



**Implementación de un sistema con inteligencia computacional para identificar  
dificultad respiratoria a partir del procesamiento digital de señales de voz**

**Sara Isabel Fernández Velasco**

**Karen Andrea Ramos Casanova**

20561722451

20561725267

**Universidad Antonio Nariño**

Programa Ingeniería Biomédica

Facultad de Ingeniería Mecánica, Electrónica y Biomédica

Popayán - Cauca, Colombia

2021

**Implementación de un sistema con inteligencia computacional para identificar dificultad respiratoria a partir del procesamiento digital de señales de voz**

**Sara Isabel Fernández Velasco**

**Karen Andrea Ramos Casanova**

Proyecto de grado presentado como requisito parcial para optar al título de:

**Ingeniero Biomédico**

Director (a):

PhD. Julián Antonio Villamarín Muñoz

Línea de Investigación:

Análisis de información biomédica

Grupo de Investigación:

Grupo de investigación bioingeniería Universidad Antonio Nariño

**Universidad Antonio Nariño**

Programa Ingeniería Mecánica

Facultad de Ingeniería Mecánica, Electrónica y Biomédica

Popayán - Cauca, Colombia

2021

## NOTA DE ACEPTACIÓN

El trabajo de grado titulado

\_\_\_\_\_, Cumple con

los requisitos para optar

Al título de \_\_\_\_\_.

\_\_\_\_\_

Firma del Tutor

\_\_\_\_\_

Firma Jurado

\_\_\_\_\_

Firma Jurado

Popayán, 02, noviembre 2021.

## Contenido

	<b>Pág.</b>
Preliminares .....	X
Resumen.....	12
Abstract .....	13
Introducción .....	14
1. Generalidades del proyecto.....	17
1.1 Planteamiento del problema.....	17
1.2 Justificación .....	19
1.3 Objetivos .....	21
1.3.1 Objetivo general.....	21
1.3.2 Objetivos específicos .....	21
1.4 Alcances .....	21
2. Estado del arte.....	23
2.1 Artículos científicos .....	23
2.2 Patentes .....	24
2.3 Tecnologías de búsqueda abierta .....	26
3. Marco teórico .....	29
3.1 Anatomía del sistema fonorespiratorio .....	29
3.2 Fisiología del sistema fonorespiratorio.....	31
3.3 Patologías del sistema fonorespiratorio .....	33
3.3.1 Patologías del sistema respiratorio.....	33
3.3.2 Patologías del sistema fonador.....	34

3.4	Parámetros diagnósticos clínico acústicos .....	36
3.4.1	Parámetros de perturbación de la voz .....	36
3.4.2	Parámetros espectrales .....	37
3.4.3	Parámetros temporales .....	38
3.4.4	Parámetros estadísticos .....	40
3.5	Indicadores aerodinámicos.....	41
3.6	Modelo cliente - servidor .....	42
3.7	Redes neuronales .....	44
3.7.1	Arquitectura red neuronal .....	45
4.	Materiales y métodos .....	47
4.1	Componentes del sistema.....	47
4.1.1	Adquisición de señales de voz .....	48
4.1.2	Método computacional.....	50
4.1.3	Red neuronal .....	52
4.1.4	Arquitectura sistema cliente servidor.....	53
4.1.5	Sistema web, captura de señales y transmisión .....	55
5.	Resultados.....	57
5.1	Diferencias gráficas sano - patológico .....	57
5.2	Resultados de parámetros en las señales.....	60
5.3	Gráficas de dispersión.....	67
5.4	Red neuronal .....	70
5.5	Sistema web .....	73
	Conclusiones .....	78

Anexos .....	79
Referencias Bibliográficas .....	80

## Lista de Figuras

Figura 3-1: Vía aérea superior .....	29
Figura 3-2: Vía aérea inferior .....	30
Figura 3-3: Cuerdas vocales durante la fonación - Respiración .....	31
Figura 3-4: Fisiología del aparato fonador.....	32
Figura 3-5: Formantes y frecuencia fundamental .....	38
Figura 3-6: Cruces por cero de una señal.....	40
Figura 3-7: Modelo cliente servidor .....	43
Figura 3-8: Estructura de una red neuronal artificial .....	45
Figura 3-9: Arquitectura red neuronal .....	46
Figura 4-10: Representación diagramática de componentes y funcionamiento general del sistema .....	47
Figura 4-11: Operación del sistema .....	48
Figura 4- 12: Protocolo para capturar la voz .....	49
Figura 4-13: Procesamiento digital de señales de voz .....	51
Figura 4-14: Estructura red reconocimiento de patrones.....	52
Figura 4-15: Agrupamiento de datos de entrenamiento, prueba y validación .....	53
Figura 4-16: Arquitectura cliente – servidor.....	54
Figura 4-17: Interfaz en el App designer de Matlab .....	56
Figura 5-18: Vocal “a”, voz sana - patológica (Paciente mujer) .....	57
Figura 5-19: Vocal “o”, voz sana - patológica (Paciente mujer) .....	58
Figura 5-20: Vocal “a”, voz sana - patológica (Paciente hombre) .....	59
Figura 5-21: Vocal “o”, voz sana - patológica (Paciente hombre) .....	59

Figura 5-22: Comparación entre la señal de la vocal “a” sana y patológica para cada parámetro en hombres.....	64
Figura 5-23: Comparación entre la señal de la vocal “o” sana y patológica para cada parámetro en hombres.....	64
Figura5- 24: Comparación entre la señal de la vocal “a” sana y patológica para cada parámetro en mujeres. ....	65
Figura 5-25: Comparación entre la señal de la vocal “o” sana y patológica para cada parámetro en mujeres. ....	65
Figura 5-26: Hombre vocal ‘a’ .....	68
Figura 5-27: Mujer vocal ‘a’ .....	68
Figura 5-28: Hombre vocal ‘o’ .....	69
Figura 5-29: Mujer vocal ‘o’.....	69
Figura 5-30: Matriz de confusión mujeres vocal a .....	71
Figura 5-31: Matriz de confusión hombres vocal o .....	72
Figura 5-32: Discriminación de sexo.....	73
Figura 5-33: Interfaz de inicio – Analizar la voz.....	74
Figura 5-34: Interfaz de inicio – Contribuir a los protocolos de bioseguridad.....	74
Figura 5-35: Interfaz de inicio – Donar la voz.....	75
Figura 5-36: Análisis de la voz - Grabar la voz .....	76
Figura 5-37: Análisis de la voz - Grabando .....	76
Figura 5-38: Análisis de la voz - Resultado.....	77

**Lista de tablas**

Tabla 2- 1: Resultados de búsqueda de artículos científicos .....	23
Tabla 2-2: Resultados de búsqueda de patentes.....	25
Tabla 2-3: Resultados de tecnologías de búsqueda abierta.....	27
Tabla 3-4: Patologías del sistema respiratorio .....	33
Tabla 3-5: Patologías del sistema fonador .....	35
Tabla 5-6: Diferencias porcentuales (%) sano - patológico en vocal a.....	60
Tabla 5-7: Diferencias porcentuales (%) sano-patológico en vocal o .....	61
Tabla 5-8: Media aritmética voz sana-voz patológica .....	61
Tabla 5-9: Desviación estándar voz sana-voz patológica .....	66
Tabla 5-10: Resultados Red neuronal con 2 entradas.....	70
Tabla 5-11: Resultados Red neuronal con 3 o más entradas .....	70
Tabla 5-12: matrices de confusión.....	72

## Preliminares

*Nuestra recompensa se encuentra en el esfuerzo y no en el resultado. Un esfuerzo total, es una victoria completa.*

*Mahatma Gandhi*

## **Agradecimientos**

A Dios, por darme salud, paciencia y sabiduría para llevar a cabo la meta propuesta. A mis padres, abuelos y hermanas por siempre brindarme su amor y apoyo incondicional. A mi compañera de tesis por su tiempo, esfuerzo y dedicación, a la Universidad Antonio Nariño, a sus docentes por ofrecernos las herramientas y conocimientos necesarios para desarrollar nuestro proyecto de grado. Un agradecimiento especial, al director de tesis, PhD. Julián Antonio Villamarín por su constante dedicación, orientación y apoyo durante el desarrollo del mismo.

*Sara Isabel Fernández V.*

A Dios, por la vida, por permitirme disfrutar de este logro. A mis padres, hermana y abuelos por su amor y apoyo incondicional. A mi compañera de tesis por el tiempo y esfuerzo compartido a lo largo de este proyecto, a la Universidad Antonio Nariño y sus docentes en especial, al asesor de tesis, el PhD. Julián Antonio Villamarín por compartir sus conocimientos y por el apoyo brindado para poder alcanzar este triunfo.

*Karen Andrea Ramos C.*

## Resumen

El sistema desarrollado incorpora inteligencia computacional, este permite identificar personas con dificultad respiratoria (causada por gripe) de forma automática y no invasiva, a partir del procesamiento digital de señales de voz, integrando el cálculo de parámetros acústicos, espectrales, temporales y estadísticos, que implementados en una arquitectura cliente servidor, permiten el análisis respectivo, obtenidos como resultado de la grabación de dos sonidos vocálicos sostenidos (vocal “a” y “o”), a través del micrófono de un dispositivo móvil de un grupo poblacional hispano hablante con edades entre los 20 y 49 años.

Los resultados obtenidos determinan que para la detección de dificultad respiratoria, la vocal “a” resulta más eficiente en mujeres con una tasa de acierto de 97.62% y la vocal “o” en hombres; alcanzando una tasa de acierto de 96.77%. Se espera que este sistema contribuya a nuevas herramientas de apoyo con potencial de aplicación en salud, al fomentar los protocolos de bioseguridad, especialmente en este contexto de pandemia por Covid-19; enfermedad causante de dificultad respiratoria, al igual que la gripe.

**Palabras clave:** Sistema computacional, voz, dificultad respiratoria, parámetros acústicos, inteligencia computacional.

### **Abstract**

The developed system incorporates computational intelligence, this allows to identify people with respiratory distress (caused by influenza) in an automatic and non-invasive way, from the digital processing of voice signals, integrating the calculation of acoustic, spectral, temporal and statistical parameters, which implemented in a client-server architecture, they allow the respective analysis, obtained as a result of the recording of two sustained vowel sounds (vowel "a" and "o"), through the microphone of a mobile device of a Spanish-speaking population group aged between 18 and 49 years old.

The results obtained determine that for the detection of respiratory distress, the vowel "a" is more efficient in women with a hit rate of 97.62% and the vowel "o" in men; reaching a hit rate of 96.77%. This system is expected to contribute to new support tools with potential for application in health, to promote biosafety protocols, especially in this context of a Covid-19 pandemic; illness that causes respiratory distress, like the flu.

**Keywords:** Computational system, voice, respiratory distress, acoustic parameters, computational intelligence.

## Introducción

Las enfermedades respiratorias afectan todos los órganos que intervienen en el proceso de la respiración, incluye desde una simple gripe hasta la enfermedad pulmonar obstructiva crónica. A partir de los sonidos generados por la respiración; torácicos o incluso la voz, se obtienen señales biológicas, que son procesadas por medio de softwares como Matlab, ANAGRAF o PRAAT, donde se extraen parámetros acústicos como la frecuencia fundamental ( $F_0$ ), frecuencia de las formantes, jitter, shimmer, coeficientes cepstrales o cepstrum, relación armónico - ruido (HNR), relación señal-ruido, entre otros. Adicionalmente, se utilizan clasificadores de alto aprendizaje, como las Redes Neuronales (RNA); Red neuronal Monocapa, Multicapa, Convolutacional (CNN), recurrente (RNN) y Redes de base radial (RBF). Cuando se realiza en conjunto la adquisición, procesamiento y clasificación de los sonidos torácicos o la voz en este caso, son de gran apoyo en el área de salud; siendo un sistema con fines diagnósticos o también para observar el comportamiento del sistema respiratorio en una condición determinada; estos sistemas son validados con la ayuda de profesionales en salud y a través de parámetros estadísticos como las matriz de confusión, sensibilidad, especificidad y tasa de acierto.

Se han encontrado investigaciones en las que se observa el comportamiento de las señales de voz, utilizadas en conjunto con sonidos respiratorios (sibilancias, crepitancias, tos) para la detección de algunas enfermedades respiratorias como asma, neumonía y posiblemente Covid-19. En el artículo AI4COVID-19: AI enabled preliminary diagnosis for COVID-19 from cough samples via an app (Imran et al., 2020), se implementa una prueba para detectar covid-19 a través de una aplicación móvil llamada AI4COVID-19; la cual requiere grabaciones de tos; de igual forma en Mayorga et al. (2010) en Valoración de Enfermedades Respiratorias Basada en

Modelación Acústica, se presenta un método para una valoración cuantitativa en la salud de pacientes relacionados con desórdenes respiratorios usando sonidos del pulmón. En cambio en Gómez et al. (2016) An insight to the automatic categorization of speakers according to sex and its application to the detection of voice pathologies: A comparative study, estudian señales de voz pero su aplicación radica en discriminar el sexo de una persona y en detectar patologías de la voz; además, en Elisei (2012) Análisis acústico de la voz normal y patológica utilizando dos sistemas diferentes: ANAGRAF y PRAAT, utilizan las señales de voz para detectar y caracterizar el habla, la voz y los desórdenes vocales; el propósito es comparar los resultados en ambos software, obtenidos con un conjunto de mediciones acústicas.

También se hallaron patentes relacionadas con señales de voz. La invención en Gómez Vilda (2011), método y sistema para la estimación de parámetros fisiológicos de la fonación, consiste en un método y sistema de cómputo para el registro y análisis de la voz, que permite calcular una serie de parámetros de la fonación; obtienen información relevante sobre influencias causadas por trastornos orgánicos o neurológicos. En Travieso et al. (2013) Método para la evaluación clínica del sistema fonador de pacientes con patologías laríngeas a través de una evaluación acústica de la calidad de la voz se refiere a un método que evalúa de forma objetiva una grabación de la fonación sostenida de una vocal por medio de un sistema informático. Este método mide la estabilidad de la voz, la riqueza espectral, la presencia de ruido, etc. La patente Edinburgh & Dunfermline (2014) sistemas y métodos para identificar al paciente que habla durante la medición de un parámetro fisiológico, la lente, presenta una unidad de monitorización de pacientes para procesar la señal de un sonido que emana el paciente para identificar información sobre la respiración.

Estos artículos científicos presentan la detección o valoración de enfermedades respiratorias a partir de sonidos respiratorios; adicionalmente los artículos que estudian la voz, la utilizan para clasificar la calidad de esta, discriminar personas por sexo, o identificar usuarios y patologías netamente de la voz.

Tanto artículos como patentes no están destinados al estudio de dificultad respiratoria específicamente a través de señales de voz, por lo que se convierte en una razón, además del contexto de pandemia por Covid-19, para el desarrollo de nuevas tecnologías. En este proyecto se decide implementar un sistema computacional para detectar dificultad respiratoria a partir del procesamiento digital de señales de voz y redes neuronales, obteniendo una evaluación objetiva y no invasiva, y siendo de gran apoyo en el área de la salud; aportando a los protocolos de bioseguridad para evitar la propagación y contagio por Covid-19 u otra enfermedad respiratoria; teniendo como objetivo implementar un sistema computacional de análisis de la voz para identificar la presencia de dificultad respiratoria causada por gripe, a través del procesamiento digital de señales y redes neuronales. Para esto se utiliza la metodología CDIO (Concebir - Diseñar - Implementar - Operar), y se plantean los alcances y limitaciones del proyecto; el sistema se limita a identificar la dificultad respiratoria causada por gripe, involucra directamente el hecho de que una persona tenga la capacidad para hablar y de esta manera recolectar su voz a través de su dispositivo electrónico, para analizar la información y generar una respuesta binaria; es decir se indica si el hablante posee dificultad respiratoria o no. El proyecto además de impactar en los protocolos de bioseguridad, aporta a la línea de investigación de análisis de informaciones biomédicas del grupo de investigación bioingeniería (UAN); en el cual se relacionan el desarrollo de técnicas de procesamiento digital de bioseñales con el propósito de extraer información de apoyo al diagnóstico clínico.

## **1. Generalidades del proyecto**

### **1.1 Planteamiento del problema**

Las enfermedades del aparato respiratorio son aquellas que afectan a los diferentes órganos y sistemas que participan en la respiración, los síntomas más frecuentes son la tos, la disnea (sensación subjetiva de dificultad en la respiración) y el dolor torácico (Galera et al., 2010).

Según nuevas estimaciones de los Centros para el Control y la Prevención de Enfermedades de los Estados Unidos, la Organización Mundial de la Salud (OMS) y asociados para la salud mundial, cada año se producen hasta 650.000 defunciones por enfermedades respiratorias (Organización Mundial de la Salud, 2017). En diciembre de 2019 se reportó a la OMS los primeros casos de neumonía (Covid-19) (OMS, 2020.); hasta la fecha, noviembre 02 del 2021, los casos confirmados por coronavirus en el mundo, son más de 247 millones y más de 5 millones de personas fallecidas. En Colombia, de las personas positivas al coronavirus el 48.59% son hombres y el restante 51.41% son mujeres. Respecto a la cifra de fallecidos, en Colombia se aproxima a más de 127 mil (MSN noticias, 2021); esta situación ha sumido a la economía mundial en una fuerte recesión (World Bank, 2020).

Para contener la propagación de la enfermedad se implementa protocolos de bioseguridad enfocados al análisis del riesgo, la mejora de la preparación y la respuesta ante las situaciones de urgencia (OMS y Organización de las Naciones Unidas para la Agricultura y la Alimentación (ONUAA), 2010.) por lo que es importante resaltar el desarrollo de nuevas herramientas de apoyo con potencial de aplicación en salud, esenciales para este contexto de pandemia (Leiva et al., 2020).

De acuerdo a la vigilancia tecnológica realizada la reducción de la presión parcial de O<sub>2</sub> en la sangre arterial (hipoxia) o un incremento de la presión parcial de CO<sub>2</sub> (hipercapnia) (Azuaje,

2019) la insuficiencia cardiaca, gripe, las infecciones virales y neumonía son algunas de las enfermedades desencadenantes de insuficiencia respiratoria (Casas et al., 2008) y por ende de dificultad respiratoria, debido a que un nivel bajo de oxígeno en la sangre puede causar dificultad para respirar; la insuficiencia se puede diagnosticar a través de la gasometría arterial (MedlinePlus enciclopedia médica, 2021), la cual es una medición de la cantidad de oxígeno y de dióxido de carbono presente en la sangre de una arteria; también la radiografía de tórax (National Heart, Lung and Blood Institute (NIH, 2019), las pruebas de función pulmonar como espirometría y oximetría de pulso son métodos utilizados para la detección de insuficiencia respiratoria (NIH, 2021).

Por otra parte, el procesamiento de señales de voz y la extracción de parámetros acústicos pueden convertirse en una herramienta potencial para detectar la dificultad respiratoria (Catalunya Vanguardista, 2020), considerando que el sistema fonatorio emplea la energía aerodinámica suministrada por los pulmones para sonorizar el aire, por tanto las afectaciones del sistema respiratorio se reflejarán en el sistema fonatorio, especialmente en la etapa de emisión acústica. Lo anterior se fundamenta en las investigaciones desarrolladas por parte del Hospital de Clínicas de la Facultad de Medicina de la Universidad de São Paulo, las cuales han evidenciado la viabilidad técnica en la implementación de un sistema para la detección temprana de insuficiencia respiratoria (Plataforma SPIRA), a través del análisis de patrones divergentes en señales de voz, al comparar muestras de voz de personas que no padecen la enfermedad con descriptores de voces de personas que exhiben insuficiencia respiratoria, al analizar pausas en el habla, pausas en la respiración entre otras características acústicas estimadas al emitir la voz (Goldman et al., 2020).

De igual forma en Rey et al. (2017) Implementación de un algoritmo para la identificación de usuarios considerando problemas fisiológicos que afectan el habla, se evidencia el desarrollo de un sistema de clasificación para identificar usuarios con problemas respiratorios, a partir de

afecciones de la voz, modelando un sistema de clasificación a partir de parámetros acústicos y coeficientes cepstrales, con tasas de acierto del 96.67% y 77.78% respectivamente. Lo anteriormente mencionado refleja el potencial de la implementación de sistemas computacionales de análisis de la voz como herramienta no invasiva en el análisis fonorespiratorio.

En este contexto, la pandemia por Covid-19 ha impulsado el surgimiento de nuevas tecnologías para la detección de algunas enfermedades respiratorias como asma, neumonía, y covid-19 (Gómez García et al., 2016), utilizando sonidos respiratorios (sibilancias, crepitancias, tos); como Brown et al. (2021), explorando el diagnóstico automático de COVID-19 a partir de datos de sonido respiratorio de fuentes colectivas, los modelos clasificatorios planteados logran un desempeño superior al 80% en todas las tareas.

En líneas generales, la especificidad del problema en el presente proyecto, radica en cómo detectar la dificultad respiratoria (causada por gripe), teniendo una evaluación objetiva y no invasiva, a través del procesamiento digital de señales de voz y redes neuronales; como herramienta de apoyo al área de la salud, y con potencial de ser extendida como instrumento de apoyo a los protocolos de bioseguridad, ya que en el mediano o futuro plazo, se contaría con una herramienta tecnológica que identificaría a personas con dificultad respiratoria, a partir del análisis fonorespiratorio, indicando alertas para la no exposición de las personas con riesgo de propagar enfermedades respiratorias.

## **1.2 Justificación**

La presente investigación se enfoca en el desarrollo de un sistema computacional con el fin de identificar la presencia de dificultad respiratoria a través del procesamiento digital de señales de voz y redes neuronales. Siendo el análisis acústico de la voz un método adecuado para obtener una evaluación objetiva y no invasiva con fines diagnósticos (Delgado et al., 2017). De esta forma,

se puede obtener un diagnóstico preliminar ante una alteración en el aparato fonorespiratorio debido a que se realiza a partir de cálculos matemáticos de parámetros comúnmente utilizados para procesar señales biológicas como jitter, shimmer, frecuencia fundamental ( $F_0$ ), relación armónico ruido (HNR), coeficientes cepstrales, entre otros. Los valores resultantes de estos parámetros son utilizados como entradas para clasificadores automáticos de alto aprendizaje en redes neuronales artificiales (RNA) (Marlina et al., 2017).

Debido a las investigaciones más recientes y a la falta de dispositivos existentes en el mercado con relación al aparato fonorespiratorio, se pretende desarrollar e implementar un sistema capaz de identificar la presencia de dificultad respiratoria a través del procesamiento de señales de voz; siendo un método no invasivo en vista de que, al obtener la señal no hay instrumentos, ni contacto directo que afecten físicamente el cuerpo del paciente. Es importante resaltar, que este sistema sería de gran utilidad para detectar de forma temprana y sencilla una posible alteración en el sistema respiratorio, que aportaría directamente en el bienestar del usuario, permitiendo asistir a un profesional en la salud antes de llegar a un problema respiratorio crónico; también a la salud pública en general, al evitar el contagio de enfermedades respiratorias con fácil propagación y por último, al personal de salud, facilitando un diagnóstico preliminar, que ahorraría tiempo, costos e incomodidades al paciente.

Para la comunidad científica este proyecto abre las puertas a nuevas investigaciones referentes al procesamiento de señales de voz, detección de dificultad respiratoria y a la creación de nuevos sistemas e instrumentos con aplicación en salud.

## **1.3 Objetivos**

### ***1.3.1 Objetivo general***

Implementar un sistema computacional de análisis de la voz para identificar la presencia de dificultad respiratoria causada por gripe, a través del procesamiento digital de señales y redes neuronales.

### ***1.3.2 Objetivos específicos***

- Establecer los descriptores fonorespiratorios apropiados en la distinción entre voces obtenidas desde pacientes con síntomas de dificultad respiratoria y personas sin alteración, a partir del procesamiento digital de señales de voz.

- Desarrollar un sistema de clasificación mediante Redes Neuronales, incorporando como descriptores de entrada parámetros fonorespiratorios estimados automáticamente y correlacionados con la presencia de dificultad respiratoria causada por gripe.

- Evaluar el desempeño clasificatorio del sistema computacional, con pacientes sanos y patológicos, a través del análisis por matriz de confusión, tasa de acierto y un consenso de profesionales especialistas en voz.

## **1.4 Alcances**

En este proyecto se pretende desarrollar un sistema computacional para identificar la presencia de dificultad respiratoria causada por gripe a través del procesamiento digital de señales de voz y redes neuronales.

El sistema se limita a identificar dificultad respiratoria causada por gripe, debido a que esta, puede desencadenarse por diversas causas como enfermedades respiratorias, problemas cardíacos, obesidad, afecciones que provocan daños en los músculos que controlan la respiración, etc.

Por otra parte, este sistema involucra directamente el hecho de que una persona tenga la capacidad para hablar y de esta manera recolectar su voz a través de su dispositivo electrónico, para analizar la información. Algunas de las limitaciones en este caso, es cuando una persona es muda o presenta cualquier otra dificultad en el habla.

Adicionalmente, cabe resaltar que el sistema genera una respuesta binaria; indica si el hablante posee dificultad respiratoria o no; el sistema no está diseñado para diagnosticar una enfermedad respiratoria en específico.

## 2. Estado del arte

### 2.1 Artículos científicos

En la actualidad las enfermedades respiratorias afectan a los diferentes órganos que participan en la respiración como las vías nasales, los pulmones y otras vías respiratorias. De este modo, las alteraciones de la función del sistema respiratorio conducen a problemas de dificultad e insuficiencia respiratoria, enfermedades respiratorias crónicas, entre otras. De esta manera surgen tecnologías para la evaluación de la función respiratoria (Tabla 1).

**Tabla 2- 1:**

#### Resultados de búsqueda de artículos científicos

Título	Descripción	Institución	Año
Lung function estimation from a monosyllabic voice segment captured using smartphones	Los autores se enfocaron en enfermedades respiratorias crónicas; con un direccionamiento conveniente basado en dispositivos móviles que utiliza un monosilábico de voz llamado sonido 'vocal A' o sonido 'Aaaa ...' para estimar la función pulmonar; para esto se desarrolló un modelo que detecta el sonido de la 'vocal A' de otro acústico y un modelo de predicción para estimar la función pulmonar usando el sonido de la vocal A. El estudio muestra que los sonidos de las vocales A se pueden detectar con un 93% de precisión, y los sonidos de las vocales A pueden estimar las funciones pulmonares con un error absoluto medio del 7,4-11,35%, (Saleheen et al, 2020).	Department of Computer Science University of Memphis, Digital Health Lab Samsung Research America Inc, Pulmonary and Critical Care Division Brigham and Women's Hospital Harvard Medical School	2020
Acoustic Methods for Pulmonary Diagnosis	Se hizo una revisión de los fundamentos de las señales acústicas pulmonares y de las características fisiopatológicas de las	Universidad de California USA	2019

---

Information system for diagnosis of respiratory system diseases	<p>enfermedades que se utilizan para su detección. Se enfocaron en diferentes métodos para la medición y creación de señales utilizadas en investigaciones recientes para el diagnóstico de enfermedades pulmonares. Estos métodos permiten la reducción de ruido y una mejor extracción de características y clasificación de señales, (Rao et al.,2019).</p> <p>Se presenta un sistema de información para el diagnóstico de pacientes con enfermedades pulmonares, en el cual se utilizan fragmentos de tos grabados con una grabadora de voz para el desarrollo de un sistema de reconocimiento de tos que puede ser utilizado por médicos especialistas para diferentes diagnósticos, (Abramov et al, 2018).</p>	<p>University Square, Voronezh, Russia Voronezh State University of Engineering Technologies</p>	2018
---	---	--	------

---

Nota: Elaboración propia, 2021

## 2.2 Patentes

El desarrollo de tecnologías destinadas a evaluar la función respiratoria se ha extendido al pasar de los años, cada una con ciertas características que la diferencian de otras, esto ha incentivado a patentar las herramientas tecnológicas, algunas de estas se presentan en la tabla 2.

**Tabla 2-2:****Resultados de búsqueda de patentes**

<b>Título</b>	<b>Descripción</b>	<b>País</b>	<b>Año</b>
Systems and methods for analyzing and monitoring lung function using voice and breath sound samples for respiratory care	Se refiere a un motor de análisis de voz, implementado por computadora, el cual recibe datos desde un dispositivo informático, analiza los datos del mensaje de transmisión para predecir el estado respiratorio del usuario a partir de un modelo de aprendizaje. (Venneti et al., 2020)	EEUU	2020
Device and method for diagnosing respiratory disease	Refiere un aparato y método para el diagnóstico de enfermedades respiratorias reconociendo una voz de un usuario, comprende de una unidad de entrada de reconocimiento de sonido, una de preprocesamiento, un detector de movimiento y una unidad de procesamiento de señales, para analizar una señal de voz relacionada con la respiración del usuario determinada por el preprocesador usando datos de entrenamiento asegurados de antemano para diagnosticar la enfermedad respiratoria. (Hong et al., 2019)	Corea	2019
Voice inspecting device	Habla de un convertidor que transforma la señal de voz ingresada desde un micrófono en una señal digital, al mismo tiempo que es almacenada en una memoria de datos, una parte de operación realiza una transformación de fourier convirtiendo los datos en datos de espectro de voz secuenciales de tiempo de frecuencia. Una parte de operación estadística opera la desviación estándar que muestra el grado o extensión de la dispersión y distribución en los datos del espectro de voz y la calcula como un valor de índice. (Kazunori Kita, 1995)	Japón	1995

---

Method and system for screening for covid-19 with a vocal biomarker	Se refiere a un método informático para examinar sujetos desconocidos para detectar covid 19 a partir del cálculo de espectrograma de un clip de voz, para la extracción de un vector de características, para finalmente aplicar un clasificador de aprendizaje automático, más específicamente una red neuronal convolucional (Aronovich y Hassan, 2020) .	EEUU	2020
Intelligent health monitoring	Incorpora un sistema de inteligencia artificial; el método permite la obtención y análisis de los diferentes síntomas respiratorios de los usuarios por medio de uno o más procesadores. El sistema procesa u opera basado en datos como la temperatura y los registros de los síntomas respiratorios, los cuales son capturados por medio de un video o de grabaciones de voz del usuario, posteriormente el sistema permite determinar si el sujeto necesita una mayor atención, (Anushiravani et al., 2020).	EEUU	2020
A method for analysis of cough sounds using disease signatures to diagnose respiratory diseases	Se identifica un método para diagnosticar una o más enfermedades del tracto respiratorio a partir de sonidos de tos, para producir señales de características de sonido de tos y obtener una o más firmas de enfermedades basadas en las señales de características del sonido de la tos. (Abeyratne y Swarnkar, 2018)	EEUU	2018

---

Nota: Elaboración propia, 2021

### 2.3 Tecnologías de búsqueda abierta

No todas las tecnologías referentes a salud respiratoria se han publicado en artículos científicos o patentes; es importante tener en cuenta el desarrollo de herramientas tecnológicas por parte de empresas o instituciones que presentan sus creaciones ante el mercado como ideas innovadoras y destinadas a brindar una solución a un problema que afecta a la sociedad, algunas de estas se mencionan en la tabla 3.

Tabla 2-3:

## Resultados de tecnologías de búsqueda abierta

Nombre	Descripción	Institución/empresa
End-to-End AIBased Point-of-Care Diagnosis System for Classifying Respiratory Illnesses and Early Detection of COVID-19	Está compuesto por componentes de hardware y software, este registra los síntomas de los pacientes o usuarios, como la temperatura corporal, el sonido de la tos y el flujo de aire, mediante sensores, apoyado por un módulo de inteligencia artificial (IA) para clasificar y diagnosticar diferentes enfermedades respiratorias, incluida la detección temprana de COVID-19, (Nasreddine et al., 2021).	Universidad de los Emiratos Árabes Unidos, Staffordshire University, Universidad de Tecnología de Tianjin
SPIRA	Es un programa de captación de grabaciones de frases, leídas por pacientes con COVID-19. Las muestras de voz reunidas se comparan con las de personas que no padecen la enfermedad para detectar patrones divergentes con programas de inteligencia artificial. Mediante técnicas de aprendizaje de máquinas como los de redes neuronales, el sistema aprenderá a detectar si una persona exhibe insuficiencia respiratoria según la cantidad de pausas, respiraciones y otros patrones presentes en la voz al emitir una frase.(Agencia Iberoamericana para difusión de la ciencia y tecnología, 2020)	Fundación de Apoyo a la Investigación Científica del Estado de São Paulo

---

VocalisCheck	Es una herramienta de detección que puede brindar una puntuación de riesgo confiable para guiar a las personas que pueden necesitar pruebas de diagnóstico para el coronavirus. Un teléfono inteligente u otro dispositivo que usa el software registra a los usuarios y transforma su grabación de voz en una imagen (espectrograma), (Vocalis health, 2021).	Vocalis Health
VocalisTrack	Es una plataforma de voz para detectar la falta de aire, la aplicación permite realizar un seguimiento de los pacientes crónicos en casa mediante la voz. Es una herramienta de software independiente basada en inteligencia artificial que analiza las grabaciones de voz de los pacientes con EPOC que graban su voz a través de una aplicación en su propio teléfono inteligente, (Vocalis health, 2021).	Vocalis Health

---

Nota: Elaboración propia, 2021.

Según la búsqueda realizada se puede observar que hay diversidad de tecnologías como aplicaciones y sistemas inteligentes, desarrolladas para evaluar la función pulmonar, desde una gripe hasta enfermedades pulmonares crónicas, cada herramienta tecnológica posee características únicas y métodos de análisis diferentes, como el análisis de sonidos respiratorios, de tos o de la voz; para el desarrollo de este trabajo se toma como principal enfoque artículos científicos, patentes y tecnologías que a través del análisis de voz detecten síntomas o enfermedades respiratorias.

### 3. Marco teórico

#### 3.1 Anatomía del sistema fonorespiratorio

El sistema fonorespiratorio está conformado por el aparato respiratorio y el aparato fonador; a partir del proceso de la respiración, se genera la voz.

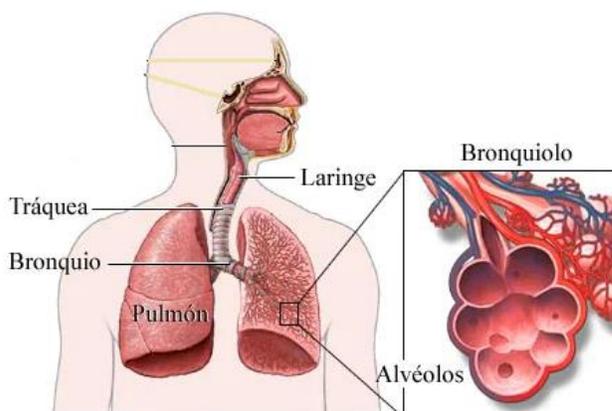
El sistema o aparato respiratorio se divide en vía aérea superior e inferior. La vía aérea superior está conformada por la nariz y fosas nasales, cavidad oral, lengua, y faringe como se ilustra en la figura 1 y la vía aérea inferior está conformada por la laringe, tráquea, bronquios, alvéolos y los pulmones como se muestra en la figura 2.

**Figura 3-1:** *Vía aérea superior*



Nota: Coloma y Álvarez, 2011

**Figura 3-2:** *Vía aérea inferior*



Nota: Banuelos, 2021

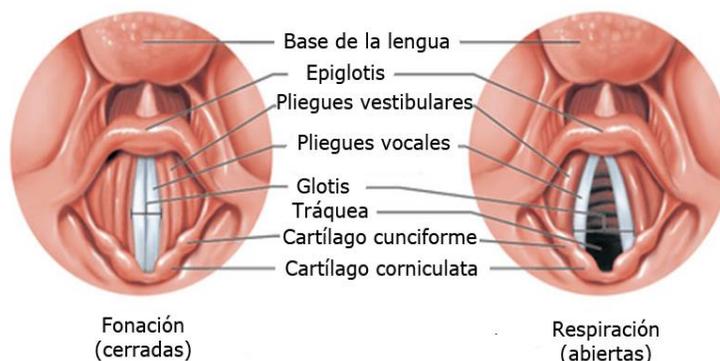
Adicionalmente, algunas estructuras asociadas para que el sistema respiratorio pueda funcionar de forma adecuada son la caja torácica, músculos inspiratorios (diafragma, intercostales, escalenos, pectorales, trapecios y abdominales) y, la pleura y espacio pleural que corresponde a una membrana que recubre ambos pulmones y la pared interna de la caja torácica, (Asenjoa y. Pintob, 2017).

Debido a que la voz humana se produce cuando las cuerdas vocales vibran a causa del aire que se expulsa a través de la laringe, el aparato fonador está conformado principalmente por estructuras del sistema respiratorio. Sin embargo, este se divide en 3 regiones: el vibrador, el fuelle y los resonadores.

El vibrador está compuesto por la laringe, la cual incluye en su interior los pliegues o también llamadas cuerdas vocales o cavidad glótica como se muestra en la figura 3. El fuelle está conformado por todas las estructuras que se encuentran por debajo de las cuerdas vocales (cavidades infraglóticas). Es decir, los pulmones, la caja torácica, el músculo diafragma y los músculos del abdomen.

Por último, los resonadores son las cavidades que se encuentran por encima de las cuerdas vocales: la faringe, la boca y la cavidad nasal (cavidades supraglóticas).

**Figura 3-3:** *Cuerdas vocales durante la fonación - Respiración*



Nota: Torres, 2013

### 3.2 Fisiología del sistema fonorespiratorio

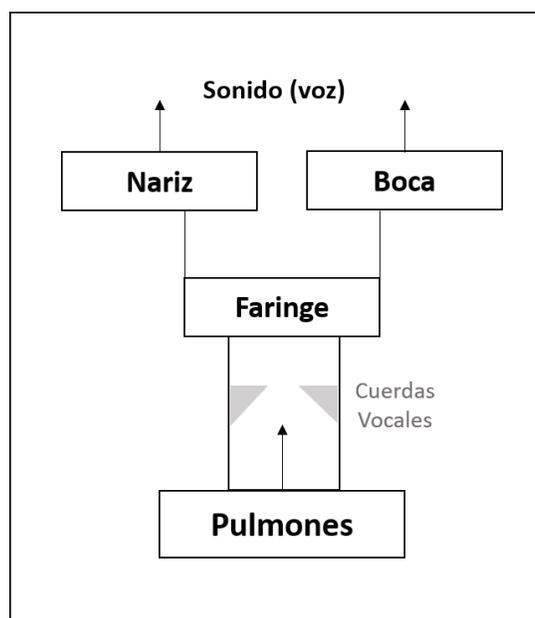
El objetivo principal del sistema respiratorio es permitir la entrada de oxígeno al cuerpo humano y la expulsión de dióxido de carbono como resultado del respectivo metabolismo. Para ello, el aire ingresa a través de la nariz o la boca (inhalación); estas cavidades se unen en la garganta o faringe, la cual sirve para transportar alimentos y también aire. Posteriormente, el aire pasa a través de la tráquea para dirigirse a los bronquios que conectan con los pulmones; los bronquios se ramifican dentro de los mismos para convertirse en bronquiolos y por último, estos terminan en sacos pequeños de aire llamados alvéolos.

De esta manera, ocurre un proceso llamado “difusión” para que el oxígeno pase de los alvéolos al torrente sanguíneo, que después va a fluir al corazón para bombearla a los tejidos que necesitan oxígeno. Por otra parte, el dióxido de carbono, se disuelve en el plasma sanguíneo para regresar al corazón y ser bombeado a los pulmones, pasar a los alvéolos y por último, ser exhalado, (Hirsch, 2019).

Al igual que la relación entre las estructuras del aparato respiratorio y el fonador, los procesos fisiológicos, también funcionan de forma complementaria. Cabe destacar que la voz en parte, es resultado del control consciente e inconsciente de la respiración.

La voz es producida principalmente en la laringe donde se encuentra la glotis, conocido como el órgano de fonación. El aire exhalado, procedente de los pulmones pasa a través de la glotis haciendo vibrar las cuerdas vocales. Mientras tanto las cavidades de la cabeza (nariz y boca), actúan como resonadores haciendo que el sistema respiratorio y nasofaríngeo actúen de forma simultánea como se ilustra en la figura 4., de los resonadores depende la amplificación, brillo y redondez del sonido; siendo dependiente del tamaño y características de las cavidades.

**Figura 3-4:** *Fisiología del aparato fonador*



Nota: Elaboración propia, 2021

El aparato fonador es controlado de forma consciente cuando una persona habla. Algunas características como la intensidad de la voz dependen de la cantidad o fuerza con que se realice la exhalación. En cuanto a la diferencia de la voz entre hombres y mujeres, esto se debe a que las

cuerdas vocales de los hombres son más largas y gruesas que en la mujer, es por ello, que la voz suena más grave (Preparadores de Oposiciones para la Enseñanza, 2005).

### 3.3 Patologías del sistema fonorespiratorio

Las patologías del sistema respiratorio alteran el funcionamiento del aparato fonador debido a que su fisiología depende del control de la respiración. Esta afectación es variable; se puede percibir en algunos casos o, por lo contrario, puede pasar desapercibida.

#### 3.3.1 Patologías del sistema respiratorio

A continuación, en la tabla 4. Se presenta la descripción y síntomas de algunas de las patologías más comunes y contagiosas que afectan el sistema respiratorio.

**Tabla 3-4:**

#### Patologías del sistema respiratorio

Patología	Descripción	Síntomas
Gripe	Es una enfermedad viral e infecciosa de las vías respiratorias. Esta se caracteriza por ser altamente transmitida a través del aire u objetos infectados por las gotas de saliva que se expulsan al hablar, toser o estornudar. A pesar de que los síntomas son conocidos por la mayoría de la población, estos suelen confundirse con los de otros virus, (Sánchez, 2021).	Malestar general, fiebre, dolor de cabeza, flujo nasal, molestias en la garganta y tos seca.
Asma	Es una enfermedad que causa la inflamación de las vías respiratorias. La cual es producida por una interacción entre el sistema nervioso y el sistema inmunológico. Está generalmente, se presenta en episodios donde las vías respiratorias al estar demasiado inflamadas causan irritación y con ello, la obstrucción del flujo de aire al momento de entrar y salir el aire de los pulmones, (Del Río, 2008).	Dificultad respiratoria, sibilancias, tos, opresión torácica.
Neumonía	Según la OMS la neumonía se define como un tipo de	Tos con flema, fatiga, fiebre,

---

	<p>infección respiratoria aguda que afecta a los pulmones, donde los alvéolos están llenos de pus y líquido, lo que hace dolorosa la respiración y limita la absorción de oxígeno. Esta enfermedad puede propagarse a través de gotas de saliva que se expulsan al hablar, estornudar o toser, también puede contagiarse a través de la sangre, (OMS, 2019).</p>	<p>escalofríos, náuseas, dificultad respiratoria.</p>
covid-19	<p>Según la OMS, la COVID-19 es la enfermedad causada por el nuevo coronavirus conocido como SARS-CoV-2. La OMS tuvo noticia por primera vez de la existencia de este nuevo virus el 31 de diciembre de 2019, al ser informada de un grupo de casos de «neumonía vírica» que se habían declarado en Wuhan (República Popular China), (OMS, 2020).</p>	<p>Fiebre, tos seca y cansancio</p>
IRA	<p>La infección respiratoria aguda (IRA) es una infección que afecta a las vías respiratorias superior y/o inferior, provocando pasar de una enfermedad leve a grave y que puede ser transmitida de persona a persona, (Muñoz Retana, 2018).</p>	<p>Tos, dolor de cabeza, fiebre. irritabilidad, ronquera, dolor o secreción de oído, nariz tapada.</p>
EPOC	<p>La enfermedad pulmonar obstructiva crónica (EPOC) según la OMS no es una sola enfermedad sino un concepto general que designa diversas dolencias pulmonares crónicas que limitan el flujo de aire en los pulmones, es prevenible, tratable y conduce a incapacidad e incluso la muerte, (Ministerio de Salud, 2020).</p>	<p>Disnea de esfuerzo (dificultad para respirar), tos crónica, producción regular de esputo, bronquitis o sibilancias frecuentes en invierno.</p>

---

Nota: Elaboración propia, 2021

### ***3.3.2 Patologías del sistema fonador***

Por otra parte, las patologías del sistema fonador son causadas por lesiones o inflamación directamente en las cuerdas vocales. A continuación, en la tabla 5 se presentan algunas patologías frecuentes que causan trastornos en la voz.

**Tabla 3-5:****Patologías del sistema fonador**

<b>Patología</b>	<b>Descripción</b>	<b>Síntomas</b>
Laringitis	La laringitis es la inflamación de la laringe a causa de una infección o uso excesivo de la voz. Al estar las cuerdas vocales en su interior, estas también se inflaman; causando la distorsión de los sonidos cuando el aire pasa a través de ellas, (Mayo clinic, 2021).	Ronquera o pérdida de la voz, tos seca, resequedad, cosquilleo o de la garganta.
Pólipos y nódulos	Los nódulos son un tejido inflamatorio ubicado en una o ambas cuerdas vocales; estos se producen por el uso excesivo de la voz. Si no se tienen los cuidados necesarios, estos pueden seguir creciendo y endurecerse, de esta forma aparecen los pólipos. Al estar ubicados en los extremos de las cuerdas vocales, estos impiden que se cierren completamente, causando trastornos en la voz, (Asociación Estadounidense del Habla, el Lenguaje y la Audición (ASHA), 2021).	Ronquera, voz entrecortada, fatiga corporal, en la garganta y de la voz, disminución en el tono de la voz.
Disfonía por tensión muscular	La disfonía por tensión muscular es producida por una contracción de todos los músculos de la laringe. Produce que parte del aire que se necesita para producir la voz se escape, causando pérdida en la intensidad de la voz y en la disminución en la energía de los armónicos, (Nieto, 2001).	Elevación de 90° entre la mandíbula y el cuello, y contracción cervical.
Parálisis de las cuerdas vocales	Este tipo de parálisis surge a causa de la interrupción de los impulsos nerviosos de la laringe. Esto ocasiona de forma parcial o completa la pérdida de la voz y en algunos casos dificultad para respirar. Puede ser causada por una infección en la zona, daños en los nervios o por células cancerosas. (Mayo clinic, 2021)	Ronquera, pérdida del tono vocal, respiración ruidosa, tos, dificultad respiratoria.

Adicionalmente a las patologías que se pueden presentar de forma frecuente en el sistema fonador, también, existen factores de riesgo que causan la aparición de trastornos en la voz. Entre

ellos se encuentran: Envejecimiento, fumar, consumo de bebidas alcohólicas, gritar, cáncer de garganta, estrés psicológico, entre otros (Mayo clinic, 2021).

### 3.4 Parámetros diagnósticos clínico acústicos

Existen parámetros utilizados para el análisis de sonidos respiratorios y de la voz. Estos son implementados a través de algoritmos para el desarrollo de aplicaciones o sistemas computacionales en la detección de enfermedades respiratorias, del sistema fonador, neurológicas, etc. Desde esta perspectiva, se toma el enfoque de análisis de la voz a través de parámetros de perturbación como jitter y shimmer, parámetros espectrales, temporales y estadísticos.

#### 3.4.1 Parámetros de perturbación de la voz

El jitter es un parámetro de perturbación respiratoria, encargado de representar la diferencia absoluta promedio entre dos períodos consecutivos, tradicionalmente se lo conoce como jitta; al dividir el resultado del procedimiento antes descrito con el periodo promedio se genera un resultado porcentual denominado jitt; tiene un valor de 1,04% como límite umbral para la detección de patologías, (Teixeira et al., 2013). Para su cálculo se aplica el modelo matemático indicado en ecuación 1.

#### Ecuación 1.

$$jitta = \frac{1}{N-1} \sum_{i=1}^{N-1} |T_i - T_{i-1}|$$

Donde N es igual al total de los datos,  $T_i$  es el periodo actual y  $T_{i-1}$  es el periodo anterior de la señal.

El valor obtenido en la ecuación 1, se reemplaza en la ecuación 2 para obtener el jitt.

**Ecuación 2.**

$$jitt = \frac{jitta}{\frac{1}{N} \sum_{i=1}^N Ti} \times 100$$

El parámetro shimmer indica la diferencia absoluta promedio entre las amplitudes de dos periodos consecutivos dividida por la amplitud promedio, el valor porcentual obtenido se denomina shim y se toma el 3,81% como límite para la detección de patologías (Teixeira et al., 2013); su cálculo se realiza a partir de la **ecuación 3**.

**Ecuación 3.**

$$shim = \frac{\frac{1}{N-1} \sum_{i=1}^{N-1} |Ai - A_{i+1}|}{\frac{1}{N} \sum_{i=1}^N Ai} \times 100$$

Donde N es igual al total de los datos,  $A_i$  es la amplitud actual y  $A_{i+1}$  la siguiente amplitud de la señal.

**3.4.2 Parámetros espectrales**

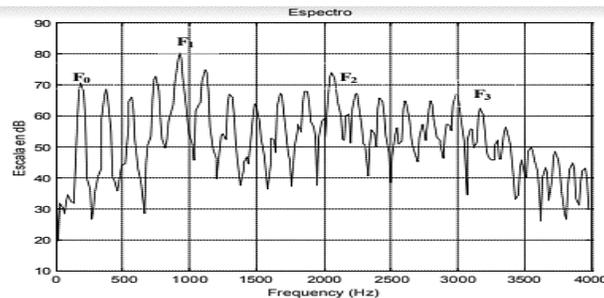
La frecuencia fundamental o pitch es determinada por la velocidad de apertura o cierre de las cuerdas vocales en la laringe durante la fonación de sonidos del tipo sonoro, es decir, es el número de veces que vibran los pliegues vocales por segundo, (Castellanos y Sepúlveda, 2004); la detección de este parámetro surge en el momento en que se trata de encontrar la presencia de las frecuencias con mayor intensidad presentes en la señal de voz, (Silva et al., 2017).

Las formantes están asociadas a las resonancias que surgen a través de las cavidades del tracto vocal. Constituyen un refuerzo en amplitud de los grupos armónicos situados alrededor de una determinada frecuencia; estas permiten identificar sonidos periódicos como vocales y nasales; cabe resaltar que la formante es una cualidad acústica propia de cada sonido y dependiente de la

configuración de su espectro, (Byron, s.f). En la figura 5 se distinguen tres formantes; F1, F2 y F3, así como la frecuencia fundamental.

### Figura 3-5:

#### *Formantes y frecuencia fundamental*



Nota: Navarro Mesa, J. L. s.f

El cepstrum es determinado como la transformada inversa de Fourier del logaritmo de la potencia espectral de la señal, es decir, es la aplicación de dos transformadas de Fourier a la señal de voz. La amplitud del pico de cepstrum proporciona información del grado de periodicidad de la señal de más baja frecuencia en relación con otros componentes periódicos o ruidosos presentes en la señal de voz, (Delgado et al., 2018). Su fórmula matemática característica se presenta en la ecuación 4.

#### **Ecuación 4.**

$$ceps = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log[X(e^{jw})] x e^{jwn} dw$$

Donde  $X$  corresponde a una secuencia a la cual se aplica el logaritmo natural complejo de la transformada de Fourier.

#### **3.4.3 Parámetros temporales**

El parámetro HNR evalúa la relación existente entre los componentes periódicos y el componente no periódico que contiene un segmento del habla sonora. El primer componente

representa la vibración de las cuerdas vocales, y el segundo se deriva del ruido glotal. El HNR incrementa cuanto mayor es el flujo de aire que es expulsado por los pulmones y llevado hacia la energía de vibración de las cuerdas vocales; es decir cuando se tiene un HNR alto significa una voz sonora y armónica, de lo contrario denota una voz asténica y con disfonía, (Teixeira et al., 2013). El HNR se calculó como el promedio de la señal entre la varianza, elevado al cuadrado. Para ello, inicialmente se calcula la varianza a partir de la ecuación 5.

**Ecuación 5.**

$$V = \frac{\sum_{i=1}^{i=n} (X_i - \mu)^2}{n - 1}$$

Donde,  $X_i$  es cada uno de los datos,  $\mu$  es la media de los datos y  $n$  es el total de los datos.

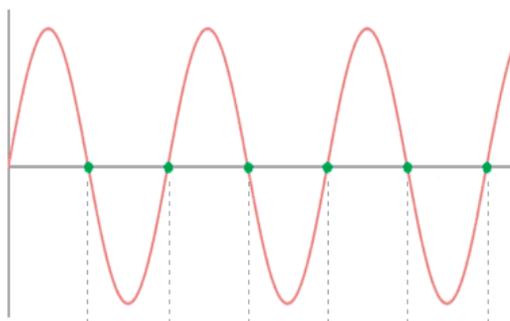
Posteriormente, se remplaza la varianza en la ecuación 6 y se obtiene el HNR.

**Ecuación 6.**

$$HNR = \left(\frac{\mu}{V}\right)^2$$

Donde  $\mu$  es la media de los datos y  $V$  es la varianza obtenida en la ecuación 5.

El parámetro cruces por cero indica el número de veces que una señal continua toma el valor de cero; específicamente se dice que hay un cruce por cero cada vez que la señal cambia de positivo a negativo o viceversa, más ruido en la señal significa mayor número de cruces por cero y menos ruido es equivalente a menor número de cruces, (Martínez, 2015). Para comprender de mejor forma este parámetro, se presenta la figura 6.

**Figura 3-6:***Cruces por cero de una señal*

Nota: Kike, 2018.

**3.4.4 Parámetros estadísticos**

La curtosis es una medida estadística encargada de determinar el grado de concentración que poseen los valores de una variable alrededor de la zona central de distribución, se calcula a partir de la ecuación 7.

**Ecuación 7.**

$$K = \frac{\sum_{i=1}^n (x_i - \mu)^4 f_i}{n s^4} - 3$$

Donde  $x_i$  pertenece al dato actual,  $\mu$  a la media de la señal,  $f_i$  la frecuencia de cada dato,  $n$  es el total de datos y  $s$  la desviación estándar.

Esta medida también es denominada como medida de apuntamiento, es decir informa de lo apuntada (mayor concentración) o lo achatada (menor concentración) que es una distribución, (Marco, 2017).

Por otra parte, la asimetría hace referencia a la curva formada por los valores de una serie de datos, tiene la misma forma a la izquierda y a la derecha con respecto a un valor central, para este caso la media aritmética; este parámetro permite observar a qué lado se agrupan más los datos para definir si existe asimetría positiva o negativa. La asimetría se calcula a partir del coeficiente

de asimetría de Fisher. Esta medida facilita comparar distribuciones y se calcula a partir de la ecuación 8, (Alvarado y Obagi, 2008).

### **Ecuación 8.**

$$AS = \frac{\sum_{i=1}^n (xi - \mu)^3 fi}{ns^3}$$

Donde  $xi$  pertenece al dato actual,  $\mu$  a la media de la señal,  $fi$  la frecuencia de cada dato,  $n$  es el total de datos y  $s$  la desviación estándar.

### **3.5 Indicadores aerodinámicos**

Las pruebas aerodinámicas o también llamadas valoración de la eficacia respiratoria son una prueba objetiva que abarca los conceptos de flujo aéreo y presión de la columna aérea. Los parámetros utilizados en esta prueba son el tiempo máximo de fonación o TMF tras una inspiración profunda, eficacia respiratoria y fonatoria (Jiménez y Espitia, 2016). Cuando se estudia la aerodinámica de la voz significa que se está analizando la eficacia de su manejo; esta resulta de la interacción entre el flujo aéreo y la laringe. La medición del TMF con sonido se realiza a través de la emisión de la vocal /a/, y el TMF sin sonido con la consonante /z/, con el fin de comprobar el control de aire ante una emisión sorda. La puntuación de TMF normal está entre los 25 y 35 segundos para los hombres y de 15 y 25 segundos para las mujeres.

Una función respiratoria anormal puede ser causada por dos razones principales, la primera es que la cantidad de aire utilizada para la fonación está disminuida, esto significa que hay un problema a nivel pulmonar o respiratorio que genera problemas en la intensidad de la emisión, la otra razón es que la fuerza que ejerce el aire sobre la laringe no es la cantidad necesaria (presión subglótica), lo que puede significar que hay una dificultad presente en el cierre glótico debido a la existencia de una lesión, que ocluye los pliegues vocales.

Otra medida aerodinámica es el cociente fonorespiratorio el cual cuantifica la relación entre la función pulmonar y la laríngea; a partir de este índice se puede inferir si la existencia de un TMF bajo es debido a una incompetencia pulmonar o glótica. Este se calcula dividiendo el TMF obtenido sin la intervención de los pliegues vocales durante la emisión, y el TMF máximo que los pliegues vocales permiten con una intensidad y tono confortables, (Garzón y Muñoz, 2018).

Prosiguiendo con la descripción de los indicadores aerodinámicos, está el tiempo Máximo de Espiración (TME). La relación entre el TME y el TMF se recoge a partir del cociente  $s/z$ , el cual se interpreta como un indicador de la eficiencia de cierre glótico, este índice resulta de dividir el tiempo máximo de producción de una consonante /s/ entre el tiempo máximo de fonación para la /a/, (Vila et al, 2011).

### **3.6 Modelo cliente - servidor**

La telemedicina hace referencia a la provisión de servicios de salud a distancia en la promoción, prevención, diagnóstico, tratamiento y rehabilitación realizada por profesionales que utilizan tecnologías de la información y la comunicación, permitiendo el intercambio de datos para facilitar el acceso a la prestación de servicios, principalmente a las personas que posee limitación de acceso debido al área geográfica en la que se encuentran, (Resolución 2654 de 2019).

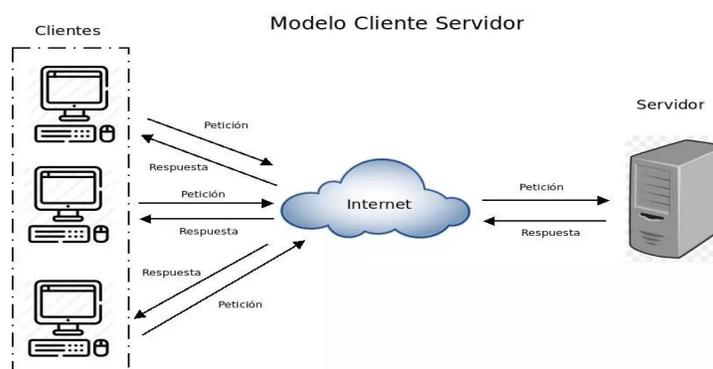
La comunicación entre computadoras se rige por el llamado modelo cliente-servidor, el cual provee usabilidad, flexibilidad, interoperabilidad y escalabilidad en las comunicaciones; este modelo reparte tareas entre los proveedores de recursos denominados servidores y los demandantes llamados clientes; estos últimos son quienes realizan las diferentes peticiones a uno o varios servidores que se encuentran en ejecución. Tanto el cliente como el servidor son entidades que pueden residir en la misma o en diferentes ordenadores.

Miles de usuarios buscan información en internet, este último lo adoptan como una parte esencial de la vida cotidiana, por lo que solo basta tener una computadora para obtener información, consultar páginas web, entablar comunicación, transferir archivos y programas entre computadoras lejanas, situadas incluso al otro lado del planeta como se muestra en la figura 7. Una computadora conectada a internet debe contar con el protocolo TCP/IP (Protocolo de Control de Transmisión/Protocolo de Internet), para que sea posible la interacción entre los usuarios y la red de redes; este protocolo está diseñado para manejar la información en paquetes, debido a la gran versatilidad de comunicación que hay entre dos computadoras.

El usuario de una computadora se convierte en cliente en el momento en que accede a una página web y solicita información a un servidor; para que el usuario pueda recibir esta, es necesario contar con ciertos datos para que la computadora sea localizada. Estos datos están contenidos en la dirección IP, (Estrada, 2004).

**Figura 3-7:**

*Modelo cliente servidor*



Nota: Schiaffarino A, 2019

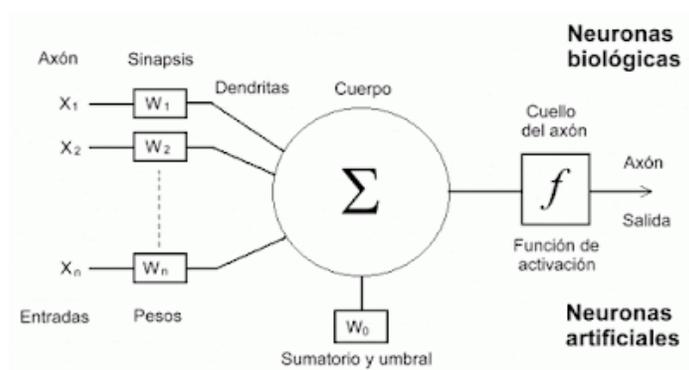
El protocolo HTTP (Hypertext Transfer Protocol) es el protocolo TCP/IP utilizado para transferir documentos de hipertexto a través de internet, por esta razón se define un conjunto de reglas necesarias para cada archivo enviado, ya sea texto, imágenes, sonido, videos, etc. El cliente

realiza una petición al servidor y este último le responde con el contenido solicitado; al ingresar a una página web, el navegador lleva a cabo una petición HTTP al servidor, quien responde con el código HTML de la página, una vez el cliente recibe el mensaje, este lo interpreta y lo muestra en pantalla, (Niño, s.f).

### **3.7 Redes neuronales**

Las redes neuronales se componen por un grupo de neuronas simuladas, las cuales se consideran análogas a las neuronas biológicas; estas se encuentran interconectadas y son capaces de aprender de la forma en cómo lo hacen las personas.

El componente mínimo de una red neuronal es una neurona, la cual se considera un elemento capaz de transformar varias señales de entrada en una única salida; las entradas pueden ser procedentes de otras neuronas o ser entradas externas a la red, lo mismo ocurre con las salidas, estas pueden transmitirse a otras neuronas o funcionar como señal de salida de la red. Las señales de entrada son moduladas por un factor denominado peso ( $W$ ), este gradúa la importancia de la conexión existente entre las neuronas. Al considerar a las neuronas como un elemento de procesamiento simple, está restringida a instrucciones elementales como sumas y productos, por lo que la actividad de una unidad es la suma ponderada de las actividades de las unidades con las que está interconectada, multiplicada por los pesos correspondientes, como se indica en la figura 8.

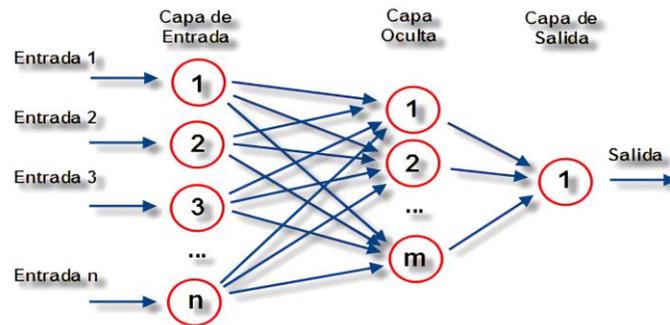
**Figura 3-8:***Estructura de una red neuronal artificial*

Nota: German, 23015.

También, se indica la función de activación  $g$ , existen diversas funciones de activación, tienen el objetivo de propagar la salida de los nodos de una capa hacia la capa siguiente, estas producen la activación de la neurona y permiten incorporar el modelado de datos de entrada no lineales a la red neuronal. Las funciones más utilizadas son: función lineal, sigmoide, tangente hiperbólica y función escalón (clasificación binaria).

### 3.7.1 Arquitectura red neuronal

Las neuronas están agrupadas en capas, y las capas están interconectadas. Existen tres tipos de capas: 1. Capa de entrada: La constituyen los datos del problema, 2. Capas intermedias: puede existir más de una, dependiendo de la complejidad del problema que se quiera solucionar con la red, mediante el tratamiento de estas capas resultan propiedades de generalización, extracción de características, etc. 3. capa de salida: el número de neuronas de esta capa depende del formato esperado de salida de la red (Pino et al., 2001), como se indica en la figura 9.

**Figura 3-9:***Arquitectura red neuronal*

Nota: Wikipedia Academic, 2010.

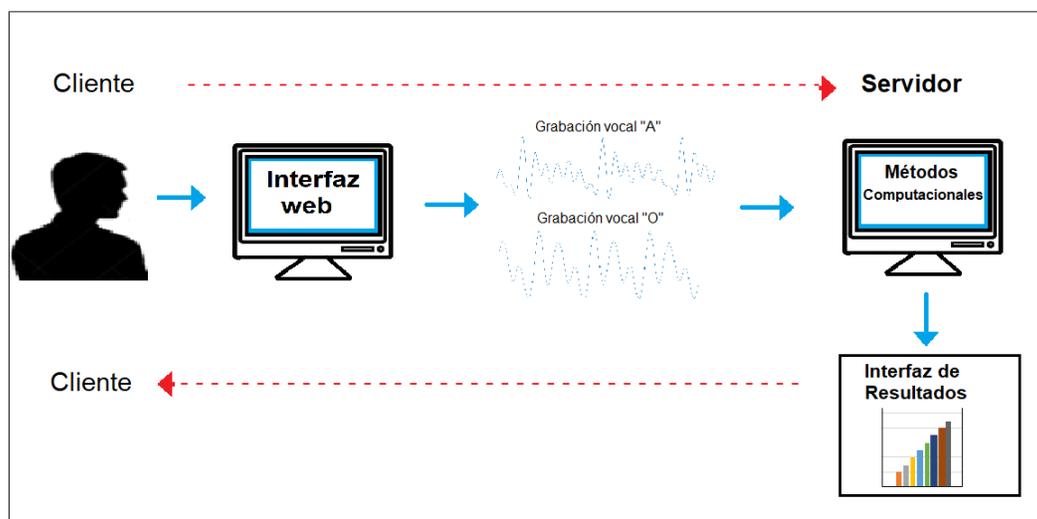
## 4. Materiales y métodos

### 4.1 Componentes del sistema

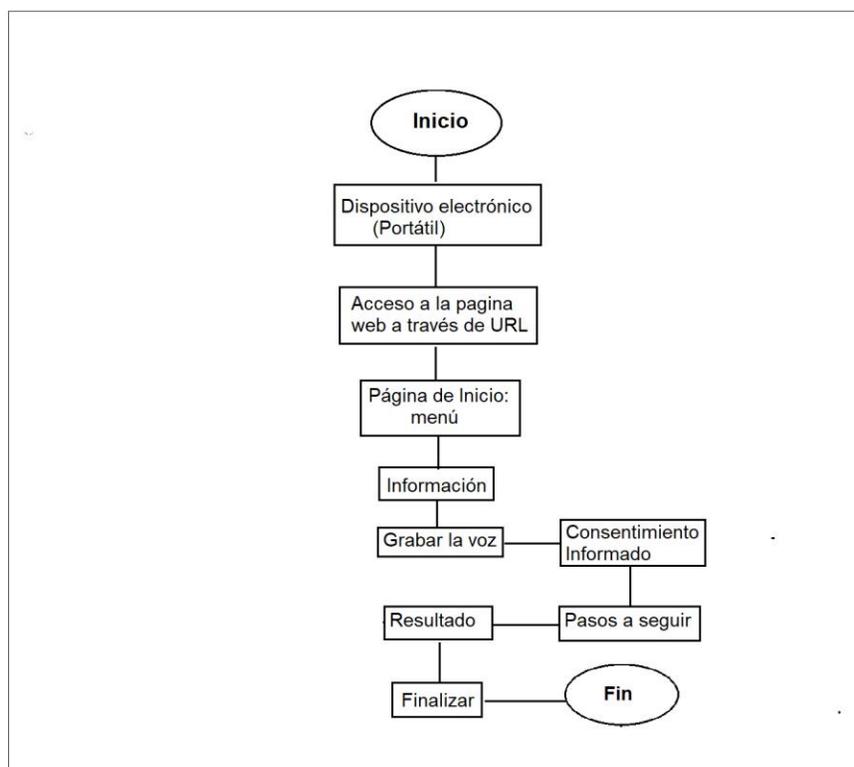
Inicialmente, el usuario ingresa a un sistema web desde su dispositivo electrónico (celular, computador o Tablet), para realizar la grabación de un sonido vocálico sostenido. Posteriormente, a través de la arquitectura cliente - servidor, las señales de voz se dirigen hacia un servidor donde se realizan métodos avanzados de procesamiento digital de señales, los cuales permiten el análisis de señales de voz. Adicionalmente, el sistema incorpora inteligencia computacional para detectar de forma automática y no invasiva, la presencia de dificultad respiratoria. En la figura 10 se ilustra de forma gráfica los componentes y funcionamiento del sistema en general, y en la figura 11. Se observa el diagrama de flujo de operación del sistema desde la perspectiva del usuario.

#### Figura 4-10:

*Representación diagramática de componentes y funcionamiento general del sistema*



Nora: Elaboración propia, 2021

**Figura 4-11:***Operación del sistema*

Nota: Elaboración propia, 2021

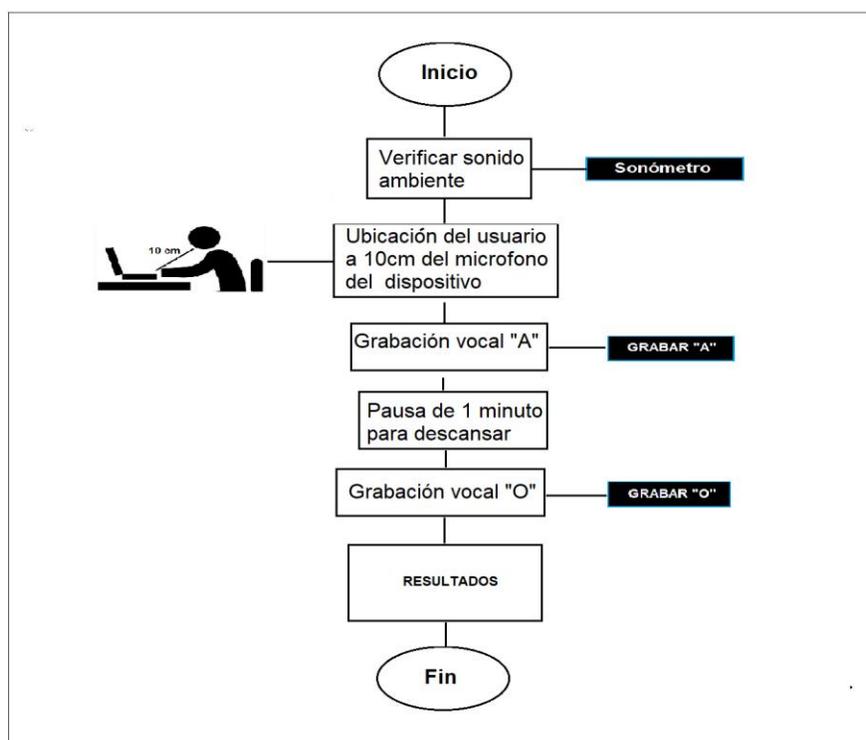
#### **4.1.1 Adquisición de señales de voz**

Para la adquisición de señales de voz se realizó la grabación de dos sonidos vocálicos sostenidos; vocal “a” y “o”, a 2 mujeres y 3 hombres. Las grabaciones se realizaron a través del dispositivo móvil Huawei Y9 2019, el cual cuenta con cancelación de ruido con micrófono dedicado. En este estudio, participaron personas hispano hablantes, entre los 20 y 49 años de forma voluntaria, en condición sana y patológica. Es decir, se realizó la grabación de la “a” y “o” cuando fueron diagnosticados con gripe a través de una evaluación clínica y posteriormente, se capturó la voz en condición sana de los mismos participantes. Cabe resaltar, que se presentó un consentimiento informado, donde el usuario acepta participar de forma libre, consciente y

voluntaria, para el análisis y donación de su voz. A la vez, se cumplió con el protocolo que se ilustra en la figura 12. para obtener las señales con el menor ruido posible y en las mejores condiciones; se verificó con una aplicación móvil “sonómetro” que el ruido ambiente estuviera por debajo de 40 dB; seguidamente, el usuario debía ubicarse a 10 cm de distancia del micrófono incorporado en el dispositivo electrónico, y, por último, mencionar de forma sostenida los sonidos vocálicos (uno a la vez).

**Figura 4- 12:**

*Protocolo para capturar la voz*



Nota: Elaboración propia, 2021

Finalmente, al realizar un ventaneo de 150 ms de las señales colectadas, se obtuvieron 208 muestras para mujeres entre la vocal “a” y “o” en condición gripal y también, 208 muestras en condición sana para ambos sonidos vocálicos. Por otra parte, para los hombres se obtuvieron

312 muestras entre vocal “a” y “o” en condición gripal y 312 muestras en condición sana para los mismos sonidos vocálicos, para un total de 1040 muestras entre hombre y mujeres.

#### **4.1.2 Método computacional**

El análisis de las grabaciones de voz de cada paciente se realizó a partir de algoritmos automatizados, implementados en Matlab como se muestra en la figura 13, que involucra el cálculo de diferentes parámetros en la señal de voz; el proceso inicia con la recepción de la señal de voz y su respectiva normalización a través del teorema de Tales como se presenta en la ecuación 9.

#### **Ecuación 9.**

$$z = 2 * \left( \frac{X - Xmin}{Xmax - Xmin} \right) - 1$$

Donde X son los datos de la señal, Xmin el valor mínimo de los mismos y Xmax el valor máximo.

Es importante resaltar que de cada grabación se tomaron 4 segundos específicamente, capturados del cuerpo de la señal, debido a que en este segmento hay mayor estabilidad de la grabación de voz, a una frecuencia de muestreo de 44100 Hz.

Una vez definido el tiempo de duración de la señal y su frecuencia se procede a la implementación de un ventaneo con una duración de 150 ms y una superposición del 50%. Para cada ventana se realizó el cálculo de la frecuencia fundamental con el objeto de discriminar si la señal de voz es de hombre o mujer.

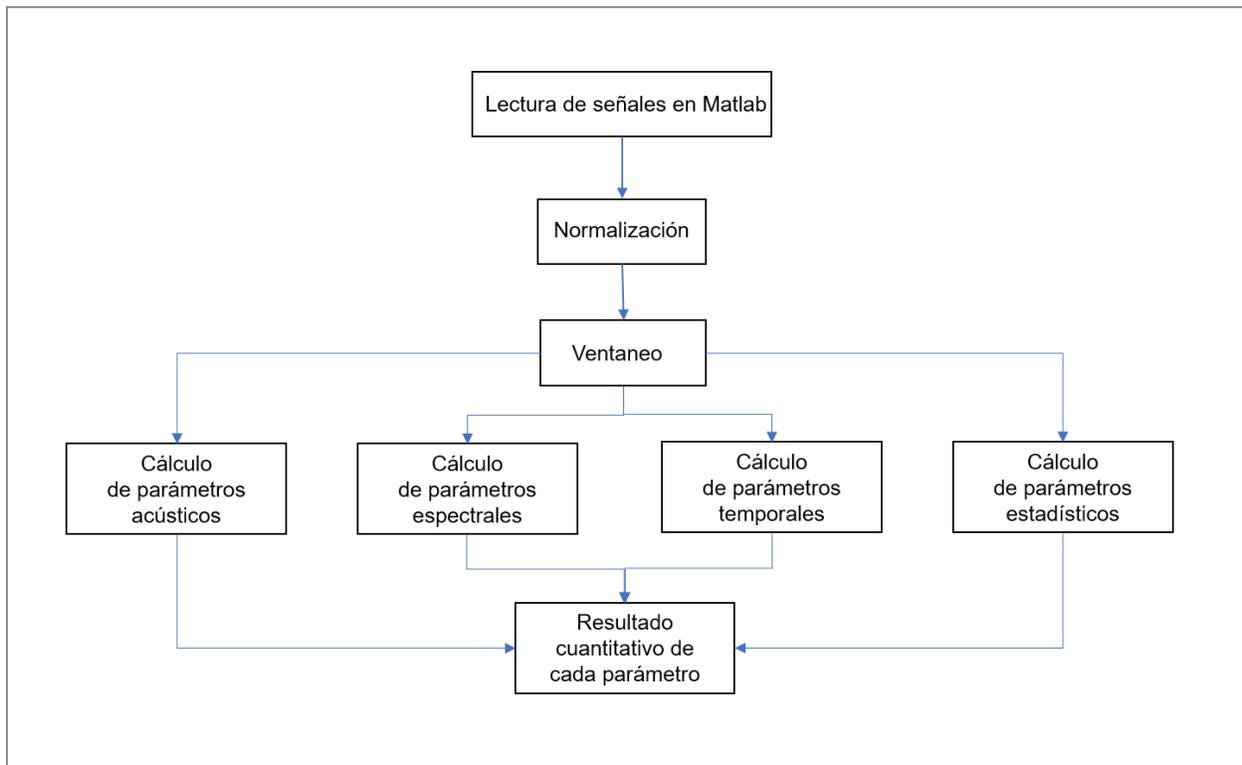
Posteriormente, se lleva a cabo el cálculo de parámetros de perturbación a partir de modelos matemáticos ya establecidos; dentro de estos se encuentra el jitter el cual evalúa la variación de los periodos de la señal, y el shimmer que indica los cambios de amplitud.

Se realiza una evaluación espectral sobre la señal de voz a través del parámetro cepstrum y el cálculo del primer formante, y un análisis temporal a través de un algoritmo de extracción

del número de cruces por cero de la señal y la aplicación del HNR o relación armónica - ruido. Además, se complementa el estudio con parámetros estadísticos como la asimetría y curtosis.

**Figura 4-13:**

*Procesamiento digital de señales de voz*



Nota: Elaboración propia, 2021

Todos los parámetros mencionados anteriormente son fundamentales para la distinción entre voz sana y voz con dificultad respiratoria, sin embargo, algunos valores de parámetros aportan más a esa diferenciación que otros, por lo que se decide realizar una selección específica.

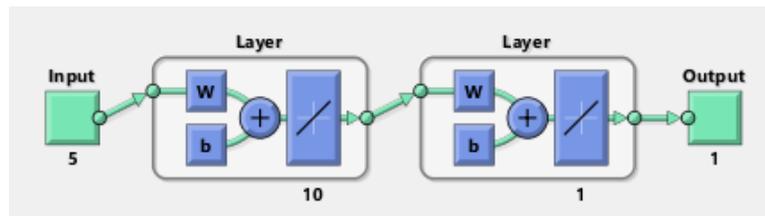
Finalmente, a través del método computacional se obtuvo un vector que contiene los valores resultantes de cada parámetro por ventana, formando una matriz por señal; estas fueron utilizadas como descriptores de entrada para el entrenamiento, prueba y validación de la red neuronal.

### 4.1.3 Red neuronal

Se implementó una red neuronal de reconocimiento de patrones utilizando una red perceptrón multicapa (PMC) de gradiente conjugado escalado, con el fin de clasificar voces sanas y patológicas, con presencia de dificultad respiratoria; esta funciona con entrenamiento supervisado, es decir, para cada entrada se realiza una comparación entre la salida obtenida por la red con la deseada. Por esta razón, el ajuste de los pesos trata de minimizar al máximo la diferencia entre estas salidas. En la figura 14 se presenta la arquitectura de la red implementada, la cual consta de 5 entradas (Jitter-shimmer-formante-curtosis-asimetría), una capa intermedia con diez neuronas y una capa de salida con una neurona, esto con el fin de obtener una salida binaria que representa presencia de dificultad respiratoria (1) o ausencia de esta (0). En cuanto a las funciones de activación, se aplicó la función rampa linear tanto para la capa oculta y de salida.

**Figura 4-14:**

*Estructura red reconocimiento de patrones*



Nota: Creada en Matlab, 2021

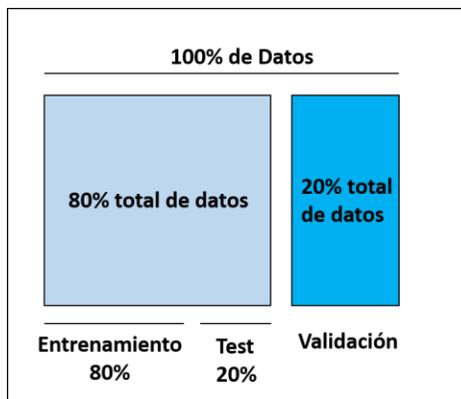
Cuando se hace referencia a una red neuronal se debe tener en cuenta la palabra optimización que equivale a minimizar la función de error global. Generalmente, se utilizan algoritmos de gradiente descendente para optimizar el sistema, sin embargo, estos tienen un índice de convergencia pobre. El algoritmo del gradiente conjugado escalado (SCG) está basado en la estrategia general de optimización, con el objetivo de acelerar la convergencia. En el algoritmo de entrenamiento se lleva a cabo una búsqueda a lo largo de direcciones conjugadas a través de un

mecanismo para decidir qué tan lejos ir en una dirección determinada, generando una convergencia más rápida; la dirección que surge del entrenamiento se restablece consecutivamente a la negativa del gradiente para que no exista la acumulación de errores en las aproximaciones.

Para la red neuronal se utilizaron dos conjuntos de datos diferentes; uno para hombres y otro para mujeres, debido a la diferencia que se presenta en los parámetros espectrales, principalmente en la frecuencia fundamental y primera formante. Para ello, se separó el 80% del total de datos para nuevamente dividirlo en entrenamiento y prueba, y el 20% para la validación como se muestra en la figura 15. Cabe resaltar que la red se ejecutó a través de validación cruzada para verificar la precisión del clasificador evitando overfitting, es decir, problemas de generalización en la red.

**Figura 4-15:**

*Agrupamiento de datos de entrenamiento, prueba y validación*



Nota: Elaboración propia, 2021

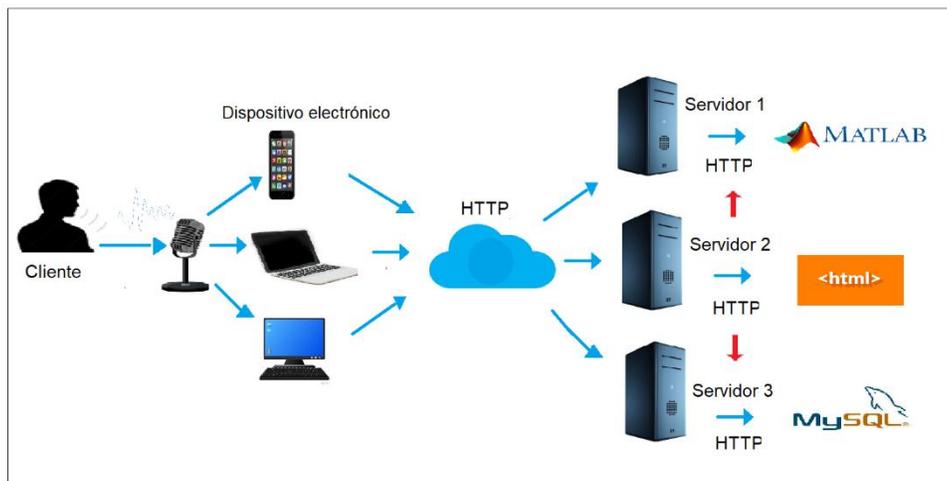
**4.1.4 Arquitectura sistema cliente servidor**

El sistema funciona bajo una arquitectura Cliente – Servidor como se muestra en la figura 16; existe un dispositivo cliente, tal como un smartphone, Tablet, portátil o computador de escritorio, que ejecuta cualquier navegador web como cliente. Por otra parte, el servidor es un computador que tiene incorporado la instalación del software de servidor (Matlab).

La comunicación entre el cliente y el servidor se implementó utilizando el protocolo HTTP o HTTPS, sobre un servicio de internet. En cuanto al software de servidor, este sistema utiliza el servidor Matlab Web Apps Server, para ejecutar el procesamiento digital de la señal de voz, la extracción de sus parámetros, y los algoritmos de inteligencia computacional. Adicionalmente, el sistema utiliza un segundo servidor para ejecutar toda la interfaz web, la grabación de la voz y presentación de resultados al usuario. El segundo servidor ejecuta parte del sistema mediante los lenguajes HTML y CSS. La interacción entre los dos servidores se realiza utilizando el protocolo HTTP y el Lenguaje HTML, ya que este permite integrar las dos interfaces de ambos servidores e intercambiar datos entre ellos. Por último, el sistema utiliza un tercer servidor para almacenamiento de datos y las señales de voz, este servidor está implementado utilizando MySQL y se comunica con los otros servidores vía HTTP.

**Figura 4-16:**

*Arquitectura cliente – servidor*



Nota: Elaboración propia, 2021

#### ***4.1.5 Sistema web, captura de señales y transmisión***

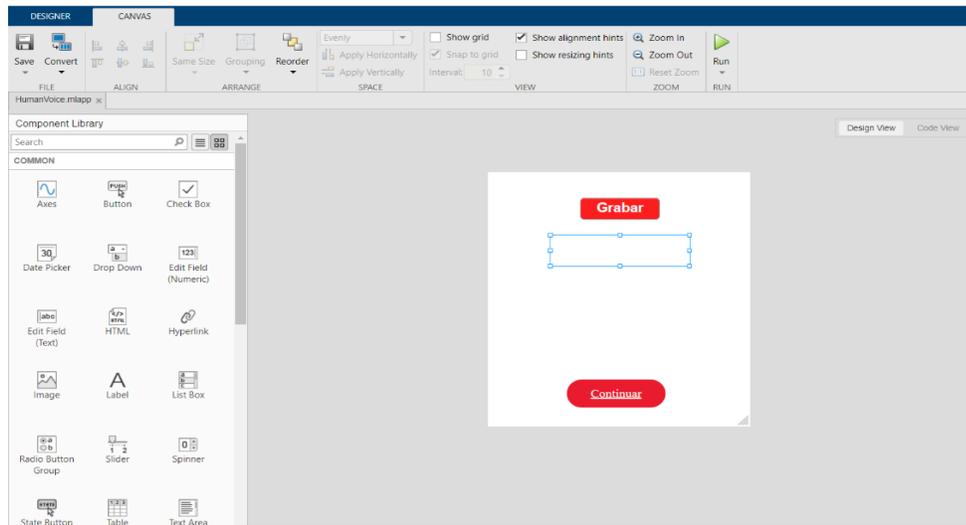
Para la implementación del sistema web, la página se realizó en lenguaje HTML y CSS, utilizando un template que permite visualizar una interfaz de inicio, páginas de información concreta sobre el equipo de trabajo y, sobre la respiración y la voz. Posteriormente, se montó en el servidor de Matlab web App server y se diseñó una interfaz utilizando el App designer de Matlab, donde se ubicaron los botones de “Grabar”, “continuar” como se muestra en la figura 17. Adicionalmente, se llamó una función para extraer los parámetros de grabación, otra para la clasificación de sexo y, por último, una función para identificar la dificultad respiratoria.

Se estableció realizar la clasificación de sexo debido a que la red neuronal es diferente para hombres y mujeres, de acuerdo a los valores y patrones de los parámetros. Por esta razón, una vez se clasifique el sexo, se dirige a una determinada red neuronal donde finalmente, ya se identifica si hay presencia de dificultad respiratoria o no.

Una vez realizada la interfaz y llamadas las funciones, se empaquetó y se compiló como aplicación del web App Server para finalmente incorporarla en el lenguaje HTML como una URL, en este caso como un botón llamado “Grabar voz”.

**Figura 4-17:**

*Interfaz en el App designer de Matlab*



Nota: Elaboración propia a través de App designer Matlab, 2021

## 5. Resultados

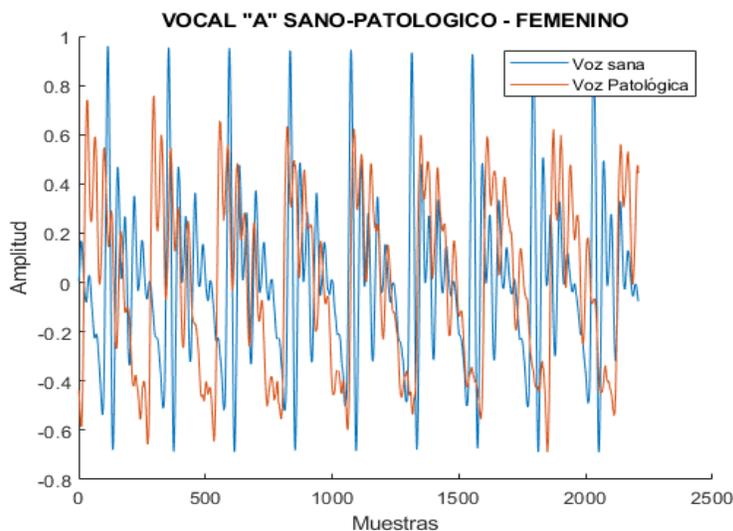
### 5.1 Diferencias gráficas sano - patológico

Al realizar el procesamiento digital de señales de voz colectadas, se obtuvieron los valores numéricos de cada parámetro acústico, espectral, temporal y estadístico. Adicionalmente, se realizó la comparación gráfica que a simple vista permite identificar características morfológicas de la señal sana y patológica, del mismo paciente, sexo y vocal.

Se realizó la comparación grafica sano-patológico de las señales raw en ventanas de 50 ms equivalente a 2205 muestras para observarse de forma más detallada, cabe resaltar que el ventaneo para la extracción de parámetros y clasificación se realizó con 150ms. En la figura 18. Se presenta el gráfico comparativo del sonido vocálico “A” de una mujer voluntaria. En esta imagen se puede identificar que la amplitud pico a pico de la voz sana presenta menor ruido, y una amplitud mayor y constante en comparación con la voz patológica

#### Figura 5-18:

*Vocal “a”, voz sana - patológica (Paciente mujer)*

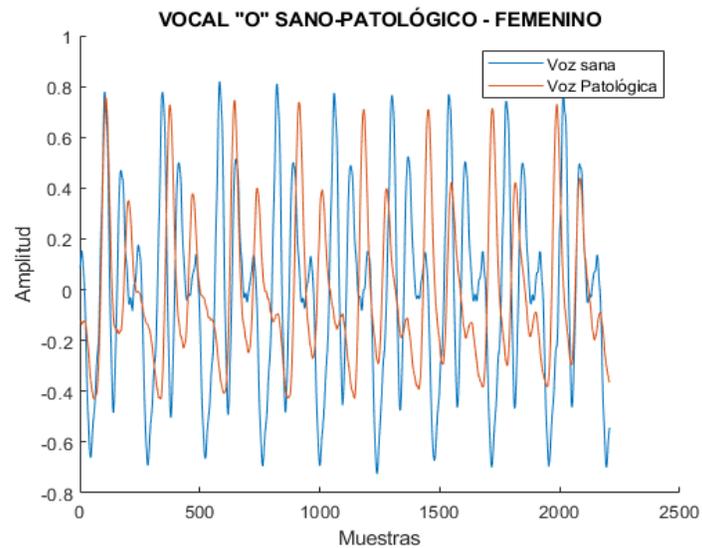


Nota: Visualización de señales a través del software Matlab, 2021

Posteriormente, en la figura 19 se ilustra el gráfico comparativo del sonido vocálico “o” de una mujer voluntaria. Se puede observar que la señal de voz patológica se encuentra en desfase, al igual que con una amplitud pico a pico menor, y con un ciclo glótico (cada pico es un ciclo) menos en comparación a la voz sana.

**Figura 5-19:**

*Vocal “o”, voz sana - patológica (Paciente mujer)*

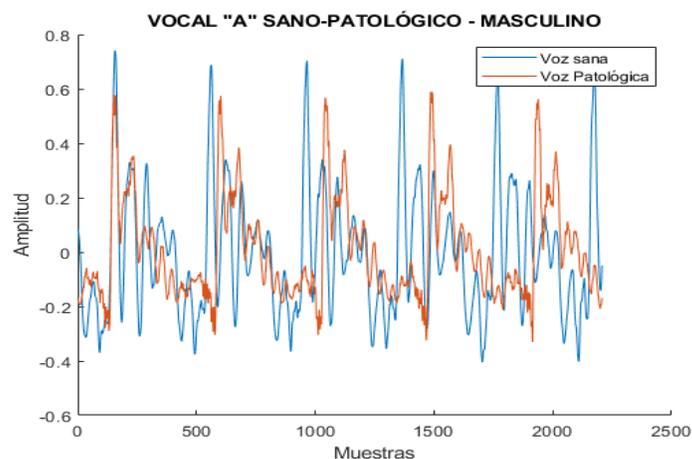


Nota: Visualización de señales a través del software Matlab, 2021

Por otra parte, en la figura 20 se observa el gráfico comparativo del sonido vocálico “a” de un hombre voluntario. Se puede identificar, que la señal de voz patológica tiene una amplitud menor e inestable en comparación a la sana. También, se observa que la señal patológica presenta mayor ruido e inestabilidad de la señal en general.

**Figura 5-20:**

*Vocal "a", voz sana - patológica (Paciente hombre)*

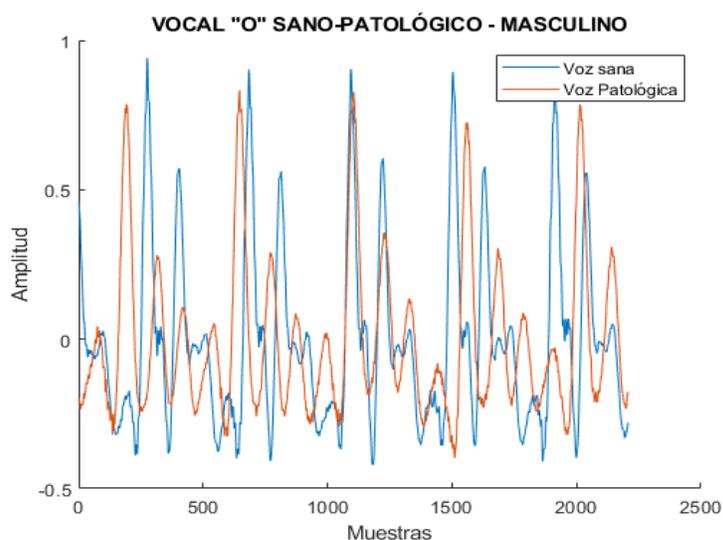


Nota: Visualización de señales a través del software Matlab, 2021

Por último, en la figura 21 se muestra el gráfico comparativo del sonido vocálico "o" de un hombre voluntario, donde se puede identificar que la señal de voz patológica presenta una amplitud menor e inestabilidad en relación a la voz en condición sana.

**Figura 5-21:**

*Vocal "o", voz sana - patológica (Paciente hombre)*



Nota: Visualización de señales a través del software Matlab, 2021

## 5.2 Resultados de parámetros en las señales

Al obtener los valores resultantes de cada parámetro en cada señal de voz, se realiza una serie de análisis para identificar la diferencia existente entre voz sana y con presencia de dificultad respiratoria de un mismo paciente, en la tabla 6 y 7 se presenta la diferencia porcentual entre cada parámetro sano y patológico obtenido del promedio de todas las ventanas de 150 ms, derivados del análisis de la vocal ‘a’ y ‘o’ respectivamente, tanto para hombres (M) como para mujeres (F).

**Tabla 5-6:**

### Diferencias porcentuales (%) sano - patológico en vocal a

Vocal 'a'	Jitter	Shimmer	Cepstrum	Formante	Curtosis	Asimetría	Cruce ceros	HNR
<b>Paciente 1 (M)</b>	12.80	18.22	29.55	13.59	23.14	44.39	26.36	84.42
<b>Paciente 2 (F)</b>	18.90	51.46	19.07	48.23	32.54	23.92	11.54	76.87
<b>Paciente 3(M)</b>	23.96	25.68	25.54	12.17	17.35	52.74	15.14	71.87
<b>Paciente 4(M)</b>	18.82	41.17	21.98	17.78	24.17	17.49	9.20	74.39
<b>Paciente 5(F)</b>	81.99	31.90	21.99	42.22	39.90	81.01	49.15	74.01

Nota: Elaboración propia, 2021

**Tabla 5-7:****Diferencias porcentuales (%) sano-patológico en vocal o**

Vocal 'o'	Jitter	Shimmer	Cepstrum	Formante	Curtosis	Asimetría	Cruce ceros	HNR
Paciente 1 (M)	58.69	19.04	26.39	18.51	24.22	21.68	22.31	87.2 7
Paciente 2 (F)	33.79	37.35	20.83	6.17	9.44	53.14	10.64	83.03
Paciente 3(M)	21.05	42.45	28.26	6.33	12.23	77.51	19.66	74.41
Paciente 4(M)	4.60	36.93	22.22	2.76	4.62	19.17	23.09	72.94
Paciente 5(F)	49.48	29.23	18.80	13.21	14.30	59.60	25.43	71.2 3

Nota: Elaboración propia, 2021

En la tabla 8, se presenta el valor promedio de las ventanas de 150 ms de cada parámetro del procesamiento digital de la señal “a” y “o” en hombres (M) y mujeres (F).

**Tabla 5-8:****Media aritmética voz sana - patológica**

Vocal 'a'	Jitter	Shimmer	Cepstrum	Formante	Curtosis	Asimetría	Cruce ceros	HNR	
Paciente 1 (M)	Sano	23,94	115	0,76	560,62	3,03	0,57	156	0,002 55
	Patológico	23,06	110,10	1,09	523,96	3,96	1,03	114	0,000 51
Paciente 2 (F)	Sano	16,45	74,55	0,93	588,96	2,68	0,50	259	0,000 35

	Patológico	15,86	115,11	1,017	467,88	3,98	0,57	293	0,000 27
<b>Paciente 3(M)</b>	Sano	32,86	76,15	1,076	642,67	3,16	0,48	219	0,000 16
	Patológico	26,76	80,51	0,94	564,23	3,85	1,02	205	0,000 21
<b>Paciente 4(M)</b>	Sano	26,35	65,08	0,84	615,90	2,76	0,423	193	0,000 09
	Patológico	30,08	110,29	0,85	506,06	2,09	0,36	186	0,000 16
<b>Paciente 5(F)</b>	Sano	23,33	86,32	0,82	364,14	3,28	0,47	294	0,000 09
	Patológico	3,79	87,55	0,84	229,33	1,96	0,09	147	0,000 15
<b>Vocal 'o'</b>		Jitter	Shimmer	Cepstrum	Formante	Curtosis	Asimetría	Cruce ceros	HNR
<b>Paciente 1 (M)</b>	Sano	33.61	90.38	0.92	259.92	3.61	1.16	127.57	7.6e- 05
	Patológico	13.81	75.19	0.76	307.19	4.77	1.48	100.63	0.001 62
<b>Paciente 2 (F)</b>	Sano	21.88	62.46	0.75	376.44	2.84	0.42	145.71	0.000 18
	Patológico	15.02	80.42	0.78	393.52	3.14	0.19	142.26	0.002 58
<b>Paciente</b>		27.80	44.54	0.84	347.50	2.58	0.19	141.05	

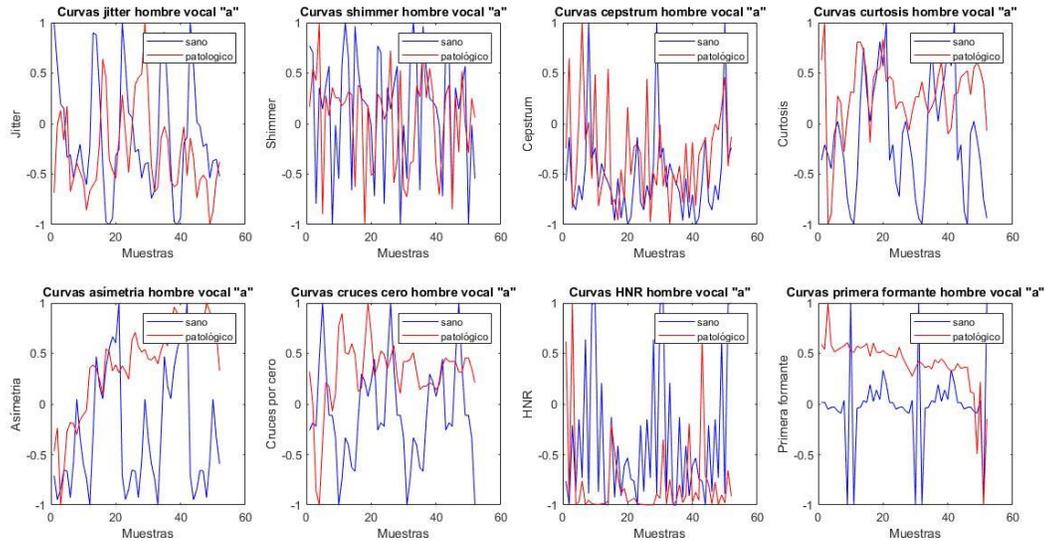
<b>3(M)</b>									0.000
	Sano								1
	Patológico	33.07	26.22	0.91	366.82	2.61	0.85	112.07	0.000 10
	Sano	32.19	14.12	0.77	384.77	2.35	0.76	110	0.000 14
<b>Paciente</b>									
<b>4(M)</b>	Patológico	32.02	22.84	0.73	380.42	2.24	0.61	142.98	0.000 10
	Sano	20.78	111.34	0.65	459.98	2.37	0.22	145.92	0.000 15
<b>Paciente</b>									
<b>5(F)</b>	Patológico	20.19	104.757	0.69	426.48	2.78	0.59	109.38	0.000 16

Nota: Elaboración propia, 2021

Adicionalmente, en la figura 22, 23, 24 y 25 se presenta las gráficas comparativas entre señal sana y patológica para cada uno de los parámetros utilizados, la comparación se realizó por sexo y por vocal. Se observan diferencias significativas entre las curvas características que permiten identificar la variación en escala y forma según la presencia de dificultad respiratoria o no.

**Figura 5-22:**

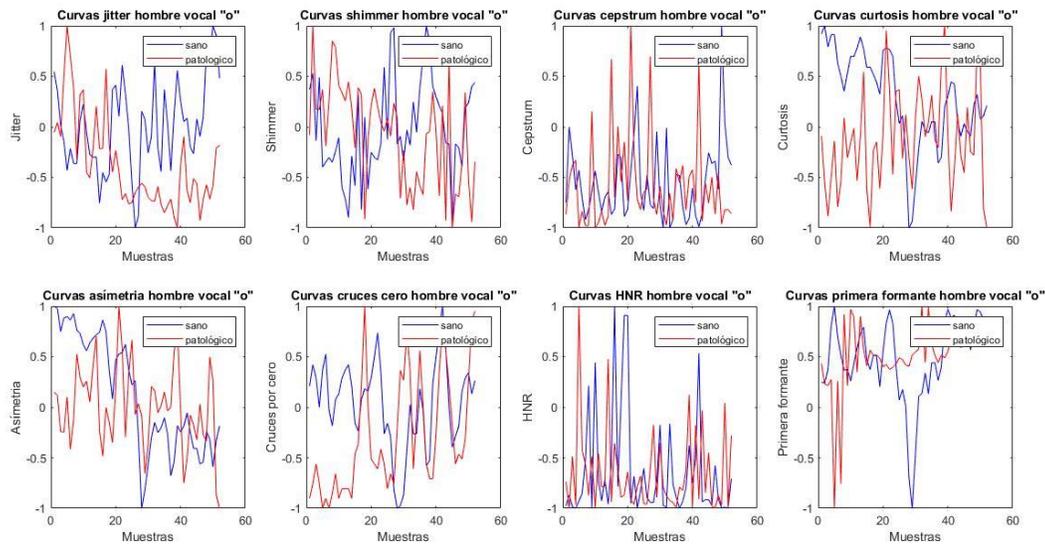
*Comparación entre la señal de la vocal "a" sana y patológica para cada parámetro en hombres.*



Nota: Visualización de señales a través del software Matlab, 2021

**Figura 5-23:**

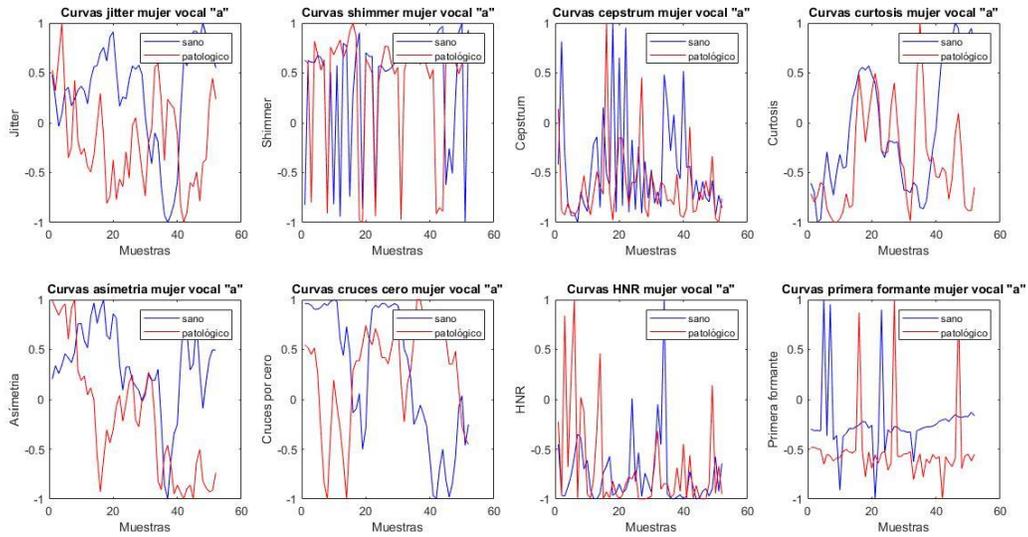
*Comparación entre la señal de la vocal "o" sana y patológica para cada parámetro en hombres.*



Nota: Visualización de señales a través del software Matlab, 2021

**Figura5- 24:**

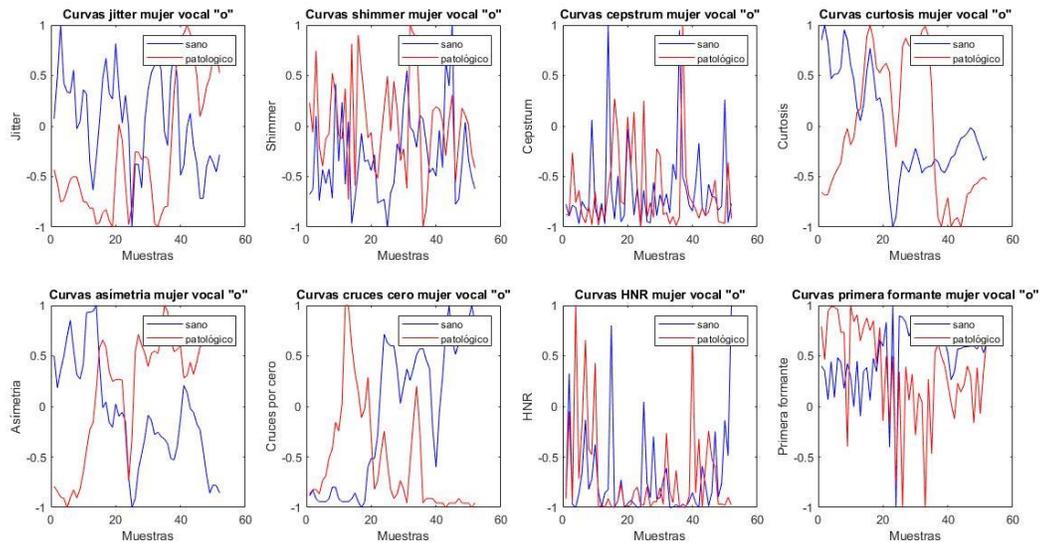
*Comparación entre la señal de la vocal "a" sana y patológica para cada parámetro en mujeres.*



Nota: Visualización de señales a través del software Matlab, 2021

**Figura 5-25:**

*Comparación entre la señal de la vocal "o" sana y patológica para cada parámetro en mujeres.*



Nota: Visualización de señales a través del software Matlab, 2021

Por otra parte, teniendo en cuenta que la desviación estándar indica la dispersión de los datos con respecto a la media; en la tabla 9, se presenta este parámetro estadístico aplicado a las señales sanas y patológicas en cada parámetro; según los resultados obtenidos la dispersión de los datos de la señal patológica es mayor que la sana, tanto en hombres como en mujeres; por lo que una voz patológica tiene mayor inestabilidad debido a que las enfermedades respiratorias se consideran la causa más común de alteración en la voz.

**Tabla 5-9:**

**Desviación estándar voz sana-voz patológica**

<b>Vocal 'a'</b>		<b>Jitter</b>	<b>Shimmer</b>	<b>Cepstrum</b>	<b>Formante</b>	<b>Curtosis</b>	<b>Asimetría</b>	<b>Cruce ceros</b>	<b>HNR</b>
<b>Paciente 1 (M)</b>	Sano	1,73	18,14	0,18	89,03	0,15	0,07	12.88	0,00271
	Patológico	3,79	23,89	0,33	82,24	0,18	0,07	19.44	0,00097
<b>Paciente 2 (F)</b>	Sano	3,76	22,22	0,22	296,64	0,10	0,11	12.04	0,00039
	Patológico	2,14	56,00	0,20	91,07	0,20	0,10	33.83	0,00042
<b>Paciente 3 (M)</b>	Sano	8,18	23,63	0,29	12,30	0,20	0,07	23.43	0,00023
	Patológico	2,91	13,28	0,23	9,06	0,28	0,08	24.38	0,00030
<b>Paciente 4 (M)</b>	Sano	3,99	21,51	0,23	20,08	0,15	0,03	12.81	0,00011
	Patológico	3,09	31,21	0,24	10,84	0,10	0,03	15.37	0,00021
<b>Paciente 5 (F)</b>	Sano	4,93	36,67	0,21	118,64	0,27	0,034	35.20	0,00013
	Patológico	1,49	36,50	0,29	95,93	0,13	0,06	15.62	0,00023
<b>Vocal 'o'</b>		<b>Jitter</b>	<b>Shimmer</b>	<b>Cepstrum</b>	<b>Formante</b>	<b>Curtosis</b>	<b>Asimetría</b>	<b>Cruce ceros</b>	<b>HNR</b>

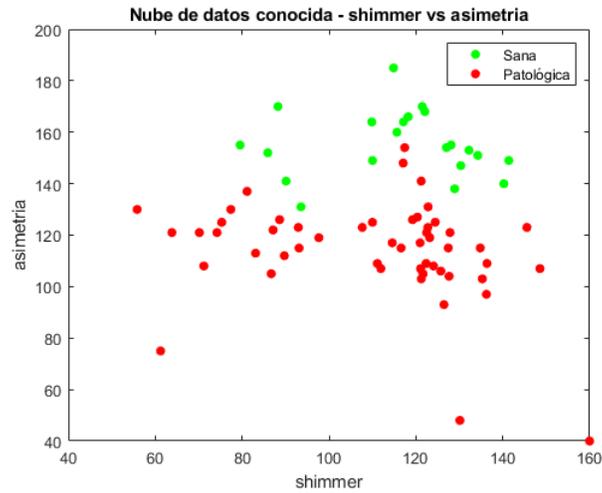
<b>Paciente 1 (M)</b>	Sano	1.81	16.46	0.28	5.42	0.14	0.056	16.47	0.00010
	Patológico	3.13	6.87	0.24	49.62	0.17	0.05	11.78	0.00190
<b>Paciente 2 (F)</b>	Sano	5.41	24.19	0.17	12.89	0.09	0.046	22.51	0.00029
	Patológico	5.95	31.93	0.23	21.99	0.14	0.10	23.93	0.00423
<b>Paciente 3 (M)</b>	Sano	6.48	9.66	0.24	23.61	0.34	0.19	15.56	0.00025
	Patológico	1.91	10.83	0.34	7.198	0.19	0.044	10.14	0.00016
<b>Paciente 4 (M)</b>	Sano	1.52	2.01	0.27	11.73	0.032	0.056	9.52	0.00020
	Patológico	1.04	5.39	0.22	4.62	0.05	0.03	6.76	0.00015
<b>Paciente 5 (F)</b>	Sano	6.03	49.69	0.16	66.84	0.14	0.075	24.64	0.00023
	Patológico	13.48	28.60	0.21	17.11	0.28	0.113	11.77	0.00025

Nota: Elaboración propia, 2021

### 5.3 Gráficas de dispersión

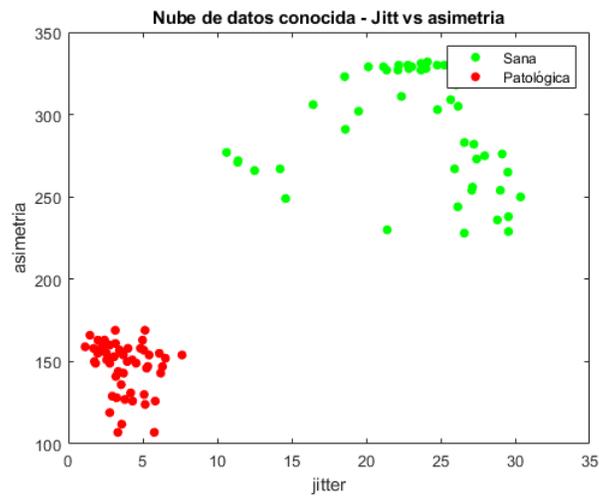
Se realizan diferentes combinaciones entre parámetros con el fin de graficar los valores correspondientes en un plano 2D a manera de gráficos de dispersión, en los cuales se puede observar la diferencia entre los valores de una voz sana y patológica del mismo paciente para determinada combinación. A continuación, se presentan algunos ejemplos de parámetros que permiten generar una separación entre las señal sana y patológica de un mismo paciente, para hombres y mujeres.

En la figura 26 se ilustra la separación de datos sano - patológico para un paciente hombre, vocal "a" con los parámetros asimetría Vs shimmer. A pesar, de que hay una separación para cada conjunto, los datos patológicos tienen una mayor dispersión que los sanos.

**Figura 5-26:***Hombre vocal 'a'*

Nota: Gráficos de dispersión a través del software Matlab, 2021

En la figura 27, se presenta el gráfico de dispersión sano – patológico de una mujer, vocal “a” con la asimetría Vs jitter; la agrupación de ambos conjuntos es clara y evidente.

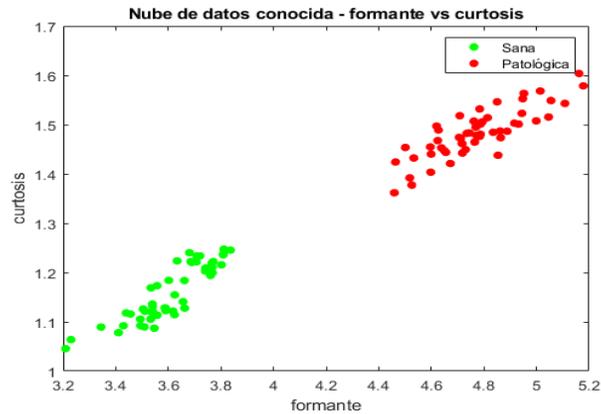
**Figura 5-27:***Mujer vocal 'a'*

Nota: Gráficos de dispersión a través del software Matlab, 2021

En la figura 28 se presenta el gráfico de dispersión de sano – patológico de un hombre, vocal “o” con la formante Vs curtosis. Donde la separación de los conjuntos de datos es clara y evidente.

**Figura 5-28:**

*Hombre vocal ‘o’*

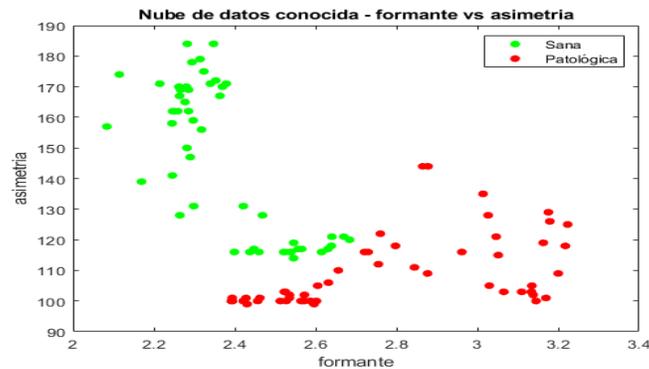


Nota: Gráficos de dispersión a través del software Matlab, 2021

En la figura 29, se observa el gráfico de dispersión de una mujer, vocal “o” con asimetría Vs formante. Donde hay una discriminación entre los conjuntos de datos. Sin embargo, algunos datos de ambos conjuntos tienden a dispersarse.

**Figura 5-29:**

*Mujer vocal ‘o’*



Nota: Gráficos de dispersión a través del software Matlab, 2021

A partir de estos gráficos de dispersión, se da el primer indicio para la elección de la vocal “a” en mujeres y la vocal “o” en hombres, debido a la clara separación de datos sano-patológico. Sin embargo, se necesitan más de 2 parámetros para obtener una tasa de acierto con mayor precisión debido a que la señal de voz varía constantemente en función del estado, edad, sexo, sonido vocálico, entre otros.

#### 5.4 Red neuronal

En la tabla 10, se presentan los resultados obtenidos a través de la red neuronal con solo con 2 entradas y en la tabla 11 se muestran los resultados con 3 o más entradas, en ambos casos, con parámetros obtenidos en un ventaneo de 150 ms de cada señal.

**Tabla 5-10:**

##### Resultados Red neuronal con 2 entradas

	Hombre 'a'	Mujer 'a'	Hombre 'o'	Mujer 'o'
<b>Asimetría-jitter</b>	82.25%	92.85%	53.22%	85.71%
<b>Asimetría-shimmer</b>	91.93%	90.48%	91.93%	85.71%
<b>Cruce cero-curtosis</b>	90.32%	81.34%	5.77%	71.15%
<b>P formante-Asimetría</b>	92.95%	83.11%	45.19%	21.74%
<b>P formante-curtosis</b>	90.38%	27.88%	27.88%	54.81%

Nota: Elaboración propia, 2021

**Tabla 5-11:**

##### Resultados Red neuronal con 3 o más entradas

Combinación	Parámetros	Hombre 'a'	Mujer 'a'	Hombre 'o'	Mujer 'o'
1	Jitter, shimmer, cepstrum, HNR, cruces por cero, primer formante, curtosis, asimetría	59.67%	88.09%	91.93%	71.42%
2	Asimetría, jitter, shimmer, curtosis, cruce por cero, primer formante	59.57%	97.53%	93.54%	69.05%
3	Jitter, shimmer, cepstrum, primera formante, curtosis, asimetría, cruce por cero	46.77%	90.47%	75.80%	78.57%
4	Jitter, shimmer, primera formante, curtosis, asimetría	66.13%	97.62 %	96.77%	76.19%

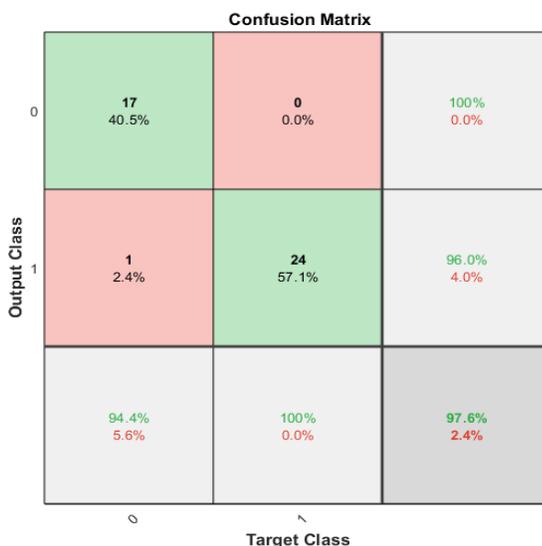
Nota: Elaboración propia, 2021

Después de realizar un análisis de las diferentes combinaciones, se determinó que para el caso de las mujeres resulta una mejor tasa de acierto la combinación 4 conformada por jitter, shimmer, primera formante, curtosis y asimetría con la vocal ‘a’, alcanzando una máxima tasa de acierto de 97.62%. En el caso de los hombres se elige la misma combinación, pero con la diferencia que para el sexo masculino se hace uso de la vocal ‘o’; se logra una máxima tasa de acierto de 96.77%

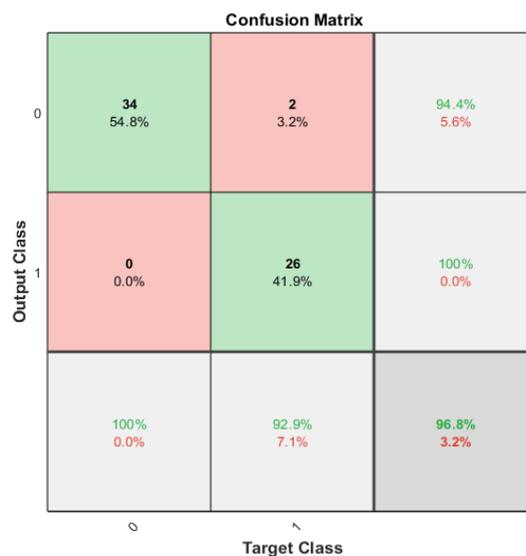
Para una visualización más efectiva de los resultados obtenidos por la red neuronal, se realizaron las respectivas matrices de confusión indicadas en la figura 30 y 31, tanto para hombres como para mujeres, con la combinación de parámetros antes descrita.

**Figura 5-30:**

*Matriz de confusión mujeres vocal a*



Nota: matriz de confusión obtenida a través de Matlab, 2021

**Figura 5-31:***Matriz de confusión hombres vocal o*

Nota: matriz de confusión obtenida a través de Matlab, 2021

A partir de las matrices de confusión como se muestra en la tabla 12, se realizó el cálculo de los siguientes parámetros: la sensibilidad, que indica la fracción de verdaderos positivos y la especificidad la fracción de verdaderos negativos, la exactitud es el cociente entre la suma de verdaderos positivos, más verdaderos negativos y el total de los datos evaluados. Y la precisión es la división entre los verdaderos positivos y la suma de verdaderos positivos más los falsos negativos.

**Tabla 5-12:****Matrices de confusión**

	Precisión	Sensibilidad	Especificidad	Exactitud
<b>Mujeres</b>	1	0.94	1	0.98
<b>Hombres</b>	0.94	1	0.93	0.97

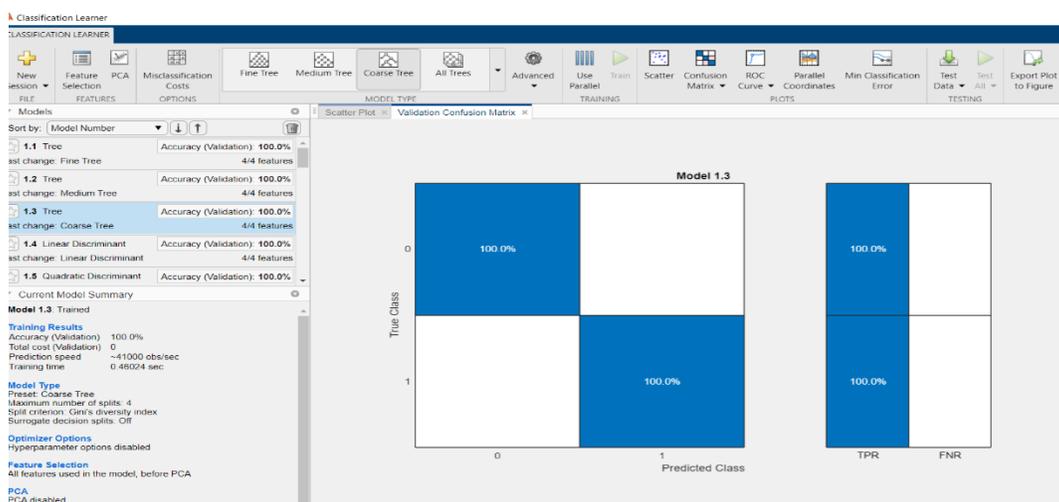
Nota: Elaboración propia, 2021

## 5.5 Sistema web

A continuación, en la figura 32 se presenta el resultado de la discriminación de sexo a través del algoritmo *classification Learner* de Matlab con el árbol de decisión; se alcanzaron resultados con tasa de acierto hasta del 100%, haciendo uso de los parámetros *f0*, *jitter*, *shimmer* y *HNR*.

**Figura 5-32:**

*Discriminación de sexo*



Nota: Elaboración propia a través de Matlab, 2021

Por otra parte, en la figura 33, se muestra la página web implementada a través de lenguaje HTML, se observa la interfaz de inicio, la cual interactúa directamente con el usuario; el sistema en general recibe el nombre “**HumanVoice**” y ofrece opciones como información acerca del equipo de trabajo, información sobre el sistema respiratorio - la voz, análisis de la voz y contacto. Adicionalmente, en esta interfaz de inicio, se presentan tres imágenes principales que cambian cada 3 segundos, con el fin de hacer alusión al objetivo; Analizar la voz, contribuir a los protocolos de bioseguridad (figura 34) y donar la voz (figura 35).

**Figura 5-33:***Interfaz de inicio – Analizar la voz*

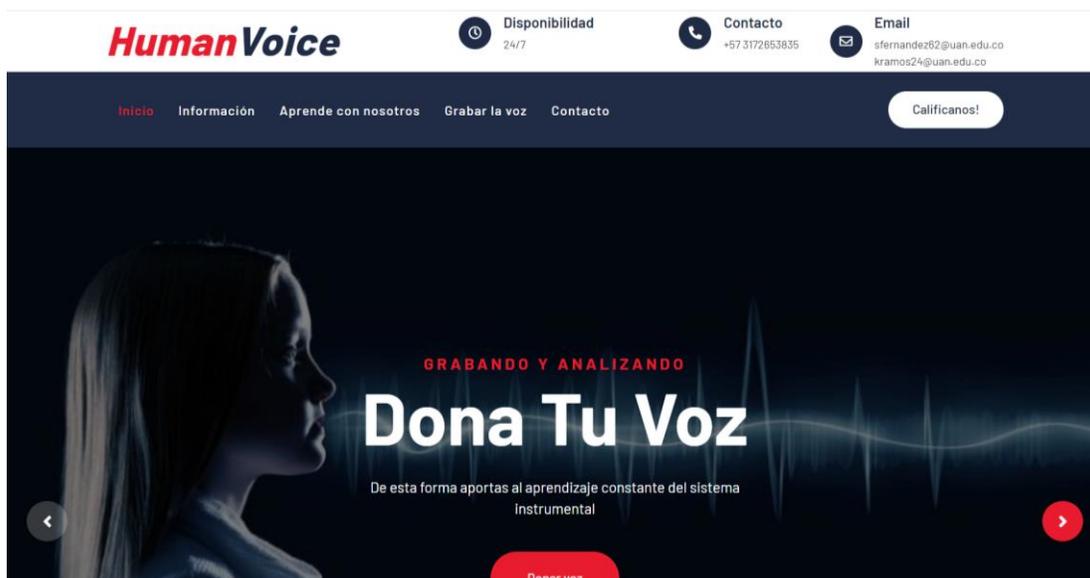
Nota: Elaboración propia a través de HTML, 2021

**Figura 5-34:***Interfaz de inicio – Contribuir a los protocolos de bioseguridad*

Nota: Elaboración propia a través de HTML, 2021

**Figura 5-35:**

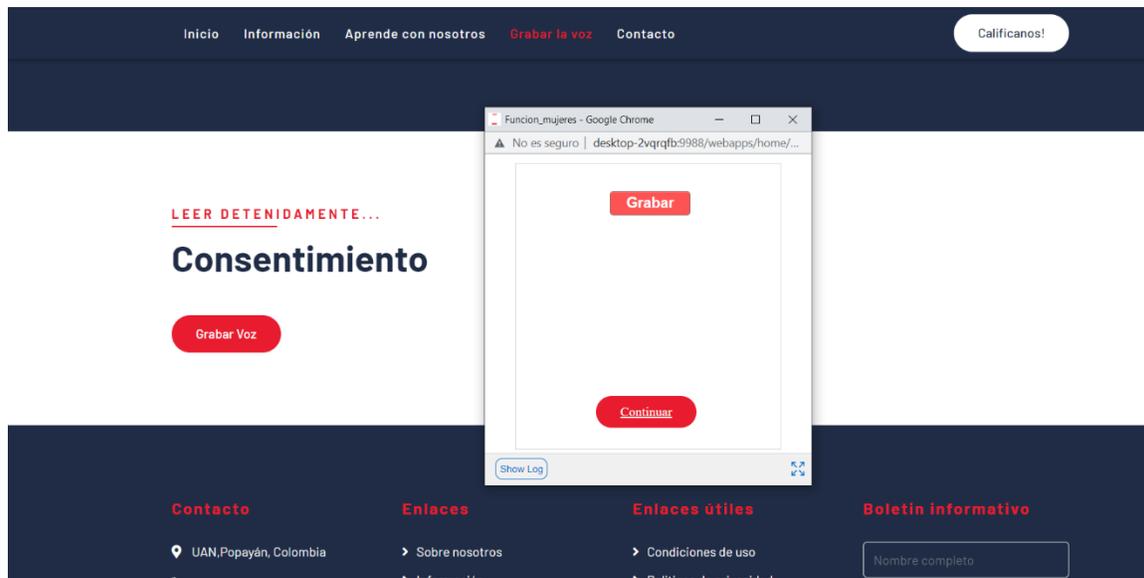
*Interfaz de inicio – Donar la voz*



Nota: Elaboración propia a través de HTML, 2021

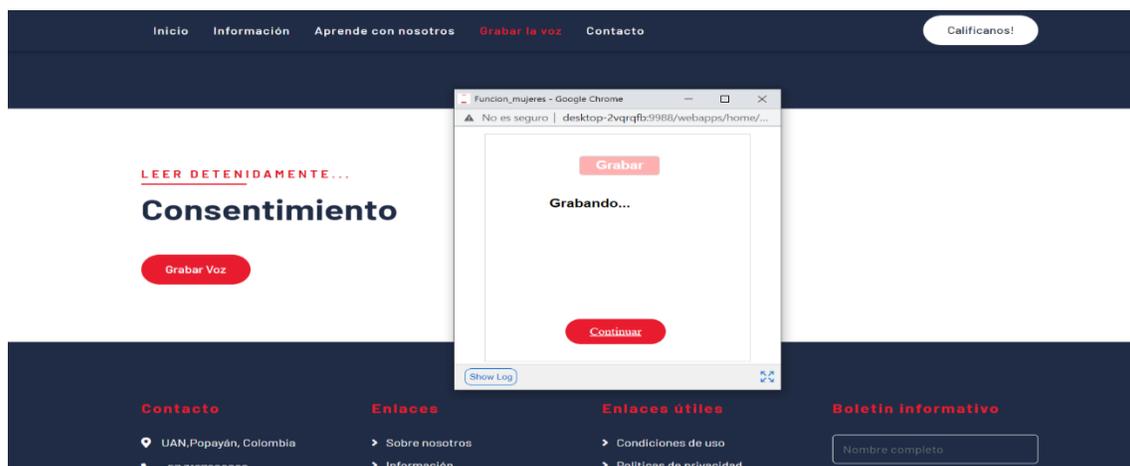
Posteriormente, una vez el usuario haya adquirido la información necesaria sobre el objetivo del sistema, accede a “Grabar la voz”, donde se presenta un consentimiento informado para dar a conocer los términos del sistema; Analizar la voz de forma libre, consiente y voluntaria, teniendo en cuenta que es un diagnóstico preliminar y también, que al realizar este análisis, automáticamente su voz será almacenada en el sistema para el aprendizaje constante del mismo.

Al aceptar las condiciones, aparece una ventana emergente como se muestra en la figura 36, donde el usuario podrá realizar la grabación al presionar el botón “Grabar”.

**Figura 5-36:***Análisis de la voz - Grabar la voz*

Nota: Nota: Elaboración propia a través de HTML, 2021

Una vez se presione el botón “Grabar”, el usuario realiza la grabación del sonido vocálico durante 6 segundos; para confirmar que está realizando el procedimiento de forma correcta, en la ventana emergente debe aparecer “Grabando...” como se muestra en la figura 37.

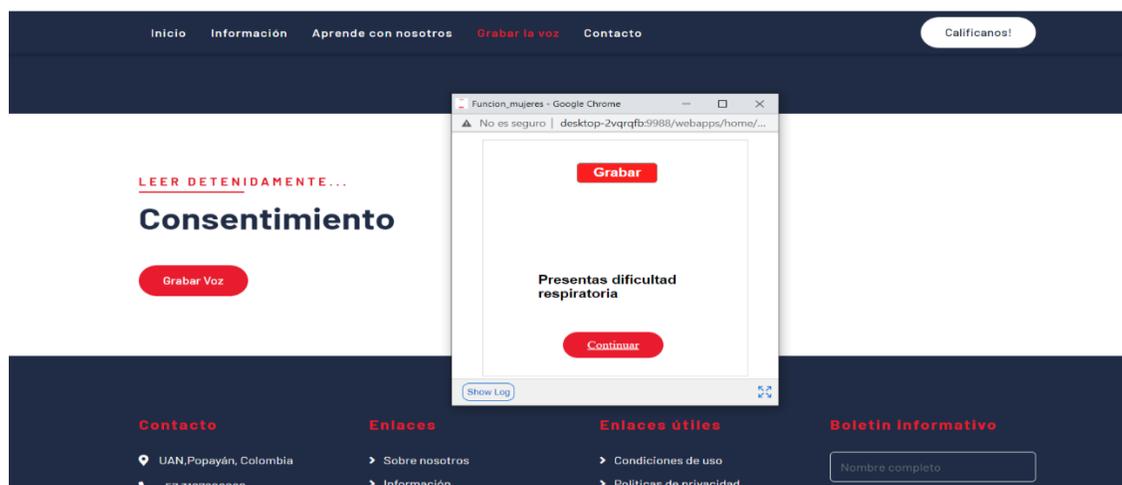
**Figura 5-37:***Análisis de la voz - Grabando*

Nota: Elaboración propia a través de HTML, 2021

Por último, al pasar los 6 segundos de grabación del sonido vocálico, la señal se envía internamente al servidor para el respectivo análisis y clasificación, y posteriormente, se presenta el resultado del análisis de la voz como: Presencia de normalidad respiratoria o dificultad respiratoria como se presenta en la figura 38.

### Figura 5-38:

#### *Análisis de la voz - Resultado*



Nota: Elaboración propia a través de HTML, 2021

## Conclusiones

Este sistema describe la implementación de un método que interrelaciona componentes software y hardware, incorpora inteligencia computacional, para detectar, de forma automática y no invasiva, la presencia de dificultad respiratoria, a través de métodos avanzados de procesamiento digital de señales, que implementados en una arquitectura cliente servidor, permiten el análisis de señales de voz de personas hispano hablantes, obtenidas como resultado de la grabación de un sonido vocálico sostenido a través del micrófono de un dispositivo electrónico.

Se logró definir los descriptores fonorespiratorios para diferenciar entre una voz sana y patológica a través del procesamiento digital de señales de voz. También, se logró implementar una red neuronal de gradiente conjugado escalado que incorpora como entradas los descriptores fonorespiratorios; alcanzando una tasa de acierto hasta del 100%. Destacando que la vocal “a” permite una mayor tasa de acierto en mujeres y la vocal “o” en hombres. Adicionalmente, el sistema se evaluó a partir de la matriz de confusión y el cálculo de la precisión, sensibilidad, especificidad y exactitud.

De acuerdo a los resultados obtenidos, se establece que la interrelación de parámetros procesados en el sistema de inteligencia computacional, contribuyen a la evaluación de la alteración o normalidad respiratoria, la cual es muy importante en el contexto de las crisis sanitarias que pueden ocasionar las enfermedades respiratorias de fácil transmisión, y que reflejan la pertinencia del sistema para suministrar soporte a los protocolos de bioseguridad.

Finalmente, se realizó la solicitud de patente a la superintendencia de Industria y Comercio de Colombia (SIC), donde se obtuvo un radicado como patente de invención nacional debido al alto componente innovador y a un bajo grado de similitud con tecnologías ya existentes.

## Anexos

### Anexo 1. Radicado de Patente ID256 “MÉTODO PARA DETECTAR DIFICULTAD RESPIRATORIA, A PARTIR DEL PROCESAMIENTO DIGITAL DE SEÑALES DE VOZ, CON INTELIGENCIA COMPUTACIONAL.”

Superintendencia de Industria y Comercio de Colombia (SIC) - Reporte detallado de solicitudes

Fecha y hora: 22 oct. 2021 05:12:08 p.m.

Número de Patente (11): [NC2021/0014164](#)

Fecha de radicación: 22 oct. 2021

Referencia de solicitante: ID256

Estado de la solicitud: Bajo Verificación de Requisitos Mínimos

Tipo de solicitud: NC Solicitud de Patente/Modelo/Trazado

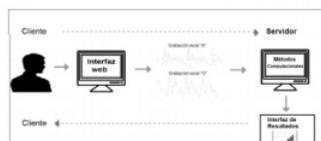
Tiempo para la publicación: 12

Área Tecnológica: Electricidad - Electrónica

Título: MÉTODO PARA DETECTAR DIFICULTAD RESPIRATORIA, A PARTIR DEL PROCESAMIENTO DIGITAL DE SEÑALES DE VOZ, CON INTELIGENCIA COMPUTACIONAL.

Tipo: Patente de Invención Nacional

#### Resúmen:



## Referencias Bibliográficas

- Abeyratne, U. y Swarnkar, V. (2018). A method for analysis of cough sounds using disease signatures to diagnose respiratory diseases. <https://patents.google.com/patent/US20210076977A1/en?q=US2021076977A1>
- Abramov, G.V, Korobova, L.A., Ivashin, A.L. & Matytsina, I.A. (2018). Information system for diagnosis of respiratory system diseases. <https://iopscience.iop.org/article/10.1088/1742-6596/1015/4/042036/meta>
- Agencia Iberoamericana para difusión de la ciencia y tecnología. (2021). Desarrollan un sistema de detección de la insuficiencia respiratoria a través de la voz. Es un trabajo cuyo objetivo consiste en ayudar en el triaje de pacientes con sospecha de COVID-19. <https://www.dicyt.com/noticias/desarrollan-un-sistema-de-deteccion-de-la-insuficiencia-respiratoria-a-traves-de-la-voz>
- Alvarado Valencia, J. A. y Obagi Araújo, J.J. (2008). Fundamentos de inferencia estadística. Edit. Pontificia Universidad Javeriana. [https://books.google.com.co/books?hl=es&lr=&id=3uhUqvF0\\_84C&oi=fnd&pg=PR7&dq=asimetria+en+estadistica&ots=DF5FibybnN&sig=zMmJ3nZfdLCTs6kqCOI-g7vMtw&redir\\_esc=y#v=onepage&q=asimetria%20en%20estadistica&f=false](https://books.google.com.co/books?hl=es&lr=&id=3uhUqvF0_84C&oi=fnd&pg=PR7&dq=asimetria+en+estadistica&ots=DF5FibybnN&sig=zMmJ3nZfdLCTs6kqCOI-g7vMtw&redir_esc=y#v=onepage&q=asimetria%20en%20estadistica&f=false)
- Anushiravani, R., Krishna Nemala, S., Kiran Yalamanchili, R. & Swetha Davuluri, N. (2020). Intelligent health monitoring. <https://patents.google.com/patent/US20200388287A1/en?q=US2020388287A1>
- Aronovich, D y Hassan, S.H. (2020). Cesionario actual Salud Vocalis. <https://patents.google.com/patent/US20200323484A1/en?q=US2020323484A1>
- Asenjoa, C.A. y Pintob, R.A., (2017). Características anatómo-funcional del aparato respiratorio durante la infancia. Revista Médica Clínica Las Condes Vol 28 N° 1. <https://www.elsevier.es/es-revista-revista-medica-clinica-las-condes-202-articulo-caracteristicas-anatomo-funcional-del-aparato-respiratorio-S0716864017300020>
- Asociación Estadounidense del Habla, el Lenguaje y la Audición (ASHA). (2021). Nódulos y Pólipos de las Cuerdas Vocales. <https://www.asha.org/public/speech/spanish/nodulos-y-polipos-de-las-cuerdas-vocales>

- Azuaje, Y. (2019). Alteración De La Relación Ventilación/perfusión. [file:///C:/Users/MIPC/Downloads/alteracion-de-la-relacion-ventilacion-perfucion\\_compress.pdf](file:///C:/Users/MIPC/Downloads/alteracion-de-la-relacion-ventilacion-perfucion_compress.pdf)
- Banuelos, D. (2021). Ilustraciones del aparato respiratorio. En: Imagui. <https://www.imagui.com/a/ilustraciones-del-aparato-respiratorio-iG6rB6pnL>
- Brown, CH., Chauha, J., Grammeno, A., Han, J., Hasthanasomb, A., Spathi, D., Xia, T., Cicuta, P. & Mascolo, C. (2020, junio 10). Exploring Automatic Diagnosis of COVID-19 from Crowdsourced Respiratory Sound Data. New York, NY, USA. <https://arxiv.org/abs/2006.05919>
- Byron, J. A. (s.f). Formantes en la fonética. acústica [https://www.academia.edu/4928142/Formantes\\_en\\_la\\_fon%C3%A9tica\\_ac%C3%BAstica#:~:text=El%20formante%20es%20una%20C2%ABcualidad%20ac%C3%BAstica%20propia%20de,numeran%20de%20abajo%20arriba%20como%20F1%2C%20F2%2C%20F3](https://www.academia.edu/4928142/Formantes_en_la_fon%C3%A9tica_ac%C3%BAstica#:~:text=El%20formante%20es%20una%20C2%ABcualidad%20ac%C3%BAstica%20propia%20de,numeran%20de%20abajo%20arriba%20como%20F1%2C%20F2%2C%20F3)
- Casas Quiroga, I. C., Contreras Zúñiga, E., Zuluaga Martínez, S.X. y J. Mejía Mantilla, J. (2008). Diagnóstico y manejo de la insuficiencia respiratoria aguda, Medigraphic, vol. 67, n° 1, pp. 24-33. <https://1library.co/document/ky663lgy-diagnostico-y-manejo-de-la-insuficiencia-respiratoria-aguda.html>
- Catalunya Vanguardista. (2020, julio 17). Detectar insuficiencia respiratoria a través de la voz. <https://www.catalunyavanguardista.com/detectar-insuficiencia-respiratoria-a-traves-de-la-voz/>
- Castellanos, G. y Sepúlveda, A. (2004). Estimación de la frecuencia fundamental de señales de voz usando transformada wavelet. Scientia et Technica, Vol. 1, N°. 24. <https://dialnet.unirioja.es/servlet/articulo?codigo=4844818>
- Coloma O. R. y Álvarez, J. P. A. (2011). Manejo avanzado de la vía aérea. Revista Médica Clínica Las Condes Vol. 22 N° 3. <https://www.elsevier.es/es-revista-revista-medica-clinica-las-condes-202-articulo-manejo-avanzado-via-aerea-S0716864011704266>
- Delgado, J., León, N., Jiménez, A., e Izquierdo, L. (2017). Análisis acústico de la voz: medidas temporales, espectrales y cepstrales en la voz normal con el Praat en una muestra de hablantes de español. Revista de Investigación en Logopedia Vol 7 N°2.108-109. <https://revistas.ucm.es/index.php/RLOG/article/view/58191>

- Delgado Hernández, J., León Gómez, N.M., Izquierdo Arteaga, L.M. y Yanira Llanos Fumero, Y. (2018). Análisis cepstral de la voz normal y patológica en adultos españoles. Medida de la prominencia del pico cepstral suavizado en vocales sostenidas versus habla conectada. Vol 69. <https://www.sciencedirect.com/science/article/abs/pii/S0001651917301577>.
- Del Río Navarro, B.E., Hidalgo Castro, E.M. y Sienna Monge, J.J. (2008). Asma. Bol. Med. Hosp. Infant. Mex. vol.66 no.1. [http://www.scielo.org.mx/scielo.php?script=sci\\_arttext&pid=S1665-11462009000100002](http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1665-11462009000100002)
- Edinburgh & Dunfermline, J.W. (2014). Systems and methods for identifying patient talking during measurement of a physiological parameter. the lens. <https://patentimages.storage.googleapis.com/f7/3f/3b/011e7be0313cda/US20140276165A1.pdf>
- Elisei, N.G. (2012). Análisis acústico de la voz normal y patológica utilizando dos sistemas diferentes: ANAGRAF y PRAAT, SCIELO, vol. 29, nº 2, pp. 271-28.
- Estrada Corona, A. (2004, septiembre 10). Protocolos TCP/IP de internet. Revista Digital Universitaria Vol.5 N°8. <https://www.ru.tic.unam.mx/bitstream/handle/123456789/791/220.pdf?sequence=1&isAllowe>
- Galera R., Gómez Carrera, L. y Ortega, B. (2010). Enfermedades del aparato respiratorio, Servicio de Neumología. Hospital Universitario La Paz., Madrid. España. Medicine Vol. 10. Núm. 63. <https://www.medicineonline.es/es-enfermedades-del-aparato-respiratorio-articulo-S0304541210702419>
- Garzón, M y Muñoz, J. (2018). Evidencias de la Técnica Vocal Flamenca: Aerodinámica y Rango de Frecuencias. Conference: V Congreso Interdisciplinar Investigación y FlamencoAt: Sevilla. [https://www.researchgate.net/profile/Marina-Garzon-4/publication/341701929\\_Evidencias\\_de\\_la\\_Tecnica\\_Vocal\\_Flamenca\\_Aerodinamica\\_y\\_Rango\\_de\\_Frecuencias/links/5ecf78ba299bf1c67d23de5a/Evidencias-de-la-Tecnica-Vocal-Flamenca-Aerodinamica-y-Rango-de-Frecuencias.pdf](https://www.researchgate.net/profile/Marina-Garzon-4/publication/341701929_Evidencias_de_la_Tecnica_Vocal_Flamenca_Aerodinamica_y_Rango_de_Frecuencias/links/5ecf78ba299bf1c67d23de5a/Evidencias-de-la-Tecnica-Vocal-Flamenca-Aerodinamica-y-Rango-de-Frecuencias.pdf)
- German. (2015). Introducción a las Redes Neuronales Artificiales. (Blog). Inteligencia artificial. <https://iagermanicovp.blogspot.com/2015/06/introduccion-las-redes-neuronales.html>
- Goldman, A., Cândido Júnior, A., Fernández Svartma, F.R., Gomes de Queiro, M. y María Aluísio, S. (2020, junio 13). Proyecto de Detección Temprana de Fallo Respiratorio a través de

Análisis de Audio (SPIRA), Investigadores brasileños crean programas informáticos para ayudar a los pacientes de Covid.

Gómez García, J.A., Moro Velázquez, L., Godino Llorente, J.I y Castellanos Domínguez, C.G. (2016). Una mirada a la categorización automática de hablantes según el sexo y su aplicación a la detección de patologías de la voz: un estudio comparativo. *Revista Facultad De Ingeniería Universidad De Antioquia*, (79), 50–62. <https://doi.org/10.17533/udea.redin.n79a06>

Gómez Vilda, P., Rodellar Biarge, V., Nieto LLuis, V., Álvarez Marquina, A., y Martínez Olalla, R. (2011). Método y sistema para la estimación de parámetros fisiológicos de la fonación.

Hirsch, L. (2019). Los pulmones y el sistema respiratorio. <https://kidshealth.org/es/teens/lungs.html>

Hong Yuan Hong., Fan Tao Wu Xuesong y Mao Wei. (2019). Device and method for diagnosing respiratory disease

Imran, A., Posokhova, I., Qureshi H. N., Masood, U., Sajid Riaz, M., Kamran Ali, K., Charles N. J., Hussain, I., & Nabeel, M. (2020). AI4COVID-19: AI enabled preliminary diagnosis for COVID-19 from cough samples via an app. *Informatics in Medicine Unlocked* Volume 20. <https://www.sciencedirect.com/science/article/pii/S2352914820303026?via%3Dihub>

Jiménez Cubides, V.P. y Espitia Camargo, N. Y. (2016). Abordaje fonoaudiológico de un desorden vocal en una estudiante de comunicación social y su impacto en la práctica académica/laboral. *Asociación colombiana de fonoaudiología. V congreso internacional*. [https://asofono.co/wp-content/uploads/2016/09/ICOV\\_p43-47\\_Jimenez\\_probvozestudcomsocial.pdf](https://asofono.co/wp-content/uploads/2016/09/ICOV_p43-47_Jimenez_probvozestudcomsocial.pdf).

Kike. G. (2018). Hacer ejemplos detector cruce por cero. En: *Forum*. <https://forum.arduino.cc/t/hacer-ejemplos-detector-cruce-por-cero/526106>

Kita, K. (1995). Dispositivo de inspección de voz. *JPH09173320A*. <https://patents.google.com/patent/JPH09173320A/en?q=JPH09173320A>

Leiva Suero, L. E., Chicaiza Tayupanta, J.O., Proaño Alulema, R.X., Hernández Navarro, E.V y Recalde Navarrete, R.J. (2020). «Tecnologías emergentes con aplicación en Salud: desarrollo de una plataforma integrada para la evaluación de factores de riesgo,

diagnóstico, tratamiento, pronóstico y geolocalización en Pandemia Co-vid-19 por SARS-Cov-2»,» Mediciencias UTA.

Marco Sanjuán, F. J. (2017). Curtosis. Economipedia.com

Marlina, M., Saputra, W., Mulyadi, B., Hayati, B., & Jaroji, J. (2017). Aplikasi sistem pakar diagnosis penyakit ispa berbasis speech recognition menggunakan metode naive bayes classifier. *Digital Zone: Jurnal Teknologi Informasi Dan Komunikasi*, 8(1), 58-70. <https://doi.org/10.31849/digitalzone.v8i1.629>.

Martínez Celdrán, E. (2015). Naturaleza fonética de la consonante ‘ye’ en español. *Revista de estudios lingüísticos hispánicos*, N°5. <https://dialnet.unirioja.es/servlet/articulo?codigo=5191567>

Mayo clinic. (2021). Laringitis. Libro de Salud Familiar de Mayo Clinic) 5.<sup>a</sup> edición. <https://www.mayoclinic.org/es-es/diseases-conditions/laryngitis/symptoms-causes/syc-20374262>

Mayo clinic. (2021). Parálisis de las cuerdas vocales. <https://www.mayoclinic.org/es-es/diseases-conditions/vocal-cord-paralysis/symptoms-causes/syc-20378873>

Mayo clinic. (2021). Trastornos de la voz. <https://www.mayoclinic.org/es-es/diseases-conditions/voice-disorders/symptoms-causes/syc-20353022>

Mayorga Ortiz, P., C. Druzgalski, O. H. González Arriaga y M. R. Ludwing. (2010). Valoración de Enfermedades Respiratorias Basada en Modelación Acústica. Congreso Internacional de AcademiaJournals.com, vol. III, pp. 112-118.

MedlinePlus enciclopedia médica. (2021). Gasometría arterial. <https://medlineplus.gov/spanish/ency/article/003855.htm>

Ministerio de Salud. (2020). Enfermedad pulmonar obstructiva crónica (EPOC) <https://www.minsalud.gov.co/salud/publica/PENT/Paginas/Enfermedad-pulmonar-obstructiva-cronica.aspx>

MSN Noticias. (2021, febrero 24). Mapa de casos y muertes por coronavirus por departamentos en Colombia. <https://www.msn.com/es-co/noticias/otras/mapa-de-casos-y-muertes-por-coronavirus-por-departamentos-en-colombia-hoy-24-de-febrero/ar-BB1dYgLO>.

Muñoz Retana, C. (2018, agosto 14). Infección respiratoria aguda (IRA). [https://www.geosalud.com/enfermedades\\_infecciosas/ira.htm](https://www.geosalud.com/enfermedades_infecciosas/ira.htm)

- Nasreddine Belkacem, A., Ouhbi, S., Laka, A., Elhadj Benkhalif, E. & Chao Chen, Ch. (2021). End-to-End AI-Based Point-of-Care Diagnosis System for Classifying Respiratory Illnesses and Early Detection of COVID-19: A Theoretical Framework. <https://www.frontiersin.org/articles/10.3389/fmed.2021.585578/full>
- National Heart, Lung and Blood Institute (NIH). (2021). Pruebas de la función pulmonar | <https://www.nhlbi.nih.gov/health-topics/espanol/pruebas-de-la-funcion-pulmonar>.
- National Heart, Lung and Blood Institute (NIH). (2019). Insuficiencia respiratoria. <https://www.nhlbi.nih.gov/health-topics/espanol/insuficiencia-respiratoria>.
- Navarro Mesa, J. L. (s.f). Procesador Acústico: El Bloque de Extracción de Características. Departamento de Señales y Comunicaciones. Tratamiento de la Señal de Audio. Universidad de Las Palmas de Gran Canaria. <http://www2.ulpgc.es/hege/almacen/download/25/25296/apuntesextraccioncaracteristicas.pdf>
- Nieto, A. (2001). Disfonía musculotensional Muscle tension dysphonia En: Revista Biomédica Revisada Por Pares. <https://www.medwave.cl/link.cgi/Medwave/PuestaDia/Cursos/3084>
- Niño, E. (s.f). Fundamentos para el desarrollo de aplicaciones en la red. Modelo cliente servidor. Ingeniería de software. <http://manglar.uninorte.edu.co/bitstream/handle/10584/2205/Modelo%20cliente%20servidor.pdf?sequence=>
- Pino Diez, R, Gómez Gómez, A. y De Abajo Martínez, N. I. (2001). Introducción a la inteligencia artificial. Sistema de expertos, Redes Neuronales artificiales y computación evolutiva. [https://books.google.com.co/books?hl=es&lr=&id=RKqLMCw3IUkC&oi=fnd&pg=PA10&dq=redes+neuronales+artificiales&ots=iHGCl3C24O&sig=rak4DGHrEPfG-ueUpQ4ExW\\_tZHM&redir\\_esc=y#v=onepage&q&f=false](https://books.google.com.co/books?hl=es&lr=&id=RKqLMCw3IUkC&oi=fnd&pg=PA10&dq=redes+neuronales+artificiales&ots=iHGCl3C24O&sig=rak4DGHrEPfG-ueUpQ4ExW_tZHM&redir_esc=y#v=onepage&q&f=false)
- Preparadores de Oposiciones para la Enseñanza. (2005). La voz humana y su fisiología. Clasificación de las voces. La voz en la adolescencia: Características y problemática. <https://www.preparadores.eu/secundaria/Musica/Musica-Tema.pdf>
- Organización Mundial de la Salud (OMS). (2017). Cada año, hasta 650.000 personas mueren por enfermedades respiratorias relacionadas con la gripe estacional, Ginebra. <https://apps.who.int/mediacentre/news/releases/2017/seasonal-flu/es/index.html>

- Organización Mundial de la Salud (OMS). (2020). Neumonía de causa desconocida – China. [https://www.who.int/es/emergencias/disease-outbreak-news/item/2020-DON229#:~:text=El%2031%20de%20diciembre%20de,provincia%20de%20Hubei%20\(China\).](https://www.who.int/es/emergencias/disease-outbreak-news/item/2020-DON229#:~:text=El%2031%20de%20diciembre%20de,provincia%20de%20Hubei%20(China).)
- Organización Mundial de la Salud y Organización de las Naciones Unidas para la Agricultura y la Alimentación. (2010). Bioseguridad: Enfoque integrado de la gestión del riesgo para la vida y la salud de las personas, los animales y las plantas, INFOSAN.
- Organización Mundial de la Salud. (2019, agosto 02). Neumonía. <https://www.who.int/es/news-room/fact-sheets/detail/pneumonia>
- Organización Mundial de la Salud (OMS). (2020, octubre 12). Información básica sobre la COVID-19. <https://www.who.int/es/news-room/q-a-detail/coronavirus-disease-covid-19>
- Rao, A., Huynh, E., Roysto, T.J., A. Kornblith, A.& Roy, S. (2019). Acoustic Methods for Pulmonary Diagnosis. in IEEE Reviews in Biomedical Engineering, vol. 12, pp. 221-239. doi: 10.1109/RBME.2018.2874353. <https://ieeexplore.ieee.org/abstract/document/8514011>
- Resolución 2654. (2019). Por la cual Establecen disposiciones para la telesalud y parámetros para la práctica de la telemedicina en el país. (Colombia)
- Rey Lancheros, D.E., Gavilán Acosta, H.J. y Espitia Cuchango, H.E. (2017). Implementación de un algoritmo para la identificación de usuarios considerando problemas fisiológicos que afectan el habla. Iteckne, vol. 14, n° 2, pp. 131-139. <http://www.scielo.org.co/pdf/itec/v14n2/1692-1798-itec-14-02-00131.pdf>
- Saleheen, N., Ahmed, T., Mahbubur Rahma, M, Nemat, E., Nathan, V., Vatanparvar, K., Blackstock, E. & Kuang, J. (2020, October 05). Lung Function Estimation from a Monosyllabic Voice Segment Captured Using Smartphones. N°10. <https://dl.acm.org/doi/10.1145/3379503.3403543>
- Sánchez Monge, M. (2021). Gripe. <https://cuidateplus.marca.com/enfermedades/infecciosas/gripe.html>
- Silva Zambran, M.M., Romo Romero, H.A., Ramírez Viáfara, J.M. y Galvis Zambrano, D.M. (2017). Técnicas de detección de la frecuencia fundamental de la voz en entornos reales. Ingeniería Solidaria, vol. 13, N.o 23, pp. 122-137.

[https://revistas.ucc.edu.co/html\\_revistas/IngSol/13%2823%29/13%2823%299/13%2823%299.html](https://revistas.ucc.edu.co/html_revistas/IngSol/13%2823%29/13%2823%299/13%2823%299.html)

Schiaffarino, A. (2019). Modelo cliente servidor. (Blog).  
<https://blog.infranetworking.com/modelo-cliente-servidor/>

Teixeira, J. P. y Carla Lopes, C.O. (2013). Vocal Acoustic Analysis – Jitter, Shimmer and HNR Parameters. Vol. 9.  
<https://www.sciencedirect.com/science/article/pii/S2212017313002788>

Torres Gallardo, B. (2013). La voz y nuestro cuerpo. Un análisis funcional. Revista de Investigaciones en Técnica Vocal, año 1, n°1. La Plata: Facultad de Bellas Artes UNLP.  
[http://sedici.unlp.edu.ar/bitstream/handle/10915/44777/Documento\\_completo.pdf?sequence=1](http://sedici.unlp.edu.ar/bitstream/handle/10915/44777/Documento_completo.pdf?sequence=1)

Travieso González, C.M., Alonso Hernández, J.B., Ferrer Ballester, M.A., y De león, J.J. (2013, diciembre 03). Método para la evaluación clínica del sistema fonador de pacientes con patologías laríngeas a través de una evaluación acústica de la calidad de la voz. En Patentados. Com. <https://patentados.com/2013/metodo-para-la-evaluacion-clinica>

Vila Rovira, J. M., Valero García, J. y González Sanvisens, L. (2011). Indicadores fonorrespiratorios de normalidad y patología en la clínica vocal. Revista de Investigación en Logopedia Vol. 1 N°1. <https://revistas.ucm.es/index.php/RLOG/article/view/58706>

Vocalis health. (2021). ¿Cómo podemos volver a la "nueva normalidad" después de la pandemia? <https://vocalishealth.com/vocalis-health-products/vocalischeck>

Vocalis health. (2021). ¿Cómo detectan los equipos de atención clínica de forma remota el deterioro en los pacientes con EPOC? <https://vocalishealth.com/vocalis-health-products/vocalistrack>

Wikipedia Academia. (2010). Red neuronal artificial. <https://esacademic.com/dic.nsf/eswiki/994505>

World Bank. (2020). Global Economic Prospects, Washington, DC: World Bank. DOI: 10.1596/978-1-4648-1553-9.