);
82     ei_printf("\tNo. of classes: %d\n", sizeof(ei_classifier_inferen-
83 cing_categories) / sizeof(ei_classifier_inferencing_categories[0]));
84
85     if (microphone_inference_start(EI_CLASSIFIER_RAW_SAMPLE_COUNT) ==
86 false) {
87         ei_printf("ERR: Failed to setup audio sampling\r\n");
88         return;
89     }
90 }
91
92 void loop()
93 {
94     BLE.poll();

```

```

95 BLEDevice central = BLE.central();
96
97 //Serial.println(readCharacteristic.value());
98 if ((digitalRead(5)==0)|| (c=="a")){
99     c="y";
100     while (digitalRead(5)==0);
101
102     monitorear_temperatura();
103     delay (500);
104 //ei_printf("b1 %d %d\n", bandera1, bandera2);
105 }
106
107
108 if ((digitalRead (3)==0)|| (c=="c")) {
109     c="y";
110     while (digitalRead(3)==0);
111
112     ei_printf("Grabando...\n");
113
114     bool m = microphone_inference_record();
115     if (!m) {
116         ei_printf("Error al grabar. Vuelva a grabar su tos\n");
117         return;
118     }
119
120     ei_printf("\n");
121     ei_printf("\n");
122
123     signal_t signal;
124     signal.total_length = EI_CLASSIFIER_RAW_SAMPLE_COUNT;
125     signal.get_data = &microphone_audio_signal_get_data;
126     ei_impulse_result_t result = { 0 };
127
128     EI_IMPULSE_ERROR r = run_classifier(&signal, &result, debug_nn);
129     if (r != EI_IMPULSE_OK) {
130         ei_printf("ERR: error al clasificar (%d)\n", r);
131         return;
132     }
133
134     ei_printf(": \n");
135     for (size_t ix = 0; ix < EI_CLASSIFIER_LABEL_COUNT; ix++) {
136 valor[ix]=result.classification[ix].value;
137     }
138
139 tos_covpos=((valor[0]>valor[1])&& (valor[0]>valor[2]));
140 tos_covneg=((valor[1]>valor[0])&& (valor[1]>valor[2]));
141 ruido=((valor[2]>valor[0])&& (valor[2]>valor[1]));
142 ambiguo=((tos_covpos==0)&&(tos_covneg==0)&& (ruido==0));
143 Serial.println (tos_covpos);
144 Serial.println (tos_covneg);

```

```

145 Serial.println (ruido);
146 Serial.println (ambiguo);
147
148 ei_printf("Analizando tos\n");
149 bandera1=1;
150
151 #if EI_CLASSIFIER_HAS_ANOMALY == 1
152     ei_printf("    Anomalía: %.3f\n", result.anomaly);
153 #endif
154
155 ei_printf("\n");
156 }
157
158 if (readCharacteristic.value() != b) {
159     b=readCharacteristic.value();
160     c=b;
161     //Serial.println(b);
162     delay(100);
163 }
164
165 if ((bandera1==1) && (bandera2==1)) {
166
167     diagnostico();
168     a=(String) (temperatura);
169 //Serial.println(readCharacteristic.value());
170 readCharacteristic.writeValue(a);
171 delay (1000);
172 readCharacteristic.writeValue(g);
173 delay (1000);
174 b=readCharacteristic.value();
175 delay (1000);
176 c="y";
177 }
178 }
179 void monitorear_temperatura() {
180
181     temperatura= mlx.readObjectTempC();
182
183 temp_covpos=(temperatura>=temperatura_limite);
184 temp_covneg=(temperatura<temperatura_limite);
185 neg=((temperatura>5.0) && (temperatura<temperatura_limite));
186
187 Serial.print("Midiendo temperatura...\n");
188
189 bandera2=1;
190
191 }
192 void diagnostico() {
193 Serial.println("revisando");
194

```



```

195  if ((tos_covpos)&&(temp_covpos)){ ei_printf("Tienes fiebre\n");
196      Serial.print(temperatura);
197      g="h";
198  ei_printf("°C");
199  ei_printf(";Tu temperatura es potencialmente riesgosa!, monito-
200  reala");
201      ei_printf("COVID positivo (exactitud mayor al 72%)\n");
202  }
203  if ((tos_covpos)&&(neg)){ ei_printf("Temperatura normal\n");
204      Serial.print(temperatura);
205      g="i";
206      ei_printf("°C");
207      ei_printf("COVID positivo (exactitud del 72%)\n");
208  }
209
210  if ((tos_covneg)&&(neg)){ ei_printf("Temperatura normal\n");
211      Serial.print(temperatura);
212      g="j";
213      ei_printf(" COVID negativo (exactitud del 72%)\n");
214  }
215  if ((tos_covneg)&&(temp_covpos)){ ei_printf("Tienes fiebre\n");
216      Serial.print(temperatura);
217      g="k";
218      ei_printf("°C");
219      ei_printf(";Tu temperatura es potencialmente riesgosa!, monito-
220  reala");
221
222      ei_printf(" COVID negativo (exactitud ligeramente menor al
223  72%)\n");
224  }
225
226  if ((valor[2]>valor[0])&& (valor[2]>valor[1])){ ei_printf("Ruido,
227  por favor vuelva a realizar la prueba.\n");
228      g="l";
229  }
230
231  if (ambiguo){ ei_printf("Valor ambiguo, por favor repita la
232  prueba.\n");
233      g="m";
234  }
235  banderal=0;
236  bandera2=0;
237  }
238  void ei_printf(const char *format, ...) {
239      static char print_buf[1024] = { 0 };
240
241      va_list args;
242      va_start(args, format);
243      int r = vsnprintf(print_buf, sizeof(print_buf), format, args);
244      va_end(args);

```

```

245
246     if (r > 0) {
247         Serial.write(print_buf);
248     }
249 }
250
251 static void pdm_data_ready_inference_callback(void)
252 {
253     int bytesAvailable = PDM.available();
254
255     // read into the sample buffer
256     int bytesRead = PDM.read((char *)&sampleBuffer[0], bytesAvaila-
257 ble);
258
259     if (inference.buf_ready == 0) {
260         for(int i = 0; i < bytesRead>>1; i++) {
261             inference.buffer[inference.buf_count++] = sampleBuf-
262 fer[i];
263
264             if(inference.buf_count >= inference.n_samples) {
265                 inference.buf_count = 0;
266                 inference.buf_ready = 1;
267                 break;
268             }
269         }
270     }
271 }
272
273 static bool microphone_inference_start(uint32_t n_samples)
274 {
275     inference.buffer = (int16_t *)malloc(n_samples * si-
276 zeof(int16_t));
277
278     if(inference.buffer == NULL) {
279         return false;
280     }
281
282     inference.buf_count = 0;
283     inference.n_samples = n_samples;
284     inference.buf_ready = 0;
285
286     // configure the data receive callback
287     PDM.onReceive(&pdm_data_ready_inference_callback);
288
289     PDM.setBufferSize(4096);
290
291     if (!PDM.begin(1, EI_CLASSIFIER_FREQUENCY)) {
292         ei_printf("Failed to start PDM!");
293         microphone_inference_end();
294

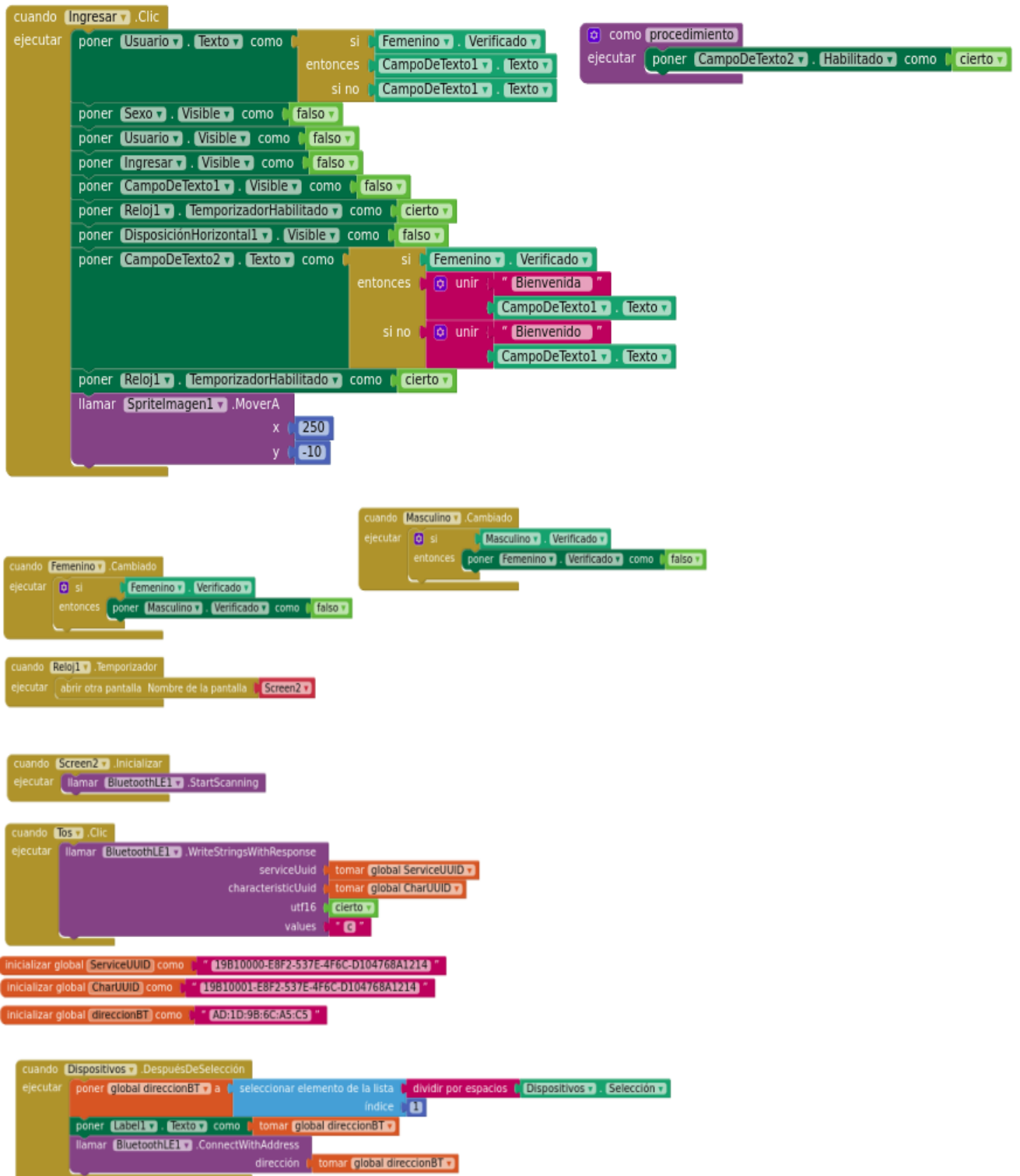
```

```

295     return false;
296 }
297
298 // set the gain, defaults to 20
299 PDM.setGain(127);
300
301 return true;
302 }
303
304
305 static bool microphone_inference_record(void)
306 {
307     inference.buf_ready = 0;
308     inference.buf_count = 0;
309
310     while(inference.buf_ready == 0) {
311         delay(10);
312     }
313
314     return true;
315 }
316
317
318 static int microphone_audio_signal_get_data(size_t offset, size_t
319 length, float *out_ptr)
320 {
321     numpy::int16_to_float(&inference.buffer[offset], out_ptr, len-
322 gth);
323
324     return 0;
325 }
326
327 static void microphone_inference_end(void)
328 {
329     PDM.end();
330     free(inference.buffer);
331
332 }
333
334 #if !defined(EI_CLASSIFIER_SENSOR) || EI_CLASSIFIER_SENSOR !=
335 EI_CLASSIFIER_SENSOR_MICROPHONE
336 #error "Invalid model for current sensor."
337 #endif

```

## 7.2. Apéndice B. Bloques de la APP.



```
cuando BluetoothLE1 .DeviceFound
ejecutar poner Dispositivos . ElementosDesdeCadena como BluetoothLE1 . DeviceList
```

```
cuando BluetoothLE1 .Connected
ejecutar si BluetoothLE1 . IsDeviceConnected
entonces llamar BluetoothLE1 .RegisterForStrings
serviceUuid tomar global ServiceUUID
characteristicUuid tomar global CharUUID
utf16 falso
poner Dispositivos . Visible como falso
llamar TextoAVoz1 .Hablar
mensaje " Conectado "
sino llamar TextoAVoz1 .Hablar
mensaje " Desconectado "
poner BluetoothLE1 . AutoReconnect como BluetoothLE1 . AutoReconnect
```

```
cuando Temp .Clic
ejecutar llamar BluetoothLE1 .WriteStringsWithResponse
serviceUuid tomar global ServiceUUID
characteristicUuid tomar global CharUUID
utf16 cierto
values " a "
```

```
cuando siguiente .Clic
ejecutar abrir otra pantalla Nombre de la pantalla Screen3
```

```

inicializar global m como " "

cuando BluetoothLE1 .StringsReceived
  serviceUuid  characteristicUuid  stringValue
  ejecutar
    poner global m a tomar stringValue
    si
      contiene texto tomar global m = cierto
      cadena "h"
    entonces poner Etiqueta3 . Texto como unir " ¡Tienes fiebre!, monitorea tu temperatura. "
      " COVID positivo (exactitud mayor al 71%) "
    si no, si
      contiene texto tomar global m = cierto
      cadena "i"
    entonces poner Etiqueta3 . Texto como unir " Tu temperatura es normal "
      " COVID positivo ( exactitud del 72%) "
    si no, si
      contiene texto tomar global m = cierto
      cadena "j"
    entonces poner Etiqueta3 . Texto como unir " Tu temperatura es normal "
      " COVID negativo (exactitud mayor al 72%) "
    si no, si
      contiene texto tomar global m = cierto
      cadena "k"
    entonces poner Etiqueta3 . Texto como unir " ¡Tienes fiebre!, monitorea tu temperatura "
      " COVID negativo (exactitud del 71%) "
    si no, si
      contiene texto tomar global m = cierto
      cadena "l"
    entonces poner Etiqueta3 . Texto como " Ruido "
    si no, si
      contiene texto tomar global m = cierto
      cadena "m"
    entonces poner Etiqueta3 . Texto como " Ambiguo "
    si no, si
      contiene texto tomar global m = cierto
      cadena "a"
    entonces
    si no, si
      contiene texto tomar global m = cierto
      cadena "c"
    entonces
  sino
    poner Label1 . Texto como tomar global m

```

### 7.3. Apéndice C. Código de las tablas.

- **Espectrograma**

```
1 \documentclass{article}
2 \usepackage[utf8]{inputenc}
3 \usepackage[spanish]{babel}
4 \usepackage{amsmath}
5 \usepackage{amsfonts}
6 \usepackage{booktabs}
7 \usepackage{array}
8 \title{Tabla 1}
9 \begin{document}
10 \begin{tabular}{*{8}{c}}
11
12 \hline
13 & & & Espectrograma & & & & \\
14 \toprule
15 \toprule
16 & Ventana & Ventana & Longitud & Paso & Frecuencia & Ruido de & \\
17 Frecuencia & de & de & de & por & de & piso & \\
18 (Hz) & tiempo(ms) & aumento(ms) & cuadro(ms) & & fotograma(f/s) & & \\
19 & bandas(Hz) & & (dB) & & & & \\
20 \midrule
21 16k & 1000 & 500 & 200 & 100 & 128 & -52 & \\
22 \bottomrule
23 \end{tabular}
24
25 \end{document}
26 \toprule
27 Ventana de aumento \\
28 \midrule
29 16k Hz & 1000 ms & 500 ms & \\
30 \bottomrule
```

- **Media de las clasificaciones de muestra de los tres modelos**

```

1 \documentclass{article}
2 \usepackage{color}
3 \usepackage{graphicx}
4 \usepackage{epsfig}
5 \usepackage{multirow}
6 \usepackage{colortbl}
7 \usepackage[table,xcdraw]{xcolor} % para el color de la primera fi-la
8 de la tabla
9 \definecolor{lightgray}{gray}{0.9}
10 \begin{document}
11 \vspace{5pt}
12 \begin{table}[h!]
13 \centering
14 \begin{tabular}{||c | c | c | c||}
15 \hline
16 \hline
17 \multicolumn{4}{|c|}{\cellcolor[HTML]{CCCCCC}Media de las clasifi-
18 caciones de muestra de los tres modelos }\\
19 \hline\hline
20 Verdadero Negativo & Falso Positivo & Verdadero Positivo & Falso Ne-
21 gativo \\ [0.5ex]
22 \hline\hline
23
24 \multicolumn{4}{|c|}{\cellcolor[HTML]{99FFCC} Modelo 505} \\
25 \hline\hline
26 84.4 & 7.8 & 49.6 & 46.6 \\
27 \hline\hline
28 \multicolumn{4}{|c|}{\cellcolor[HTML]{99FFCC}Modelo 1310} \\
29 \hline\hline
30 80 & 6.2 & 42.8 & 51.4 \\
31 \hline\hline
32 \multicolumn{4}{|c|}{\cellcolor[HTML]{99FFCC}Modelo 903} \\
33 \hline\hline
34 0 & 72.6 & 87.8 & 0 \\ [1ex]
35 \hline
36 \end{tabular}
37 \end{table}
38 \end{document}

```



- **Media de las probabilidades para VAPP, VAPN, LHRP, LHRN y EX**

```

1
2 \documentclass{article}
3 \usepackage{color}
4 \usepackage{graphicx}
5 \usepackage{epsfig}
6 \usepackage{multirow}
7 \usepackage{colortbl}
8 \usepackage[table,xcdraw]{xcolor} % para el color de la primera fila
9 de la tabla
5 \definecolor{lightgray}{gray}{0.9}
6
7 \begin{document}
8 \begin{table}[h!]
9 \begin{tabular}{|r|r|r|r|r|}
10 \hline
11 \hline
12 \multicolumn{5}{|c|}{\cellcolor[HTML]{CCCCCC}Media de las probabilidad-
13 des para VAPP, VAPN, LHRP, LHRN y EX}
14 \\\ \hline\hline
15 \textbf{Valor Predictivo} & \textbf{Valor Predictivo} & \textbf{Like-
16 lihood ratio} & &
17 \textbf{Likelihood ratio} & \textbf{Exactitud} \\\
18 \textbf{Positivo} \hspace{0.5 cm} & \textbf{Negativo} \hspace{0.5
19 cm} & \textbf{Positivo} \hspace{0.5 cm} & \textbf{Negativo}
20 \hspace{0.5 cm} & \\\
21 \hline
22 \multicolumn{5}{|c|}{\cellcolor[HTML]{CCFFCC} Modelo 505} \\\
23 \hline
24 0.86 & 0.64 & & 6.12 & 0.52 & 0.71 \\\
25 \hline
26 \multicolumn{5}{|c|}{\cellcolor[HTML]{CCFFCC}Modelo 1310} \\\
27 0.87 & 0.60 & & 6.31 & 0.58 & 0.68\\
28 \hline
29 \multicolumn{5}{|c|}{\cellcolor[HTML]{CCFFCC}Modelo 903} \\\
30 \hline
31 0.54 & 0 & & 1 & 0 & 0.54\\
32 \hline
33 \end{tabular}
34 \end{table}
35 \end{document}

```

- Rangos de Likelihood

```

1 \documentclass{article}
2 \usepackage{color}
3 \usepackage{graphicx}
4 \usepackage{epsfig}
5 \usepackage{multirow}
6 \usepackage{colortbl}
7 \vspace{5pt}
8 \usepackage[table,xcdraw]{xcolor} % para el color de la primera fila
9 de la tabla
10 \definecolor{lightgray}{gray}{0.9}
11 \begin{document}
12 \begin{table}[h!]
13 \begin{tabular}{|c|c|c|}
14 \hline
15 \hline
16 \multicolumn{3}{|c|}{\cellcolor[HTML]{CCCCCC} Rangos Likelihood Ratios
17 y su utilidad} \\ \hline \hline
18 Likelihood Ratio Positivo&Likelihood Ratio Negativo&Clasificación \\
19 [0.5ex]
20 \hline \hline
21 10&  $\$ < \$0.1$  & Altamente relevante \\ \hline
22 5-10& 0.1-0.2 & Buena \\ \hline
23 2-5& 0.5-0.2 & Regular \\ \hline
24  $\$ < \$2$  &  $\$ > 0.5$  & Mala \\ \hline
25 \end{tabular}
26 \end{table}
27 \end{document}

```

- Costo del material

```

1 \documentclass{article}
2 \usepackage{multirow}
3 \begin{document}
4 \begin{tabular}{|p{3cm}|p{3cm}|p{3cm}|p{3cm}| }
5 \hline
6 \multicolumn{4}{|c|}{Costo por material.} \\
7 \hline
8 &\hspace{0.8cm} Precio (\$) &\hspace{0.8cm} Numero &\hspace{0.5cm}
9 Precio (\$)\\
10 \hspace{0.8 cm} Material &\hspace{1.2cm} por &\hspace{1.2cm} de &
11 \hspace{0.8cm} total \\
12 &\hspace{1cm} unidad &\hspace{1cm} piezas & \\
13 \hline
14 Sensor MLX90614ESF-DCC &\hspace{1.2cm} 590 &\hspace{1.3cm} 1 &
15 \hspace{1.2cm} 590\\
16 \hline
17 Arduino Nano 33 BLE Sense&\hspace{1.1cm} 1835 &\hspace{1.3cm} 1 &
18 \hspace{1.1cm} 1835\\
19 \hline
20 Switch Push Buton&\hspace{1.1cm} 1.50 &\hspace{1.3cm} 2&
21 \hspace{1.3cm} 3\\
22 \hline
23 Resistencia 10k $\Omega$ &\hspace{1.3cm} 1 &\hspace{1.3cm} 2&
24 \hspace{1.3cm} 2\\
25 \hline
26 Resistencia 4.7k $\Omega$ &\hspace{1.3cm} 1 &\hspace{1.3cm} 2 &
27 \hspace{1.3cm} 2\\
28 \hline
29 Protoboard &\hspace{1.2cm} 55 &\hspace{1.3cm} 1 &
30 \hspace{1.2cm} 55\\
31 \hline
32 Alambre estañado&\hspace{1cm} \$6/m. &\hspace{1.2cm} 2 m. &
33 \hspace{1.2cm} 12\\
34 \hline
35 Cables Dupont &\hspace{1.3cm} 2 &\hspace{1.3cm} 6 &\hspace{1.2cm}
36 12\\
37 \hline
38 Cautín (kit) &\hspace{1.2cm} 150&\hspace{1.3cm} 1&\hspace{1.1cm}
39 150\\
40 \hline
41 Pasta para soldar (kit)&\hspace{1cm} (kit)&\hspace{1.3cm} 1&
42 \hspace{1cm} (kit)\\
43 \hline
44 Soldadura de estaño (kit)&\hspace{1cm} (kit)&\hspace{1.3cm} 1&
45 \hspace{1cm} (kit)\\
46 \hline
47 Cable USB macho A mini USB
48 &\hspace{1.3cm}50&\hspace{1.4cm}1&\hspace{1.3cm}50\\
49 \hline
50 Placa de cobre &\hspace{1.3cm}10&\hspace{1.4cm}1&\hspace{1.3cm}10

```

```
51 \hline
52 \textbf{Precio Total} &&&\hspace{0.7 cm} \textbf{\$ \cal$ 2721.00}\} \\
53 \hline
54 \end{tabular}
55 \end{document}
```

- **Número de fragmentos**

```
1 \documentclass{article}
2 \usepackage[utf8]{inputenc}
3 \usepackage[spanish]{babel}
4 \usepackage{amsmath}
5 \usepackage{amsfonts}
6 \usepackage{booktabs}
7 \usepackage{array}
8 \title{Tabla 1}
9 \begin{document}
10 \begin{tabular}{*{3}{c}}
11 \hline
12 & Modelo 505, Tiempo de recolección. & \\
13 \toprule
14 \toprule
15 Covid Positivos & Covid Negativos & Ruido \\
16 (s) & (s) & (s) \\
17 \midrule
18 82 & 82 & 80 \\
19 \bottomrule
20 \end{tabular}
21
22 \end{document}
23
24
25 \toprule
26 Ventana de aumento \\
27 \midrule
28 16k Hz & 1000 ms & 500 ms \\
29 \bottomrule
30
```

- **Medias VP, FP, VN y FN**

```

1 \documentclass{article}
2 \usepackage{color}
3 \usepackage{graphicx}
4 \usepackage{epsfig}
5 \usepackage{multirow}
6 \usepackage{colortbl}
7 \vspace{5pt}
8 \usepackage[table,xcdraw]{xcolor} % para el color de la primera fila
9 de la tabla
10 \definecolor{lightgray}{gray}{0.9}
11 \begin{document}
12 \begin{table}[h!]
13 \begin{tabular}{|c|c|c|c|}
14 \hline
15 \hline
16 \multicolumn{4}{|c|}{\cellcolor[HTML]{CCCCCC}Medias para VN, FP,} \\
17 \multicolumn{4}{|c|}{\cellcolor[HTML]{CCCCCC}VP y FN} \\
18 \hline\hline
19 VN & FP & VP & FN \\
20 \multicolumn{4}{|c|}{\cellcolor[HTML]{009999}Modelo 505} \\
21 \hline
22 84.8 & 7.8 & 49.6 & 46.6 \\
23 \hline
24 \multicolumn{4}{|c|}{\cellcolor[HTML]{009999}Modelo 1310} \\
25 \hline
26 80 & 6.2 & 42.8 & 51.4 \\
27 \hline
28 \multicolumn{4}{|c|}{\cellcolor[HTML]{009999}Modelo 903} \\
29 \hline
30 0 & 72.6 & 87.8 & 0 \\
31 \hline
32 \end{tabular}
33 \end{table}
34 \end{document}

```

## • Sensibilidad y especificidad

```
1 \documentclass{article}
2 \usepackage{color}
3 \usepackage{graphicx}
4 \usepackage{epsfig}
5 \usepackage{multirow}
6 \usepackage{colortbl}
7 \usepackage[table,xcdraw]{xcolor} % para el color de la primera fila
8 de la tabla
9 \definecolor{lightgray}{gray}{0.9}
10 \begin{document}
11 \vspace{5pt}
12 \begin{table}[h!]
13 \centering
14 \begin{tabular}{||c|c||}
15 \hline
16 \hline
17 \multicolumn{2}{c}{\cellcolor[HTML]{CCCCCC} Sensibilidad y Especifici-
18 cidad }\\
19 \multicolumn{2}{c}{\cellcolor[HTML]{CCCCCC} relación con el número
20 de}\\
21 \multicolumn{2}{c}{\cellcolor[HTML]{CCCCCC} pacientes con y sin CO-
22 VID-19 }\\
23 \hline\hline
24 Sensibilidad & Especificidad \\ [0.5ex]
25 \hline\hline
26 \multicolumn{2}{|c|}{\cellcolor[HTML]{FFCC99} Modelo 505} \\
27 \hline\hline
28 0.51 & 0.91 \\
29 \hline\hline
30 \multicolumn{2}{|c|}{\cellcolor[HTML]{FFCC99}Modelo 1310} \\
31 \hline\hline
32 0.45 & 0.92 \\
33 \hline\hline
34 \multicolumn{2}{|c|}{\cellcolor[HTML]{FFCC99}Modelo 903} \\
35 \hline\hline
36 1 & 0 \\ [1ex]
37 \hline
38 \end{tabular}
39 \end{table}
40 \end{document}
```

## 7.4. Apéndice D. Código de los histogramas

### · Clasificaciones preliminares del modelo 505"

```
1 void clas505 () {
2 TCanvas *c = new TCanvas("c","Clasificaciones preliminares del modelo
3 505",100,100,900,500);
4   c->SetGrid();
5
6   const int bien1 = 118, indet1 = 39, mal1 = 42, bien2 = 120, indet2
7 = 25, mal2 = 54, bien3 = 107, indet3 = 47, mal3 = 45, bien4 = 117, in-
8 det4 = 43, mal4 = 39, bien5 = 123, indet5 = 25, mal5 = 51;
9   const char *clas[3] = {"Bien", "Indeterminado", "Mal"}, *clas2[3]
10 = {"Bien", "Indeterminado", "Mal"}, *clas3[3] = {"Bien", "Indetermi-
11 nado", "mal"}, *clas4[3] = {"Bien", "Indeterminado", "Mal"},
12 *clas5[3] = {"Bien", "Indeterminado", "Mal"};
13
14   TH1F *h = new TH1F("h","Clasificaciones preliminares del modelo
15 505 ",0, 50, 0);
16
17   for (Int_t i=0;i<bien1;i++) {
18     h->Fill(clas[0],1);
19   }
20   for (Int_t i=0;i<indet1;i++) {
21     h->Fill(clas[1],1);
22   }
23
24   for (Int_t i=0;i<mal1;i++) {
25     h->Fill(clas[2],1);
26   }
27   h->GetXaxis()->SetTitle("Clasificaciones de las muestras");
28   h->GetYaxis()->SetTitle("Fragmentos de Tos (FT)");
29   h->SetMarkerColor(20);
30   h->SetLineWidth(2);
31   h->SetMarkerStyle(22);
32   h->SetMarkerSize(2);
33   h->SetAxisRange(0, 130, "x");
34   h->SetLineColorAlpha(20, 0.35);
35   h->SetMaximum(130);
36   h->SetMinimum(0);
37   TH1F *h2 = new TH1F("h2", "Clasificaciones preliminares del modelo
38 505", 0, 0, 0);
39
40
41   for (Int_t i=0;i<bien2;i++) {
42     h2->Fill(clas2[0],1);
```



```

43 }
44 for (Int_t i=0;i<indet2;i++) {
45     h2->Fill(clas2[1],1);
46 }
47 for (Int_t i=0;i<mal2;i++) {
48     h2->Fill(clas2[2],1);
49 }
50
51 h2->SetMarkerColor(34);
52 h2->SetMarkerStyle(22);
53 h2->SetLineWidth(2);
54 h2->SetMarkerSize(2);
55 h2->SetAxisRange(0, 130, "x");
56 h2->SetLineColorAlpha(34, 0.35);
57
58 TH1F *h3 = new TH1F("h3","Clasificaciones preliminares del modelo
59 505",0, 0, 0);
60
61 for (Int_t i=0;i<bien3;i++) {
62     h3->Fill(clas3[0],1);
63 }
64 for (Int_t i=0;i<indet3;i++) {
65     h3->Fill(clas3[1],1);
66 }
67 for (Int_t i=0;i<mal3;i++) {
68     h3->Fill(clas3[2],1);
69 }
70 h3->SetMarkerColor(49);
71 h3->SetMarkerStyle(22);
72 h3->SetMarkerSize(2.0);
73 h3->SetLineWidth(2);
74 h3->SetAxisRange(0, 130, "Y");
75 h3->SetLineColorAlpha(49, 0.35);
76
77 TH1F *h4 = new TH1F("h4","Clasificaciones preliminares del modelo
78 505",0, 0, 0);
79
80 for (Int_t i=0;i<bien4;i++) {
81     h4->Fill(clas4[0],1);
82 }
83 for (Int_t i=0;i<indet4;i++) {
84     h4->Fill(clas4[1],1);
85 }
86 for (Int_t i=0;i<mal4;i++) {
87     h4->Fill(clas4[2],1);
88 }
89 h4->SetMarkerColor(40);
90 h4->SetMarkerStyle(22);
91 h4->SetMarkerSize(2.0);
92 h4->SetLineWidth(2);

```

```

93  h4->SetAxisRange(0, 130, "Y");
94  h4->SetLineColorAlpha(40, 0.35);
95  TH1F *h5 = new TH1F("h5", "Clasificaciones preliminares del modelo
96  505", 0, 0, 0);
97
98  for (Int_t i=0; i<bien5; i++) {
99      h5->Fill(clas5[0], 1);
100 }
101 for (Int_t i=0; i<indet5; i++) {
102     h5->Fill(clas5[1], 1);
103 }
104
105 for (Int_t i=0; i<mal5; i++) {
106     h5->Fill(clas5[2], 1);
107 }
108
109 h5->SetMarkerColor(45);
110 h5->SetMarkerStyle(22);
111 h5->SetMarkerSize(2.0);
112 h5->SetLineWidth(2);
113 h5->SetAxisRange(0, 130, "Y");
114 h5->SetLineColorAlpha(45, 0.35);
115 gStyle->SetOptStat(11);
116 h->Draw("PL");
117 h2->Draw("PL SAME");
118 h3->Draw("PL SAME");
119 h4->Draw("PL SAME");
120 h5->Draw("PL SAME");
121
122 auto legend = new TLegend(1.0, 1.0, 1.0, 1.0);
123 legend->AddEntry(h, "Muestra 1", "p");
124 legend->AddEntry(h2, "Muestra 2", "p");
125 legend->AddEntry(h3, "Muestra 3", "p");
126 legend->AddEntry(h4, "Muestra 4", "p");
127 legend->AddEntry(h5, "Muestra 5", "p");
128 legend->SetFillStyle(0);
129 legend->SetBorderSize(0);
130 legend->Draw("");
131 c->SaveAs("lu.png");
132
    }

```

## • Contraste de muestras promedio filtradas del modelo 505

```
1 void filtrado505() {
2
3     TCanvas *c = new TCanvas("c","Contraste de muestras promedio fil-
4 tradas del modelo 505",10,10,900,500);
5     c->SetGrid();
6
7     auto mg = new TMultiGraph();
8
9     const int total = 188.8, bien = 134.4, mal = 54.4, nulas = 10.2;
10    const char *datos[2] = {"Muestras totales (188.8)","Bien clasifica-
11 das (134.4)"}, *datos2[1] = {"Mal clasificadas (54.4)"};
12
13    TH1F *h = new TH1F("h","Contraste de muestras promedio filtradas
14 del modelo 505",3, 0, 3);
15
16    h->GetXaxis()->SetTitle("Total de muestras");
17    h->GetYaxis()->SetTitle("FT nulos");
18    h->SetFillColor(47);
19    h->SetLineColor(1);
20    h->SetMaximum(203);
21    h->SetMinimum(0);
22
23    for (Int_t i=0;i<total;i++) {
24        h->Fill(datos[0],1);
25    }
26    for (Int_t i=0;i<bien;i++) {
27        h->Fill(datos[1],1);
28    }
29
30
31    for (Int_t i=0;i<mal;i++) {
32        h->Fill(datos2[0],1);}
33
34
35    TH1F *h2 = new TH1F("h2", "Histograma 2", 3, 0, 3);
36    h2->SetFillColor(49);
37    h2->SetLineColor(1);
38
39    for (Int_t i=0;i<nulas;i++) {
40        h2->Fill(datos2[0],1);
41    }
42
43    gStyle->SetOptStat(false);
44    h->Draw("same");
45    h2->Draw("same");
46
47    auto legend = new TLegend(1.0,1.0,1.0,1.0);
48    //legend->AddEntry(h,"Total", "f");
```

```
49 legend->AddEntry(h2,"Nulas=10.2", "f");
50 legend->SetFillStyle(0.2);
51 legend->SetBorderSize(4);
52 legend->Draw("");
53
c->SaveAs("filtrado505.png");

}
```

- **Pruebas promedio filtradas de los 3 modelos**

```
1 void filtrado123() {
2
3     TCanvas *c = new TCanvas("c","Clasificaciones preliminares de los 3
4 modelos",100,100,900,500);
5     c->SetGrid();
6
7     const int bien1 = 134.4, mal1 = 54.4 , null1 = 10.2 , bien2 = 86.8
8 , mal2 = 72.6 , null2 = 38.6, bien3 = 122.6 , mal3 = 57.6 , null3 =
9 18.6;
10    const char *clas[3] = {"Bien", "Mal", "Nulas"}, *clas2[3] =
11 {"Bien", "Mal", "Nulas"}, *clas3[3] = {"Bien", "mal", "Nulas"};
12
13    TH1F *h = new TH1F("h","Pruebas promedio filtradas de los 3 modelos
14 ",100, 50, 0);
15
16    for (Int_t i=0;i<bien1;i++) {
17        h->Fill(clas[0],1);
18    }
19    for (Int_t i=0;i<mal1;i++) {
20        h->Fill(clas[1],1);
21    }
22
23    for (Int_t i=0;i<null1;i++) {
24        h->Fill(clas[2],1);
25    }
26
27    h->SetMarkerColor(20);
28    h->SetLineWidth(2);
29    h->SetMarkerStyle(22);
30    h->SetMarkerSize(2);
31    h->SetAxisRange(0, 130, "x");
32    h->SetLineColorAlpha(20, 0.35);
33    // h->SetMaximum(130);
34    //h->SetMinimum(0);
35    TH1F *h2 = new TH1F("h2", "Clasificaciones preliminares del modelo
36 505", 0, 0, 0);
37
38
39    for (Int_t i=0;i<bien2;i++) {
40        h2->Fill(clas2[0],1);
41    }
42    for (Int_t i=0;i<mal2;i++) {
43        h2->Fill(clas2[1],1);
44    }
45    for (Int_t i=0;i<null2;i++) {
46        h2->Fill(clas2[2],1);
47    }
48
49    h2->SetMarkerColor(34);
```

```

50  h2->SetMarkerStyle(22);
51  h2->SetLineWidth(2);
52  h2->SetMarkerSize(2);
53  h2->SetAxisRange(0, 130, "x");
54  h2->SetLineColorAlpha(34, 0.35);
55
56  TH1F *h3 = new TH1F("h3", "Clasificaciones preliminares del modelo
57 505", 0, 0, 0);
58
59  for (Int_t i=0;i<bien3;i++) {
60      h3->Fill(clas3[0],1);
61  }
62  for (Int_t i=0;i<mal3;i++) {
63      h3->Fill(clas3[1],1);
64  }
65  for (Int_t i=0;i<null3;i++) {
66      h3->Fill(clas3[2],1);
67  }
68  h3->SetMarkerColor(49);
69  h3->SetMarkerStyle(22);
70  h3->SetMarkerSize(2.0);
71  h3->SetLineWidth(2);
72  h3->SetAxisRange(0, 130, "Y");
73  h3->SetLineColorAlpha(49, 0.35);
74
75  auto legend = new TLegend(1.0,1.0,1.0,1.0);
76  legend->AddEntry(h, "Modelo 505", "p");
77  legend->AddEntry(h2, "Modelo 903", "p");
78  legend->AddEntry(h3, "Modelo 1310", "p");
79
80  legend->SetFillStyle(0);
81  legend->SetBorderSize(0);
82  legend->Draw("");
83
84  c->SaveAs("123.png");
85
86 }

```

## • Pruebas totales anuladas del modelo 505

```
1 void depurar505() {
2
3     const int var1 =18, var2 = 11, var3 = 10, var4 = 9, var5= 3;
4
5     TCanvas *c = new TCanvas("c", "Graficas", 1200, 800);
6     c->SetGrid();
7     gStyle->SetOptStat(1111);
8
9
10    TH1I *h = new TH1I("Valores","Pruebas totales anuladas del modelo
11 505",3500,0,6);
12
13    for (Int_t i=0;i<var1;i++) {
14        h->Fill(1);
15    }
16    for (Int_t i=0;i<var2;i++) {
17        h->Fill(2); }
18    for (Int_t i=0;i<var3;i++) {
19        h->Fill(3);
20    }
21
22    for (Int_t i=0;i<var4;i++) {
23        h->Fill(4);
24    }
25
26    for (Int_t i=0;i<var5;i++) {
27        h->Fill(5);
28    }
29
30    h->GetXaxis()->SetTitle("Ciclo");
31    h->GetYaxis()->SetTitle("FT nullos");
32    h->SetMarkerColor(28);
33    h->SetMarkerStyle(3);
34    h->SetMarkerSize(3.0);
35    h->SetMaximum(20);
36    h->SetMinimum(0);
37
38
39    h->Draw("P");
40    c->SaveAs("depu.png");
41
42 }
```

## • Valores promedio del modelo 505

```
1 void promediosen505() {
2
3     TH1F *h = new TH1F("h", "Valores promedio finales del modelo
4 505", 4, 0, 3);
5
6     const int sen1 = 0.51, sen2 = 0.45, sen3 = 1, espe1 = 0.91, espe2 =
7 0.92, espe3 = 0, vapo1 = 0.86, vapo2 = 0.87, vapo3 = 0.54, vaneg1 =
8 0.64, vaneg2 = 0.60, vaneg3 = 0, exa1 = 0.71, exa2 = 0.68, exa3 = 0.54;
9     const char *datos[5] = {"Sensibilidad (0.51)", "Especificidad
10 (0.91)", "Valor predictivo positivo (0.86)", "Valor predictivo nega-
11 tivo (0.64)", "Exactitud (0.71)"};
12
13     TCanvas *c = new TCanvas("c", "Valores promedio finales del modelo
14 505", 10, 10, 900, 500);
15     c->SetGrid();
16     gStyle->SetOptStat(false);
17
18
19
20     for (Int_t i=0; i<=sen1; i++) {
21         h->Fill(datos[0], 0.51);
22         h->SetBinContent(4, 0.51);
23         h->SetMarkerColor(2);
24     }
25
26     for (Int_t i=0; i<=sen1; i++) {
27         h->Fill(datos[1], 0.91);
28         h->SetBinContent(4, 0.91);
29         h->SetMarkerColor(2);
30     }
31
32     for (Int_t i=0; i<=sen1; i++) {
33         h->Fill(datos[2], 0.86);
34         h->SetBinContent(4, 0.86);
35         h->SetMarkerColor(2);
36     }
37
38     for (Int_t i=0; i<=sen1; i++) {
39         h->Fill(datos[3], 0.64);
40         h->SetBinContent(4, 0.64);
41         h->SetMarkerColor(2);
42     }
43
44
45     for (Int_t i=0; i<=sen1; i++) {
46         h->Fill(datos[4], 0.71);
47         h->SetBinContent(4, 0.71);
48         h->SetMarkerColor(2);
```



```
49 }
50
51 h->GetXaxis()->SetTitle("Criterio evaluado");
52 h->GetYaxis()->SetTitle("Probabilidad");
53 h->SetLineWidth(3);
54 h->SetMarkerStyle(2);
55 h->SetMarkerSize(2.5);
56 h->SetMarkerColor(5);
57 h->SetAxisRange(0, 130, "x");
58 h->SetBinContent(4, 0.64);
59   h->SetLineColor(4);
60 h->SetMaximum(1.2);
61 h->SetMinimum(0);
62
63 h->Draw("PL");
64 c->SaveAs("promediosen505.png");
65
66 }
```

## • Resultados finales modelo 505

```
1 void final505() {
2     TCanvas *c = new TCanvas("c","Resultados finales modelo
3 505",100,100,900,500);
4     c->SetGrid();
5     gStyle->SetOptStat(0);
6
7     const int var1 = 87, var2 = 2, var3 = 44, var4 = 48, var5 = 79,
8 var6 = 11, var7 = 45, var8 = 53, var9 = 85, var10 = 8, var11 = 49,
9 var12 = 47, var13 = 84, var14 = 9, var15 = 58, var16 = 39, var17 =
10 89, var18 = 9, var19 = 52, var20 = 46;
11     const char *clas[4] = {"Verdaderos Negativos", "Falsos Positivos",
12 "Verdaderos Positivos", "Falsos Negativos"}, *clas2[4] = {"Verdaderos
13 Negativos", "Falsos Positivos", "Verdaderos Positivos", "Falsos Nega-
14 tivos"}, *clas3[4] = {"Verdaderos Negativos", "Falsos Positivos",
15 "Verdaderos Positivos", "Falsos Negativos"}, *clas4[4] = {"Verdaderos
16 Negativos", "Falsos Positivos", "Verdaderos Positivos", "Falsos Nega-
17 tivos"}, *clas5[5] = {"Verdaderos Negativos", "Falsos Positivos",
18 "Verdaderos Positivos", "Falsos Negativos"};
19
20 TH1F *h = new TH1F("h","Resultados finales modelo 505",180, 50, 0);
21
22     for (Int_t i=0;i<var1;i++) {
23         h->Fill(clas[0],1);
24     }
25     for (Int_t i=0;i<var2;i++) {
26         h->Fill(clas[1],1);
27     }
28
29     for (Int_t i=0;i<var3;i++) {
30         h->Fill(clas[2],1);
31     }
32
33     for (Int_t i=0;i<var4;i++) {
34         h->Fill(clas[3],1);
35     }
36     h->GetXaxis()->SetTitle("Valores predictivos");
37     h->GetYaxis()->SetTitle("FT");
38     h->SetMarkerColor(30);
39     h->SetLineWidth(2);
40     h->SetMarkerStyle(45);
41     h->SetMarkerSize(2);
42     h->SetAxisRange(0, 130, "x");
43     h->SetLineColorAlpha(30, 0.35);
44     h->SetMaximum(92);
45     h->SetMinimum(0);
46     TH1F *h2 = new TH1F("h2", "Resultados finales modelo 505", 0, 0,
47 0);
48     for (Int_t i=0;i<var5;i++) {
49         h2->Fill(clas2[0],1);
```

```

50  }
51  for (Int_t i=0;i<var6;i++) {
52    h2->Fill(clas2[1],1);
53  }
54  for (Int_t i=0;i<var7;i++) {
55    h2->Fill(clas2[2],1);
56  }
57  for (Int_t i=0;i<var8;i++) {
58    h2->Fill(clas2[3],1);
59  }
60
61  h2->SetMarkerColor(34);
62  h2->SetMarkerStyle(45);
63  h2->SetLineWidth(2);
64  h2->SetMarkerSize(2);
65  h2->SetAxisRange(0, 130, "x");
66  h2->SetLineColorAlpha(34, 0.35);
67
68  TH1F *h3 = new TH1F("h3","Resultados finales modelo 505",0, 0, 0);
69
70  for (Int_t i=0;i<var9;i++) {
71    h3->Fill(clas3[0],1);
72  }
73  for (Int_t i=0;i<var10;i++) {
74    h3->Fill(clas3[1],1);
75  }
76  for (Int_t i=0;i<var11;i++) {
77    h3->Fill(clas3[2],1);
78  }
79  for (Int_t i=0;i<var12;i++) {
80    h3->Fill(clas3[3],1);
81  }
82  h3->SetMarkerColor(49);
83  h3->SetMarkerStyle(45);
84  h3->SetMarkerSize(2.0);
85  h3->SetLineWidth(2);
86  h3->SetAxisRange(0, 130, "Y");
87  h3->SetLineColorAlpha(49, 0.35);
88
89 TH1F *h4 = new TH1F("h4","Resultados finales modelo 505",0, 0, 0);
90
91  for (Int_t i=0;i<var13;i++) {
92    h4->Fill(clas4[0],1);
93  }
94  for (Int_t i=0;i<var14;i++) {
95    h4->Fill(clas4[1],1);
96  }
97  for (Int_t i=0;i<var15;i++) {
98    h4->Fill(clas4[2],1);
99  }

```

```

100     for (Int_t i=0;i<var16;i++) {
101         h4->Fill(clas4[3],1);
102     }
103
104     h4->SetMarkerColor(20);
105     h4->SetMarkerStyle(45);
106     h4->SetMarkerSize(2.0);
107     h4->SetLineWidth(2);
108     h4->SetAxisRange(0, 130, "Y");
109     h4->SetLineColorAlpha(20, 0.35);
110
111 TH1F *h5 = new TH1F("h5","Resultados finales modelo 505",0, 0, 0);
112
113     for (Int_t i=0;i<var17;i++) {
114         h5->Fill(clas5[0],1);
115     }
116     for (Int_t i=0;i<var18;i++) {
117         h5->Fill(clas5[1],1);
118     }
119
120     for (Int_t i=0;i<var19;i++) {
121         h5->Fill(clas5[2],1);
122     }
123     for (Int_t i=0;i<var20;i++) {
124         h5->Fill(clas5[3],1);
125     }
126
127     h5->SetMarkerColor(45);
128     h5->SetMarkerStyle(45);
129     h5->SetMarkerSize(2.0);
130     h5->SetLineWidth(2);
131     h5->SetAxisRange(0, 130, "Y");
132     h5->SetLineColorAlpha(45, 0.35);
133
134     h->Draw("PL");
135     h2->Draw("PL SAME");
136     h3->Draw("PL SAME");
137     h4->Draw("PL SAME");
138     h5->Draw("PL SAME");
139
140     auto legend = new TLegend(1.0,1.0,1.0,1.0);
141     legend->AddEntry(h,"Valores Ciclo 1 : 87, 2, 44, 48", "p");
142     legend->AddEntry(h2,"Valores Ciclo 2: 79, 11, 45, 53", "p");
143     legend->AddEntry(h3,"Valores Ciclo 3: 85, 8, 49, 47", "p");
144     legend->AddEntry(h4,"Valores Ciclo 4: 84, 9, 58, 39", "p");
145     legend->AddEntry(h5,"Valores Ciclo 5: 89, 9, 52, 46", "p");
146
147     legend->SetFillStyle(6);
148     legend->SetBorderSize(2);
149     legend->Draw("");

```

```
150 c->SaveAs("final.png");  
151  
152 }
```

## 8. ANEXOS

### 8.1. Anexo A. Carta de consentimiento informado y motivos.



Puebla, Puebla, a 08 de Julio de 2021

#### **CARTA DE CONSENTIMIENTO INFORMADO Y MOTIVOS.**

Por medio de este documento que forma parte para la obtención del consentimiento informado, me permito invitarlo a ser partícipe de este estudio denominado ***Diseño e instrumentación de un sistema autónomo de señales biológicas para la mitigación en la propagación del COVID-19.*** Bajo la responsabilidad principal de Guadalupe Sandoval Moreno (pasante de Lic. en Física) y con la supervisión de: Benito de Celis Alonso (Doctor en Física Médica y Post-doc en la misma área y Eduardo Moreno Barbosa (Doctor en Ciencias y Post-doc Física Médica). Antes de decidir si nos brindará su valiosa participación es de suma importancia saber el porqué de este estudio y en qué consistirá su participación. Por favor lea detenidamente y con cuidado la siguiente información; si es necesario puede consultar con personas de su confianza como familiares o médicos acerca del presente estudio.

#### **1. Motivo**

El SARS-CoV-2 es un virus de origen asiático que produce la enfermedad COVID-19. Según datos de la ONU hasta el 07 de Julio se han presentado 184,105,272

casos confirmados en el mundo de COVID-19, con 3,988,565 muertes. En México se han presentado 2,549,862 casos totales de COVID-19, de los cuales resultaron 233,958 defunciones. A pesar de que se han creado vacunas no se ha podido vacunar a la población mundial entera. En memoria de las personas que se han ido se está creando una aplicación que ayude a detectar la enfermedad de la manera más precisa posible para completar el diagnóstico oportuno de esta enfermedad con la llamada e-salud o medicina con apoyo en dispositivos móviles.

## **2. Objetivo**

Recabar información auditiva con el fin de crear un algoritmo que relacione patrones de tos de personas sanas y personas que padezcan COVID-19.

## **3. Criterios de participación.**

Usted ha sido invitado a formar parte del grupo de estudio porque cumple con las características y parámetros necesarios.

**Sexo: indistinto.**

**Edad: 18 – 65 años**

**Diagnosticado con COVID-19 (preferentemente diagnosticado por PCR) o ser una persona “sana” pero que presente tos en ambos casos.**

## **4. ¿En qué consiste su participación?**

- Deberá contestar un breve cuestionario de 8 preguntas.
- Deberá mandar un audio de 5-20 segundos grabado preferentemente en un lugar silencioso donde se escuche su tos de manera nítida.

## **5. Tiempo de participación.**

Su participación termina después de mandar el audio, lo que no le causaría mayor problema posterior al envío.

Su participación es **voluntaria, anónima y confidencial**; nadie puede obligarle a participar bajo ninguna circunstancia. Es de suma importancia mencionar que este estudio busca investigar una metodología de detección de **COVID-19 Y NO SUSTITUIR NINGUNA PRUEBA**. La participación de cada persona es de suma importancia para el desarrollo de la aplicación móvil mencionada con anterioridad. **¡Ojalá pueda ayudarnos a hacer ciencia, muchas gracias!**

Si está de acuerdo en participar, le pediremos que escriba su nombre y firme el formato de Consentimiento Informado y firme al final del mismo.

\_\_\_\_\_

**NOMBRE Y FIRMA DEL PARTICIPANTE**



Guadalupe Sandoval Moreno

**NOMBRE Y FIRMA DEL RESPONSABLE**

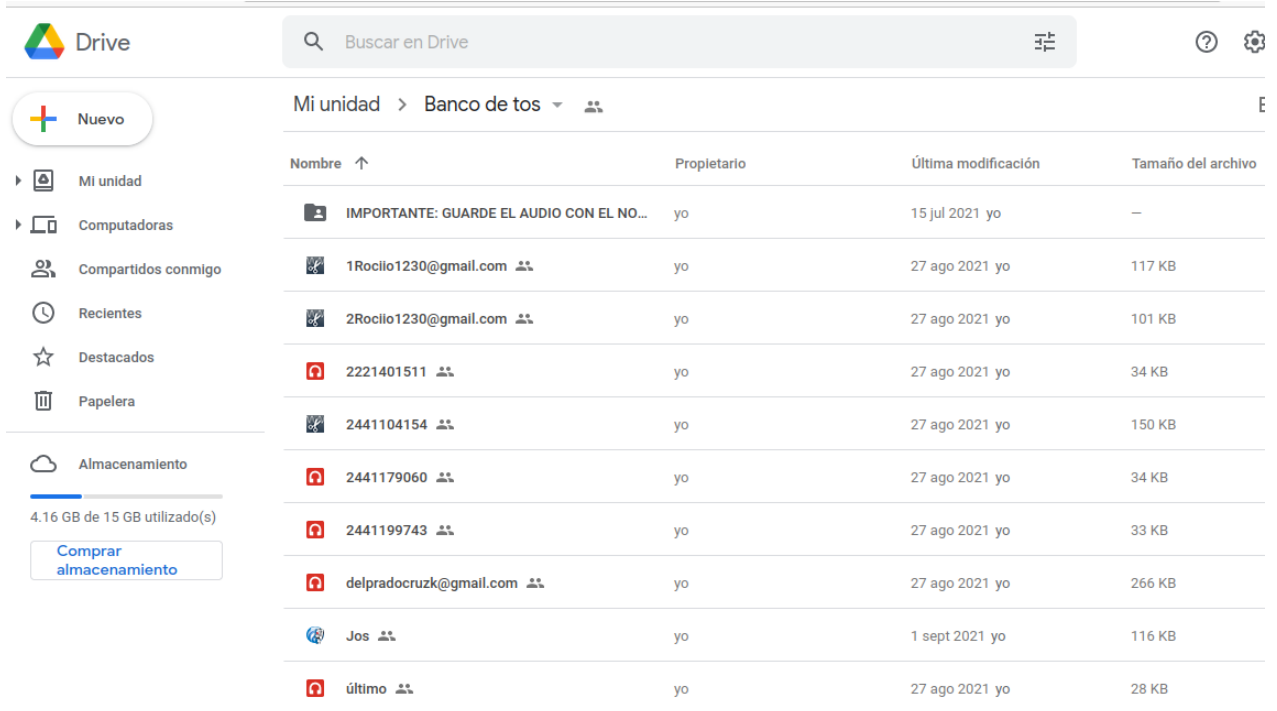


Eduardo Moreno Barbosa

Handwritten signature in black ink, consisting of the letters 'EMB' in a stylized, cursive font.

**NOMBRE Y FIRMA DEL INVESTIGADOR  
SUPERVISOR**

## 8.2. Anexo B. Banco de tos.



The screenshot shows the Google Drive interface. On the left, there is a sidebar with navigation options: 'Nuevo', 'Mi unidad', 'Computadoras', 'Compartidos conmigo', 'Recientes', 'Destacados', 'Papeler', and 'Almacenamiento' (4.16 GB de 15 GB utilizado(s) with a 'Comprar almacenamiento' button). The main area shows a search bar 'Buscar en Drive' and a breadcrumb 'Mi unidad > Banco de tos'. Below this is a table of files with columns: 'Nombre', 'Propietario', 'Última modificación', and 'Tamaño del archivo'.

| Nombre                                   | Propietario | Última modificación | Tamaño del archivo |
|--|-------------|---------------------|--------------------|
| IMPORTANTE: GUARDE EL AUDIO CON EL NO... | yo          | 15 jul 2021 yo      | —                  |
| 1Rocio1230@gmail.com                     | yo          | 27 ago 2021 yo      | 117 KB             |
| 2Rocio1230@gmail.com                     | yo          | 27 ago 2021 yo      | 101 KB             |
| 2221401511                               | yo          | 27 ago 2021 yo      | 34 KB              |
| 2441104154                               | yo          | 27 ago 2021 yo      | 150 KB             |
| 2441179060                               | yo          | 27 ago 2021 yo      | 34 KB              |
| 2441199743                               | yo          | 27 ago 2021 yo      | 33 KB              |
| delpradocruz@gmail.com                   | yo          | 27 ago 2021 yo      | 266 KB             |
| Jos                                      | yo          | 1 sept 2021 yo      | 116 KB             |
| último                                   | yo          | 27 ago 2021 yo      | 28 KB              |

8.3. Anexo C. Portada de NanNa APP.



#### 8.4. Anexo D. Fondo de despedida NanNa APP.



8.5. Anexo E. Logotipo de NanNa APP.



## 9. BIBLIOGRAFÍA

1. Arias V., Salazar J., Garicano C., Contreras J., Chacón G., Chacín-González M., Añez R., Rojas J. & Bermúdez-Pirela V. (2019) Una introducción a las aplicaciones de la inteligencia artificial en Medicina: Aspectos históricos. Revista Latinoamericana de Hipertensión. 14(5). pág. 590-609.
2. Barry J. S., Dane D. A., Morice H.A. & Walmsley D. A. (2006). The automatic recognition and counting of cough. BioMed Central, 2(8), pág. 1-9.  
doi:10.1186/1745-9974-2-8
3. Birring S.S., Fleming T., Matos S., Raj A.A., Evans D.H. & Pavord I.D. (2008). The Leicester Cough Monitor: preliminary validation of an automated cough detection system in chronic cough. European Respiratory Society, 31 (5), pág.1013-1018.  
DOI: 10.1183/09031936.00057407
4. Bravo G. S. & Cruz Q. J. P. (2015). Estudios de exactitud diagnóstica: Herramientas para su interpretación. Revista Chilena de Radiología. 21(4) pág 158-164.
5. Chaudhari G., Jiang X., Fakhry A., Han A., Xiao J., Shen S. & Khanzada A. (2021). Virufy: A Multi-Branch Deep Learning Network for Automated Detection of COVID-19. ArXiv.org. Pág. 1-9.  
<https://doi.org/10.48550/arXiv.2103.01806>
6. Consejo Nacional de Evaluación de la Política de Desarrollo Social. (5 de abril del 2022). Líneas de Pobreza por Ingresos en México, 1992 (enero) a 2022 (febrero), valores monetarios mensuales por persona a precios corrientes. Coneval. <https://www.coneval.org.mx/Medicion/MP/Paginas/Lineas-de-Pobreza-por-Ingresos.aspx>
7. Cortes O. J. A., Knott A. M., Osorio C. J. A. (2012). Aproximación a la síntesis de la música a través del análisis de fourier. Scientia et Technica Año XVII, 17(52), pág. 129-135

8. Cui, J., Li, F. & Shi, Z-L.(2019), Origin and evolution of pathogenic coronaviruses. Nature reviews microbiology, 17(3), pág.181-192.  
<https://doi.org/10.1038/s41579-018-0118-9>
9. Díaz Jiménez, I.V. (2020). Interpretación de las pruebas diagnósticas del virus SARS-CoV-2. Acta Pediátrica de México, 41 (4S1), pág. S51-S57.
- 10.Edge Impulse. (2020). Edge impulse processing blocks.  
<https://github.com/edgeimpulse/processing-blocks>
- 11.Fernández S. M. (2014). La Salud 2.0 y la atención de la salud en la era digital. Revista Médica de Risaralda, 20(1), pág. 41-46.  
DOI: <https://doi.org/10.22517/25395203.8483>
- 12.Fuente-Alba C. S. & Villagra M. M. (2016). Likelihood ratio (razón de verosimilitud): definición y aplicación en Radiología. 8(3). Pág. 204-208.
- 13.Gérard A. A. (2011). El sonograma: una representación práctica de los sonidos. Revista Boliviana de Física, 18(18), pág. 37-49.
- 14.Guo, Y-R., Cao Q-D., Hong Z-S., Tan Y-Y., Chen S-D., Jin H-J., Tan K-S., & Wang D-Y,. & Yan Y. (2020). The origin, transmission and clinical therapies on coronavirus disease 2019 (COVID-19) outbreak – an update on the status. Military Medical Research, 7(11), pág.1-10.  
<https://doi.org/10.1186/s40779-020-00240-0>
- 15.Gutiérrez Cuerva J. M. (2021). Desarrollo de aplicaciones basadas en aprendizaje automático en sistemas empotrados STM32 [Tesis de Licenciatura, Universidad de Sevilla].  
[https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwi69djWwbr3AhUhEEQIHxyjD2EQFnoECAYQAQ&url=https%3A%2F%2Fbiblus.us.es%2Fbibing%2Fproyectos%2Fabreproy%2F93362%2Fdescargar\\_fichero%2FTFG-3362%2BCUERVA%2BGUTIERREZ%252C%2BJUAN%2BMANUEL.pdf&usg=AOvVaw0GLWMISWRwj0fgnQdlmxwK](https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwi69djWwbr3AhUhEEQIHxyjD2EQFnoECAYQAQ&url=https%3A%2F%2Fbiblus.us.es%2Fbibing%2Fproyectos%2Fabreproy%2F93362%2Fdescargar_fichero%2FTFG-3362%2BCUERVA%2BGUTIERREZ%252C%2BJUAN%2BMANUEL.pdf&usg=AOvVaw0GLWMISWRwj0fgnQdlmxwK)
- 16.Guyton C., A. & Hall E., J. (2011). Tratado de fisiología médica. Elsevier.

17. Hollmann, S., Frohme M., Endrullant C., Kremer A., D'Elia D., Regierer B., & Nechyporenko A. (2020). Ten simple rules on how to write a standard operating procedure. *PLOS Computational Biology*, 16 (9), pág. 1-10.  
<https://doi.org/10.1371/journal.pcbi.1008095>
18. Instituto Nacional de Estadística y Geografía. (5 de abril del 2022). Población. Número de habitantes en México 2020 INEGI.  
<https://www.inegi.org.mx/temas/estructura/>
19. Kantor N. I. (2020). Sobre diagnóstico, testeos y prevalencia de COVID-19. *MEDICINA (Buenos Aires)*. 80(Supl. III). Pág. 77-78
20. Lee H. & Song J. (2019). Introduction to convolutional neural network using Keras; an understanding from a statistician. *Communications for Statistical Applications and Methods*. 26(6), 591-610
21. Méndez C. O. (2021) Metodología para la evaluación del poder descriptivo de un MARCADOR mediante el análisis ROC (Receiver-Operating Characteristic) [Tesis de Licenciatura, Benemérita Universidad Autónoma de Puebla].  
<https://repositorioinstitucional.buap.mx/bitstream/handle/20.500.12371/12692/2/0210324175347-2130-TL.pdf?sequence=2&isAllowed=y>
22. Procuraduría Federal del Consumidor. (5 de abril del 2022). Pruebas COVID. Costo de pruebas diagnósticas pcr y antígenos covid19.  
<https://www.gob.mx/profeco/es/articulos/pruebas-covid?idiom=es>
23. Sarmiento R. J. L. (2020), Aplicaciones de las redes neuronales y el deep learning a la ingeniería biomédica. *Revista UIS Ingenierías*. 19(4). pág 1-18  
DOI: <https://doi.org/10.18273/revuin.v19n4-2020001>
24. Sociedad Americana para Pruebas y Materiales. (Re-aprobada en 2003). Standard Specification for Infrared Thermometers for Intermittent Determination of Patient Temperature [Especificaciones estándar para termómetros infrarrojos para la determinación intermitente de la temperatura del paciente]. ASTM International.



25. Soldevila L. L., Valerio S. L. & Roure D. S. (2021). Interpretación de las pruebas diagnósticas de la COVID-19. *Formación Médica Continuada en Atención Primaria*. 28(3). pág 167-173
26. Suárez V., Suarez Quezada M., Oros Ruiz S., Ronquillo De Jesús E. (2020). Epidemiology of COVID-19 in Mexico: From the 27th of February to the 30th of April 2020. *Revista Clínica Española*. Vol. 220 (8), pág 461-473
27. Tortora J. G. y Anagnostakos P. N. (1981) *Principios de anatomía y fisiología*. Editorial médica Panamericana.
28. Virufy-data. (2020). Virufy. Bases de datos de tos. [https://github.com/virufy/virufy-data?fbclid=IwAR2UmggDeB\\_\\_NnAoYeSqHLFw9yPzQtxVf4s3hh63huDO29w\\_hSLH8LYSJ4w](https://github.com/virufy/virufy-data?fbclid=IwAR2UmggDeB__NnAoYeSqHLFw9yPzQtxVf4s3hh63huDO29w_hSLH8LYSJ4w)
29. Instituto de Diagnóstico y Referencia Epidemiológicos “Dr. Manuel Martínez Báez”. (2020). Lineamientos para la toma, manejo y envío de muestras para diagnóstico a la red nacional de laboratorios de salud pública [Archivo PDF]. [https://www.gob.mx/cms/uploads/attachment/file/558702/Lineamientos\\_TMEM\\_2020\\_180620.pdf](https://www.gob.mx/cms/uploads/attachment/file/558702/Lineamientos_TMEM_2020_180620.pdf)
30. Wang, H., Li, X., Li, T. Zhang S., Wang, L., Wu, X. & Liu, J. (2020). The genetic sequence, origin, and diagnosis of SARS-CoV-2. *European Journal of Clinical Microbiology & Infectious Diseases*, 39 (9), pág. 1629-1635. <https://doi.org/10.1007/s10096-020-03899-4>
31. Wang J. J.Y. & Gao X. (2015). Max-min distance nonnegative matrix factorization. *Neural Networks* 61 pág 75-84. <https://doi.org/10.1016/j.neunet.2014.10.006>.