



**ESPE**  
**UNIVERSIDAD DE LAS FUERZAS ARMADAS**  
**INNOVACIÓN PARA LA EXCELENCIA**

**Desarrollo de un Sistema Inteligente de Sensado como apoyo para la  
Discriminación de Cacao Sano y Contaminado en sus Procesos de Control de  
Calidad**

Catucumbamba Taco, Carlos Andrés y Molina Peláez, Jorge David

Departamento de Eléctrica, Electrónica y Telecomunicaciones

Carrera de Ingeniería en Electrónica, Automatización y Control

Trabajo de Titulación, previo a la obtención del título de Ingeniero en Electrónica,  
Automatización y Control

Ing. Guamán Novillo, Ana Verónica, Ph. D.

23 de Julio del 2021



### Document Information

Analyzed document	Capitulos del Trabajo de Titulación (1).pdf (D110696292)
Submitted	7/23/2021 3:49:00 PM
Submitted by	Guaman Novillo Ana Veronica
Submitter email	avero.guaman@gmail.com
Similarity	1%
Analysis address	avguaman.espe@analysis.orkund.com

### Sources included in the report

<b>W</b>	URL: <a href="http://posgrado.lapaz.tecnm.mx/uploads/archivos/AguirreEstrada.pdf">http://posgrado.lapaz.tecnm.mx/uploads/archivos/AguirreEstrada.pdf</a> Fetched: 5/10/2021 6:01:24 AM	 1
<b>SA</b>	<b>Tesis Karolay De La Cruz_orkund.pdf</b> Document Tesis Karolay De La Cruz_orkund.pdf (D61973260)	 2
<b>W</b>	URL: <a href="https://repositorio.espe.edu.ec/bitstream/21000/21638/1/T-ESPE-040950.pdf">https://repositorio.espe.edu.ec/bitstream/21000/21638/1/T-ESPE-040950.pdf</a> Fetched: 9/13/2020 12:57:00 AM	 1
<b>SA</b>	<b>Tesis_SEBASTIANJACOME_analisis (1).pdf</b> Document Tesis_SEBASTIANJACOME_analisis (1).pdf (D62031427)	 1

**Ing. Guamán Novillo, Ana Verónica Ph. D.**

**DIRECTORA**



**DEPARTAMENTO DE ELÉCTRICA, ELECTRÓNICA Y  
TELECOMUNICACIONES**

**CARRERA DE INGENIERÍA EN ELECTRÓNICA, AUTOMATIZACIÓN Y  
CONTROL**

**CERTIFICACIÓN**

Certifico que el trabajo de titulación, “**Desarrollo de un Sistema Inteligente de Sensado como apoyo para la Discriminación de Cacao Sano y Contaminado en sus Procesos de Control de Calidad**” fue realizado por los señores **Catucumbamba Taco, Carlos Andrés y Molina Peláez, Jorge David** el cual ha sido revisado y analizado en su totalidad por la herramienta de verificación de similitud de contenido; por lo tanto cumple con los requisitos legales, teóricos, científicos, técnicos y metodológicos establecidos por la Universidad de las Fuerzas Armadas ESPE, razón por la cual me permito acreditar y autorizar para que lo sustente públicamente.

Sangolquí, 23 de julio de 2021

.....

**Ing. Guamán Novillo, Ana Verónica Ph. D.**

C. C 1103996946



**DEPARTAMENTO DE ELÉCTRICA, ELECTRÓNICA Y  
TELECOMUNICACIONES**

**CARRERA DE INGENIERÍA EN ELECTRÓNICA, AUTOMATIZACIÓN Y  
CONTROL**

**RESPONSABILIDAD DE AUTORÍA**

Nosotros, **Catucumbamba Taco, Carlos Andrés y Molina Peláez, Jorge David**, con cédulas de ciudadanía n° 1725097735 y n° 0107277915, declaramos que el contenido, ideas y criterios del trabajo de titulación: **Desarrollo de un Sistema Inteligente de Sensado como apoyo para la Discriminación de Cacao Sano y Contaminado en sus Procesos de Control de Calidad**, es de nuestra autoría y responsabilidad, cumpliendo con los requisitos legales, teóricos, científicos, técnicos, y metodológicos establecidos por la Universidad de las Fuerzas Armadas ESPE, respetando los derechos intelectuales de terceros y referenciando las citas bibliográficas.

**Sangolquí, 23 de julio de 2021**

**Catucumbamba Taco Carlos Andrés**  
C.C.: 1725097735

**Molina Peláez Jorge David**  
C.C.: 0107277915



**DEPARTAMENTO DE ELÉCTRICA, ELECTRÓNICA Y  
TELECOMUNICACIONES**

**CARRERA DE INGENIERÍA EN ELECTRÓNICA, AUTOMATIZACIÓN Y  
CONTROL**

**AUTORIZACIÓN DE PUBLICACIÓN**

Nosotros **Catucumbamba Taco, Carlos Andrés y Molina Peláez, Jorge David**, con cédulas de ciudadanía n° 1725097735 y n° 0107277915, autorizamos a la Universidad de las Fuerzas Armadas ESPE publicar el trabajo de titulación: **Desarrollo de un Sistema Inteligente de Sensado como apoyo para la Discriminación de Cacao Sano y Contaminado en sus Procesos de Control de Calidad**, en el Repositorio Institucional, cuyo contenido, ideas y criterios son de nuestra responsabilidad.

**Sangolquí, 23 de julio de 2021**

**Catucumbamba Taco Carlos Andrés**  
C.C.: 1725097735

**Molina Peláez Jorge David**  
C.C.: 0107277915

## DEDICATORIA

Este trabajo se lo dedico principalmente a mis padres, quienes me han apoyado durante todo el transcurso de esta etapa de mi vida, a mi hermano que a pesar de todo siempre está a mi lado, y a mi abuelita Gloria que siempre ah estado pendiente de mi cuidándome.

A toda mi familia en mi casa, que siempre se han preocupado por mi bienestar ayudándome a cumplir esta meta en mi vida, ayudándome día a día a prevalecer y no decaer frente a las adversidades de la vida.

También a mi mascota Shaska que ya no esta conmigo, fuiste mi apoyo durante casi toda mi vida, te prometí que iba a cumplir este sueño, quisiera que estés aquí conmigo, este trabajo va para ti también.

*Andrés*

## DEDICATORIA

Este proyecto se lo dedico en especial a mi tío Orlando, a mis abuelitos Rodolfo y Eloisa, quienes desde el cielo sé que me están apoyando en todo momento, y que están muy orgullosos de lo que hago.

A mis padres y mi hermano, que siempre estuvieron a mi lado y pese a mis varias caídas, me han apoyado para levantarme y continuar.

Los quiero, este trabajo va para ustedes.

*Jorge*

## AGRADECIMIENTO

Agradezco a mis padres, en especial a mi mami Olga que siempre me acompaño en esta larga travesía, siempre ha estado a mi lado hasta cuando tenia que amanecerme ella estuvo todo el tiempo acompañándome, tengo la suerte que nunca me ha dejado solo. A mi hermano, por todas las veces que ah estado a mi lado. A mi abuelita por siempre cuidarme y escucharme cuando tenia problemas y hacerme entender que la vida a veces es difícil, pero se tiene que seguir adelante. A mis tíos Javier y Juan que han sido como unos hermanos mayores para mí. A mis mascotas, aunque ellos no hablen me han entregado tanta felicidad cuando mas lo necesitaba en especial a ti Shaska que aunque ya no estés conmigo agradezco todo lo que hiciste por mi.

A mis mejores amigos Edwin y Kevin, pese a que no estamos en la misma universidad siempre me han acompañado en cada logro y en cada tropiezo de mi vida, a ti Jonathan que has estado conmigo en todo momento en la universidad en las buenas y en las malas.

A mis compañeros y amigos de la universidad, en especial a ustedes Alex, Pedro y Jorge, aunque no los conocía desde el inicio se convirtieron en la mejor compañía que pude encontrar dentro de toda la gente en la carrera.

A mi tutora la Ing. Ana Guamán por todo su apoyo y paciencia al momento de desarrollar este proyecto.

*Andrés*

## AGRADECIMIENTO

Primero agradecido con Dios por permitirme culminar esta gran etapa.

Les agradezco en especial a ustedes Papi, Mami y Santy que me dan fuerzas para lograr lo que me proponga, siempre han sido mi motivo para ser quien soy ahora, me llena de alegría saber que los tengo a mi lado compartiendo este valioso momento; a mis primos, tíos, abuelitos que me dieron el aliento para soportarlo todo y con sus consejos fue más fácil alcanzar lo que buscaba.

A ti mi Diani querida, que estuviste en todo momento apoyándome y entregando todo de ti por verme alcanzar mis sueños, pese a mis varios problemas ahí estuviste, de corazón, Dios te pague.

A todos mis amigos, en especial a ti hermano Pedrinho que desde el inicio de la universidad supiste acompañarme y soportarlo todo, como un Dream Team. A mi amigo Andrés por trabajar juntos en este proyecto y entregarme su confianza.

A todos los docentes de la universidad, en especial a la Ing. Ana Guamán por su apoyo incondicional y su paciencia, pues estuvo presente en todo momento que necesitábamos un consejo.

*Jorge*

## Índice de Contenidos

CERTIFICACIÓN .....	3
RESPONSABILIDAD DE AUTORÍA .....	4
AUTORIZACIÓN DE PUBLICACIÓN.....	5
DEDICATORIA.....	6
AGRADECIMIENTO.....	8
Índice de Contenidos.....	10
Índice de Tablas .....	13
Índice de Figuras.....	14
RESUMEN .....	17
ABSTRACT .....	18
Capítulo I.....	19
Introducción .....	19
Antecedentes.....	19
Justificación e Importancia.....	20
Objetivos.....	22
Objetivo General .....	22
Objetivos Específicos .....	22
Descripción del Proyecto.....	23
Capítulo II.....	27
Marco Conceptual.....	27

	11
Control de Calidad de Cacao en Ecuador .....	27
Extracción de la baba .....	28
Fermentación .....	29
Ecurrido .....	29
Secado .....	30
Envasado .....	31
Control de Calidad.....	32
Desinfección y preparación para contenedor.....	36
Arquitectura de Prototipo de Nariz Electrónica .....	37
Redes neuronales artificiales .....	40
Arquitectura de una red neuronal.....	40
Entrenamiento de redes neuronales .....	43
Capítulo III .....	45
Diseño e Implementación del Prototipo.....	45
Diseño del Prototipo.....	45
Instrumentación y acondicionamiento de la matriz de sensado químico .....	45
Adquisición de señales .....	47
Diseño del sistema neumático .....	48
Calibración del sistema .....	50
Diseño funcional y estructural del prototipo.....	54
Creación de Base de Datos .....	59

	12
Pre-procesados de señales .....	60
Filtrado .....	61
Corrección de Línea Base .....	64
Concatenación de Sensores.....	65
Alineación de Picos .....	66
Detección y Eliminación de Outliers .....	67
Modelo de la Red Neuronal .....	69
Capítulo IV .....	73
Validación, Pruebas y Resultados.....	73
Métricas de evaluación del modelo de red neuronal .....	73
Pruebas de desempeño funcional del prototipo .....	76
Análisis de Resultados.....	84
Capítulo V .....	86
Conclusiones y Recomendaciones.....	86
Referencias .....	89

## Índice de Tablas

<b>Tabla 1</b> <i>Características y Clasificación del Cacao Fino</i> .....	21
<b>Tabla 2</b> <i>Número de muestras por tamaño de lote</i> .....	33
<b>Tabla 3</b> <i>Características para la clasificación del cacao.</i> .....	34
<b>Tabla 4</b> <i>Características de los sensores.</i> .....	39
<b>Tabla 5</b> <i>Funciones de Activación.</i> .....	42
<b>Tabla 6</b> <i>Resistencias de calibración</i> .....	52
<b>Tabla 7</b> <i>Tiempos de accionamientos.</i> .....	53
<b>Tabla 8</b> <i>Numero de datos antes y después de la eliminación de Outliers.</i> .....	69
<b>Tabla 9</b> <i>Matriz de Confusión</i> .....	71
<b>Tabla 10</b> <i>Entrenamiento del modelo de red neuronal con 2 neuronas en la capa de entrada.</i> .....	74
<b>Tabla 11</b> <i>Entrenamiento del modelo de red neuronal con 3 neuronas en la capa de entrada.</i> .....	75
<b>Tabla 12</b> <i>Parámetros para el modelo de red neuronal a implementar.</i> .....	77
<b>Tabla 13</b> <i>Comparación del Valor Real esperado y el Valor Predicho por el prototipo.</i> .....	82
<b>Tabla 14</b> <i>Matriz de confusión obtenida en las pruebas del prototipo.</i> .....	83

## Índice de Figuras

<b>Figura 1</b> Diagrama de Bloques Implementación del Sistema .....	23
<b>Figura 2</b> Parámetros Modelo de Aprendizaje.....	25
<b>Figura 3</b> Implementación Red Neuronal.....	25
<b>Figura 4</b> Etapas de preparación granos de cacao.....	28
<b>Figura 5</b> Extracción de la Baba.....	28
<b>Figura 6</b> Fermentación natural del cacao.....	29
<b>Figura 7</b> Escurrido de los granos .....	30
<b>Figura 8</b> Sección de secado .....	30
<b>Figura 9</b> Almacenamiento del Cacao .....	31
<b>Figura 10</b> Transporte de los sacos de cacao .....	32
<b>Figura 11</b> Inspección de los granos de cacao.....	36
<b>Figura 12</b> Almacenamiento del cacao para exportación.....	37
<b>Figura 13</b> Comparación entre olfato biológico y electrónico.....	38
<b>Figura 14</b> Circuito de acondicionamiento.....	46
<b>Figura 15</b> Diseño de placa PCB.....	47
<b>Figura 16</b> Pines GPIO Raspberry Pi3 .....	47
<b>Figura 17</b> Sistema neumático de accionamiento.....	49
<b>Figura 18</b> Esquema de conexión bombas de aire .....	49
<b>Figura 19</b> Accionamiento de las bombas .....	50
<b>Figura 20</b> Curva de respuesta resistencias entre 30 y 40 k $\Omega$ .....	51
<b>Figura 21</b> Curva de funcionamiento sensores.....	52

<b>Figura 22</b> Proceso de olfato artificial.....	54
<b>Figura 23</b> Diseño CAD del prototipo .....	55
<b>Figura 24</b> Cámara de almacenamiento de las muestras. ....	55
<b>Figura 25</b> Cámara de sensores.....	56
<b>Figura 26</b> Bombas de flujo de aire e inyección de olores .....	57
<b>Figura 27</b> Sección de adquisición y control de datos .....	57
<b>Figura 28</b> Suministro de energía del prototipo .....	58
<b>Figura 29</b> Vista general del prototipo implementado .....	59
<b>Figura 30</b> Grupos de grano de cacao .....	59
<b>Figura 31</b> Señales en bruto de los sensores.....	61
<b>Figura 32</b> Espectro de frecuencias .....	62
<b>Figura 33</b> Rendimiento del filtro aplicado .....	63
<b>Figura 34</b> Línea base estimada por sensor.....	64
<b>Figura 35</b> Señales ajustadas al origen.....	65
<b>Figura 36</b> Señales concatenadas de los 6 sensores.....	66
<b>Figura 37</b> Resultado de la alineación de picos en las señales procesadas.....	67
<b>Figura 38</b> Conjunto de datos final .....	68
<b>Figura 39</b> Descenso del error cuadrático medio durante el entrenamiento del modelo con dos y tres neuronas en la capa de entrada. ....	76
<b>Figura 40</b> Granos de cacao en la cámara de olores.....	77
<b>Figura 41</b> Conexión de prototipo a las fuentes de alimentación. ....	78
<b>Figura 42</b> Menú de selección para comenzar el proceso de olfato.....	79

<b>Figura 43</b> Niveles de voltaje de los sensores químicos durante el funcionamiento del prototipo. ....	79
<b>Figura 44</b> Archivo .txt con las mediciones realizadas por cada sensor a las muestras de cacao.....	80
<b>Figura 45</b> Señales medidas por los sensores por cada muestra de granos de cacao. ....	81
<b>Figura 46</b> Resultado del preprocesamiento de señales para una muestra de granos de cacao.....	81

## RESUMEN

El control de calidad para la exportación de cacao en Ecuador es un proceso que actualmente se realiza de forma manual, con lo que ofrecen a sus clientes un producto correctamente cultivado, cosechado y preparado para su exportación. En el presente trabajo se desarrolla un prototipo capaz de asistir el control de calidad, que, con ayuda de una nariz electrónica conformada por una matriz de sensores químicos, mida sustancias volátiles emanadas por los granos de cacao como amoníaco, vapores orgánicos solventes, metano, monóxido de carbono y óxido nitroso. El prototipo consta de cinco partes: una cámara de concentración de olores donde se colocan las muestras de granos de cacao, una cámara de medición con el arreglo de sensores químicos en su interior, un sistema neumático para el bombeo temporizado de aire controlando el paso de las sustancias volátiles entre cámaras, el área de adquisición-almacenamiento de señales manejado por una tarjeta Raspberry Pi3 y un modelo de red neuronal clasificador diseñado, entrenado e implementado en Python dentro de un ordenador, desde donde se controla el prototipo a través de una terminal. Luego de realizar 98 experimentos con 50 granos de cacao sanos y 48 granos contaminados, el prototipo permite discriminar las muestras de cacao sano de los que se encuentran contaminados con una exactitud del 95%.

### **PALABRAS CLAVE:**

- **REDES NEURONALES.**
- **NARIZ ELECTRÓNICA.**
- **CONTROL DE CALIDAD DE CACAO.**
- **MATRIZ DE SENSORES QUÍMICOS.**

## **ABSTRACT**

The quality control for the export of cocoa in Ecuador is a process that is currently carried out manually, with which they offer their customers a product that is properly cultivated, harvested, and prepared for export. In this work, a prototype capable of assisting in quality control is developed, which, with the help of an electronic nose made up of a matrix of chemical sensors, measures volatile substances emanating from cocoa beans such as ammonia, organic solvent vapors, methane, carbon monoxide and nitrous oxide. The prototype consists of five parts: an odor concentration chamber where the cocoa bean samples are placed, a measurement chamber with the arrangement of chemical sensors inside, a pneumatic system for timed pumping of air controlling the passage of the volatile substances between chambers, the signal acquisition-storage area managed by a Raspberry Pi3 card and a classifier neural network model designed, trained and implemented in Python within a computer, from where the prototype is controlled through a terminal. After carrying out 98 experiments with 50 healthy cocoa beans and 48 contaminated beans, the prototype allows us to discriminate healthy cocoa samples from those that are contaminated with an accuracy of 95%.

### **KEYWORDS:**

- **NEURAL NETWORKS.**
- **ELECTRONIC NOSE.**
- **COCOA QUALITY CONTROL.**
- **MATRIX OF CHEMICAL SENSORS.**

## CAPÍTULO I

### Introducción

#### Antecedentes

El uso de una nariz electrónica tiene sus orígenes en los años 80 en Gran Bretaña, en donde se empleó un conjunto de sensores semiconductores de óxidos metálicos que permitirían detectar, identificar y medir varios productos químicos transportados en ferrocarriles, barcos y camiones (Persaud & Dodd, 1982). Es en base a esta investigación que, en el año de 1982 Gardner y Bartlett dan la primera definición de una nariz electrónica como un sistema inteligente diciendo que es: "Instrumento que comprende una agrupación de sensores químicos con sensibilidades parcialmente solapadas junto a un sistema de reconocimiento de patrones, capaz de analizar y reconocer aromas simples o complejos" (Gardner & Barlett, 1999).

Actualmente, una nariz electrónica no sólo se utiliza para la clasificación de olores (Moreno, Caballero, Galán, Matía, & Jiménez, 2009) (Distante, Ancona, & Siciliano, 2003), sino también para la detección de estos, de esta manera su aplicación se extendió a diversos campos como: la agroindustria (Benedetti, Mannino, Sabatini, & Luigi, 2004), con temas enfocados en la determinación de la calidad de productos alimenticios (Lebrun, Plotto, Goodner, Ducamp, & Baldwin, 2008) (Rosales & Reyes, 2016) (Rodríguez, 2001) el medio ambiente, haciendo énfasis en el aspecto de contaminación del aire (Nagle, Gutierrez, & Schiffman, 1998); medicina, utilizada principalmente en la detección de infecciones (Persaud, Pisanelli, Evans, & Travers, 2006); y seguridad, siendo empleada en controles de explosivos (Brudzewski, Osowski, & Pawlowski, 2012).

Tal es el caso, del proyecto realizado por la Universidad Nacional Abierta y a Distancia de Colombia, el cual consiste en un sistema capaz de discriminar los granos de cacao sanos de los que se encuentran infectados con el hongo monilia. El sistema logró alcanzar una varianza del 87% en la discriminación y es apto para ser

implementado utilizando redes neuronales (Flórez, Vargas, & Harold, 2020). Sin embargo, las investigaciones existentes tienen como principal objetivo la automatización de los procesos en la cadena de producción de granos de cacao, mientras que el control de calidad se lo realiza de forma manual ya sea por análisis químico, visual o por catación (Quiñónez, Salvatierra, Cadena, & Baque, 2020).

Tomando en cuenta los antecedentes mencionados, el presente trabajo se enfoca en el diseño e implementación de un sistema basado en una matriz de sensores químicos, combinada con un modelo de aprendizaje que sea capaz de asistir en la clasificación de granos de cacao sanos de los que presenten elementos extraños en su aroma, asegurando así la calidad del producto.

### **Justificación e Importancia**

Ecuador es considerado como productor por excelencia del tipo de cacao denominado Cacao Arriba fino y de aroma, debido a sus condiciones geográficas y recursos biológicos ocupando el 63% de la producción mundial en este tipo de cacao.

Los procesos de producción de los granos de cacao se basan en una estrategia que utiliza sistemas automatizados dentro de las etapas post cosecha, mientras que para el control de calidad del grano se realizan: pruebas de corte luego de la fermentación, análisis químico para la observación de niveles de pH y cadmio, además de una etapa de catación; los cuales son métodos manuales. De este modo se obtienen distintos tipos de granos beneficiados (en correcto estado) con características mostradas en la Tabla 1.

**Tabla 1**

*Características y Clasificación del Cacao Fino (Servicio Ecuatoriano de Normalización, 2006)*

Requisitos	Cacao Fino		
	A.S.S.S	A.S.S	A.S.E
<b>Humedad, máximo, %</b>	7	7	7
<b>Peso de 100 granos, g</b>	>130	>120 a 130	100 a 120
<b>Granos fermentados, mínimo, %</b>	75	65	53
<b>Granos violetas, máximo, %</b>	15	21	25
<b>Granos Pizarrosos, máximo, %</b>	9	12	18
<b>Granos Mohosos, máximo, %</b>	1	2	4
<b>TOTALES (análisis sobre 100 granos), mínimo</b>	100	100	100
<b>Granos defectuosos, máximo, %</b>	0	1	3
<b>Material relacionado al cacao, máximo, %</b>	1	1	1
<b>Material extraño, máximo, %</b>	0	0	0

*Nota:* En la Tabla se observan los requisitos que deben cumplir los granos de cacao según la categoría.

En donde las siglas tienen los siguientes significados:

- ASSS: Arriba Superior Summer Selecto.
- ASS: Arriba Superior Selecto.
- ASE: Arriba Superior Época.

El 90% del cacao ecuatoriano se exporta en grano y el 54% del cacao fino producido en América Latina pertenece al Ecuador, del cual, existe una demanda en crecimiento que tiene una preferencia en los sectores más exigentes fabricantes de chocolates finos. Aproximadamente un 70% de las exportaciones de granos de cacao es del tipo Arriba (Plaza, 2016).

Los requisitos establecidos en la norma técnica ecuatoriana INEN 176, indican que el grano de cacao beneficiado no debe estar infestado, debe estar libre de olores, tales como: ácido butírico (podrido), moho, humo o químicos agrícolas principalmente, esto aplica para cacao fino o de aroma como para cacao corriente (Servicio Ecuatoriano de Normalización, 2006).

Por lo tanto, son necesarios sistemas electrónicos automáticos capaces de aumentar la eficiencia en los controles de calidad para la clasificación de los granos de cacao. Es así, que se propone el diseño de un sistema basado en un arreglo de sensores químicos combinado con un modelo de aprendizaje, como una herramienta que ayude en la discriminación entre cacao sano y contaminado, para mantener un nivel óptimo de calidad en la producción de cacao arriba.

## **Objetivos**

### **Objetivo General**

Diseñar e implementar un sistema electrónico de medición basado en sensores químicos y redes neuronales, como apoyo tecnológico para discriminar granos de cacao sanos o contaminados en sus procesos de control de calidad.

### **Objetivos Específicos**

- Implementar una cámara de extracción, con una matriz de sensores químicos para la detección de olores de las muestras de granos de cacao.
- Implementar un algoritmo en Python para la adquisición y almacenamiento de las mediciones.
- Diseñar e implementar la etapa de adquisición y procesamiento de señales basadas en instrumentación electrónica.
- Implementar una base de datos de las muestras de granos de cacao para la optimización de los parámetros en la red neuronal.

- Analizar el desempeño del sistema de medición para su validación.

### Descripción del Proyecto

El presente proyecto de investigación tiene como finalidad desarrollar un modelo de red neuronal capaz de discriminar entre una muestra de cacao sano de una muestra contaminada, para esto el prototipo de sistema está compuesto por dos fases, la primera corresponde al diseño e implementación de un sistema de medición basado en instrumentación electrónica, y la segunda fase es sobre la obtención de los parámetros de un modelo de aprendizaje.

### Figura 1

*Diagrama de Bloques Implementación del Sistema (Elaboración propia)*



*Nota:* En la figura se muestra las acciones asociadas a cada etapa del desarrollo del sistema.

Dentro de la **primera fase** se encuentran las etapas: cámara de extracción de olores, sistema de adquisición de datos, control-procesamiento de señales, almacenamiento de datos en el ordenador y visualización de resultados en el ordenador, tal como se observa en la Figura 1.

En primer lugar, se obtendrán las muestras de grano de cacao sanos y otras contaminadas. Luego, como etapa de medición, dentro de una cámara extractora de olores, se deja reposar la muestra de grano de cacao para que se concentre el aroma. Posteriormente ingresa y sale el aire necesario a la cámara para realizar el proceso de olfato, analizando así la sustancia de interés mediante un arreglo de sensores químicos.

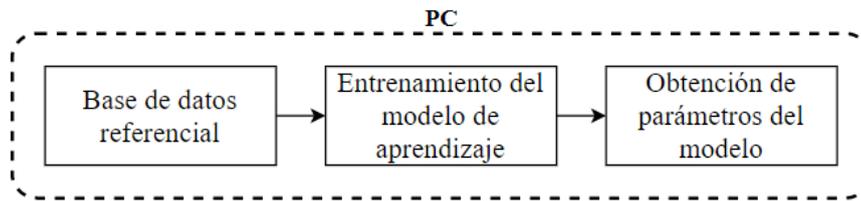
Una vez efectuada la transducción, las señales son acondicionadas y se transforman mediante un conversor análogo-digital en señales que sean compatibles con la tarjeta controladora Raspberry, quien almacena los datos y luego los envía hacia un ordenador en donde de igual manera se almacena la información de los sensores.

Mientras que, como **segunda fase** se encuentra la generación de una base de datos referencial con los valores medidos de las muestras de granos de cacao, información que servirá para el cálculo de los parámetros de un modelo de aprendizaje basado en redes neuronales, capaz de discriminar los granos sanos de los contaminados.

En la Figura 2 se puede apreciar el esquema para obtener los parámetros del modelo de aprendizaje, a partir de un entrenamiento de la red neuronal, la cual tiene como entrada (patrones) la base de datos referencial.

## Figura 2

*Parámetros Modelo de Aprendizaje (Elaboración Propia).*

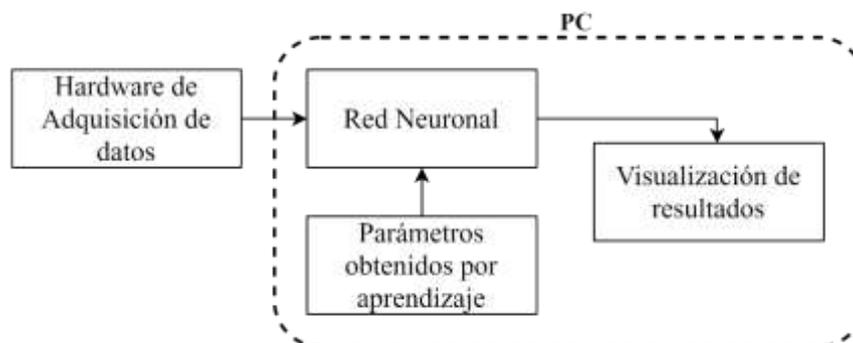


*Nota:* En la figura se muestra un diagrama de cómo se encuentra conformado el proceso de entrenamiento del modelo de aprendizaje.

Por último, en la Figura 3 se muestra la implementación de la red neuronal en el ordenador, utilizando los parámetros obtenidos del modelo de referencia, que, en conjunto con el sistema de instrumentación, serán capaces de identificar si una muestra determinada se la puede considerar o no como cacao sano.

## Figura 3

*Implementación Red Neuronal (Elaboración Propia)*



*Nota:* En la figura se muestra un diagrama de bloques del funcionamiento general del prototipo.

Para el desarrollo de las fases antes descritas el proyecto se divide en 4 capítulos, en cada uno se detalla el proceso para la creación del prototipo y los resultados obtenidos en su implementación.

**Capítulo II: Marco Conceptual**

En este capítulo se presenta la información sobre el proceso actual para la obtención de los granos de cacao y su control de calidad, además se incluyen datos técnicos sobre la arquitectura del prototipo, así como conceptos para la implementación de una red neuronal.

**Capítulo III: Diseño e Implementación del Prototipo**

En este capítulo se describen los parámetros que se utilizarán para el desarrollo del prototipo de nariz electrónica, además de las diferentes etapas para el tratamiento de las señales de interés para el entrenamiento de la red neuronal, del mismo modo se describe los modelos para el desarrollo de la red neuronal.

**Capítulo IV: Validación, Pruebas y Resultados**

El capítulo 4 muestra las métricas para determinar la efectividad del prototipo implementado realizando un análisis de las pruebas realizadas en muestras aleatorias de granos de cacao.

**Capítulo V: Conclusiones y Recomendaciones**

En este capítulo se muestran las conclusiones obtenidas en base a los objetivos propuestos, del mismo modo se presentan recomendaciones para el desarrollo de trabajos futuros relacionados a la implementación de narices electrónicas en procesos de control de calidad.

## Capítulo II

### Marco Conceptual

En este capítulo se abordan los conceptos, definiciones y situación actual del proceso de obtención de los granos de cacao. Además, se realiza la descripción de cada elemento que conforma el prototipo y las métricas utilizadas para el entrenamiento de la red neuronal. Finalmente se obtiene un diseño base para el prototipo y el modelo de red neuronal a implementar.

### Control de Calidad de Cacao en Ecuador

El proceso de preparación de los granos de cacao en Ecuador para su exportación depende de cada una de las exportadoras, y para observar el control de calidad se realizó una visita técnica a la exportadora de cacao ECO-KAKAO S.A. ubicada en San Jacinto de Buena Fé – Los Ríos, en donde ANECACAO es la asociación encargada de corroborar si el producto cumple o no con la calidad acordada entre el exportador y el cliente.

La exportadora asegura que cada etapa de tratamiento de los granos de cacao sea idónea, desde el cuidado en el campo agrícola hasta el momento en que se embarcan en sacos hacia los contenedores para su exportación. En la Figura 4 se muestran las etapas por las que pasan los granos de cacao:

## Figura 4

*Etapas de preparación granos de cacao (Elaboración Propia)*



*Nota:* En la figura se muestra los pasos del procesado de los granos de cacao para su exportación.

### Extracción de la baba

Se deposita el cacao en baba cosechado en una banda transportadora, la cual asciende hacia una tolva como se observa en la Figura 5, que dirige el cacao hacia los mecanismos de extracción y separación de baba. El producto terminado es el grano de cacao con un 30% de baba aproximadamente, que se transporta en carretas hacia su proceso de fermentación.

## Figura 5

*Extracción de la Baba (Elaboración Propia)*



*Nota:* En la figura se muestra una fotografía de la máquina extractora de baba del cacao.

## Fermentación

Se almacena un cúmulo de granos dentro de una capa de yute y de plástico como se puede apreciar en la Figura 6, en donde se deja reposar de 4 a 5 días bajo un invernadero con temperatura ambiente de 30 a 43 grados centígrados, mientras mayor sea la temperatura mejor es la fermentación. Se registra la temperatura del lugar tres veces al día de forma manual, se “voltean” los granos cada 24 horas con motivo de que la fermentación sea uniforme, y se realizan pruebas de corte de los granos cada 48 horas para comprobar el grado de fermento.

### Figura 6

*Fermentación natural del cacao (Elaboración Propia)*



*Nota:* La figura muestra el almacenamiento del grano de cacao para su fermentación.

## Ecurrido

Previo a llevarlos hacia las secadoras, se deja reposar los granos esparcidos en el suelo con palas de madera como se muestra en la Figura 7 (para no maltratar los granos) de uno a dos días.

**Figura 7**

*Escurreo de los granos (Elaboración Propia)*



*Nota:* La figura muestra los granos de cacao esparcidos para su escurrido natural.

**Secado**

Con ayuda de máquinas alimentadas a gas como se puede apreciar en la Figura 8, se realiza el proceso de secado durante 2 a 3 días, con temperaturas entre 45 a 48 grados centígrados. El control es manual por parte del operador que consta de dos horas de calor y una hora de ventilación.

**Figura 8**

*Sección de secado (Elaboración propia)*



*Nota:* La figura presenta la maquinaria utilizada para el proceso de secado del grano.

## Envasado

Dentro de este proceso se realiza una prueba de corte de los granos por guillotina donde se examinan las propiedades organolépticas a través de su sabor y olor (ya sea de nuez, flores, frutales, etc.). El siguiente paso es almacenar los granos de cacao en sacos de yute. En el caso de ser un cacao orgánico o grano selecto, como se puede apreciar en la Figura 9, se realiza un cuidado extra guardándolos dentro de fundas GRAINPRO absorbiendo la mayor cantidad de aire para permitir que el grano mantenga sus características y que todo tipo de insecto muera debido a la falta de oxígeno.

### Figura 9

*Almacenamiento del Cacao (Elaboración Propia)*



*Nota:* La figura muestra los sacos de granos de cacao almacenados por lote dentro de fundas GRAINPRO.

## Control de Calidad

Los granos que se encuentren correctamente fermentados y secados se llevan hacia tolvas que permiten llenar los sacos controlando su peso como se observa en la Figura 10, y con ayuda de un montacargas se transportan los sacos hacia el lugar de paletizado.

### Figura 10

*Transporte de los sacos de cacao (Elaboración Propia)*



*Nota:* La figura muestra el transporte de los granos de cacao al área de paletizado.

Una vez almacenados los sacos en un lote, se procede a realizar el muestreo según la NTE INEN 177 en donde se toma una muestra elemental a cada saco de forma aleatoria, la muestra elemental consta de obtener un puñado de granos de cualquier parte del saco. En la visita, el exportador tenía un lote con aproximadamente 1400 sacos, por lo que fue necesario escoger 125 sacos al azar y tomar una muestra elemental de cada uno, como se indica en la Tabla 2 (Servicio Ecuatoriano de Normalización, 1995).

**Tabla 2**

*Número de muestras por tamaño de lote (Servicio Ecuatoriano de Normalización, 2006)*

<b>Tamaño del lote (número de sacos)</b>	<b>Número mínimo de muestras elementales</b>
2 - 8	2
9 - 15	3
16 - 25	5
26 - 50	8
51 - 90	13
91 - 150	20
151 - 280	32
281 - 500	50
501 - 1.200	80
1.201 - 3.200	125
3.201 - 10.000	200
10.001 - 35.000	315
35.001 - 150.000	500
150.001 - 500.000	800
Mayor a 500.001	1250

*Nota:* La tabla presenta el número de muestras que se deben tomar dependiendo del número de sacos obtenidos por lote.

Se guardan aquellas muestras elementales en un recipiente y se las lleva al control de calidad, que actualmente en Ecuador se lo realiza de forma visual basado en la tabla de requisitos de la norma NTE INEN 176, que clasifica a los granos de cacao en distintos tipos (Servicio Ecuatoriano de Normalización, 2006).

El exportador le ofreció a su cliente granos de calidad A.S.S. o también llamados Doble "S", por lo que primero se pesan 500 gramos del recipiente de muestras elementales con granos escogidos al azar. Luego de esto se separan los granos en tres grupos con cien pepas cada uno con el fin de obtener su peso y calcular una media de aquellas tres mediciones. Si aquel promedio se encuentra dentro de los 120 a 125 gramos, se puede proceder con las siguientes pruebas. En el control de calidad efectuado, se obtuvo un peso para calidad de grano A.S.S.S. o Triple "S", que

es mucho mejor que el grano Doble acordado por el exportador quien pudo haber vendido como uno de mejor calidad.

Para el análisis de los granos según su grado de fermentación, se escoge un grupo de cien granos como se indica en la Tabla 3 y se procede a cortar de forma longitudinal cada grano, donde se observa el color que presenta en su interior y se realiza un conteo de los granos con:

- Buena fermentación: Se observa un color marrón dentro del grano.
- Ligera fermentación: Presenta una mezcla entre marrón y violeta.
- Color violeta en su interior.
- Apariencia pizarrosa: Presenta un color violeta oscuro.
- Presencia de moho.

**Tabla 3**

*Características para la clasificación del cacao. (Servicio Ecuatoriano de Normalización, 2006)*

Requisitos	Unidad	Cacao Arriba					CCN51
		A.S.S.P.S	A.S.S.S	A.S.S.	A.S.N.	A.S.E.	
<b>Cien granos pesan</b>	<b>g</b>	<b>135-140</b>	<b>130-135</b>	<b>120-125</b>	<b>110-115</b>	<b>105-110</b>	<b>135-140</b>
Buena fermentación (mín.)	%	75	65	60	44	26	***65
Ligera fermentación* (mín.)	%	10	10	5	10	27	11
Violeta (máx.)	%	10	15	21	25	25	18
Pizarroso (máx.)	%	4	9	12	18	18	5
Moho (máx.)	%	1	1	2	3	4	1
Totales (análisis sobre 100 pepas)	%	100	100	100	100	100	100

Requisitos	Unidad	Cacao Arriba					CCN51
		A.S.S.P.S	A.S.S.S	A.S.S.	A.S.N.	A.S.E.	
Cien granos pesan	g	135-140	130-135	120-125	110-115	105-110	135-140
Defectuosos (análisis sobre 500 gramos) (máx.)	%	0	0	1	3	**4	1
Total fermentado (mín.)	%	85	75	65	54	53	76
A.S.S.P.S.	Arriba Superior Summer Plantación Selecta						
A.S.S.S.	Arriba Superior Summer Selecto						
A.S.S.	Arriba Superior Selecto						
A.S.N.	Arriba Superior Navidad						
A.S.E.	Arriba Superior Época						
* Coloración marrón violeta							
** Se permite la presencia de granza solamente para el tipo A.S.E.							
*** La coloración varía de marrón a violeta							

*Nota:* La tabla presenta los porcentajes de características que determinan la clasificación del cacao.

El número obtenido por cada requisito es directamente un porcentaje que debe ser registrado según la NTE INEN 175, mientras que para contar los granos defectuosos se lo debe realizar dentro de la muestra total de los 500 gramos como se aprecia en la Figura 11 (Servicio Ecuatoriano de Normalización., 1987).

Luego se puede identificar el tipo de cacao al que pertenecen los granos del lote seleccionado ya sea de tipo Cacao Arriba o CCN51, según los porcentajes mínimos en fermentación y máximos en granos defectuosos, de color violeta, pizarrosos y con moho especificados en la Tabla 3. Mientras que, en caso de que el lote no cumpla los requisitos del control de calidad, se debe reprocesar todo el lote desde la etapa de fermentación.

**Figura 11**

*Inspección de los granos de cacao (Elaboración Propia)*



*Nota:* La figura muestra el proceso de control de calidad que se lleva a cabo actualmente.

A partir de los granos recolectados al azar inicialmente, se escoge al final una pequeña muestra de 1 kg para un último análisis que depende del país de origen del comprador:

- Japón requiere un análisis sobre la concentración de productos químicos que presenten los granos.
- La Unión Europea, antes de ser despachado el cacao, necesita un análisis de concentración de cadmio y metales pesados.

**Desinfección y preparación para contenedor**

Los sacos que se encuentren listos dentro del lote para exportación son nuevamente aislados con las fundas GRAINPRO sobre pallets de madera como se muestra en la Figura 12, para conservarlos hasta el embarque dentro del contenedor.

**Figura 12**

*Almacenamiento del cacao para exportación (Elaboración Propia)*



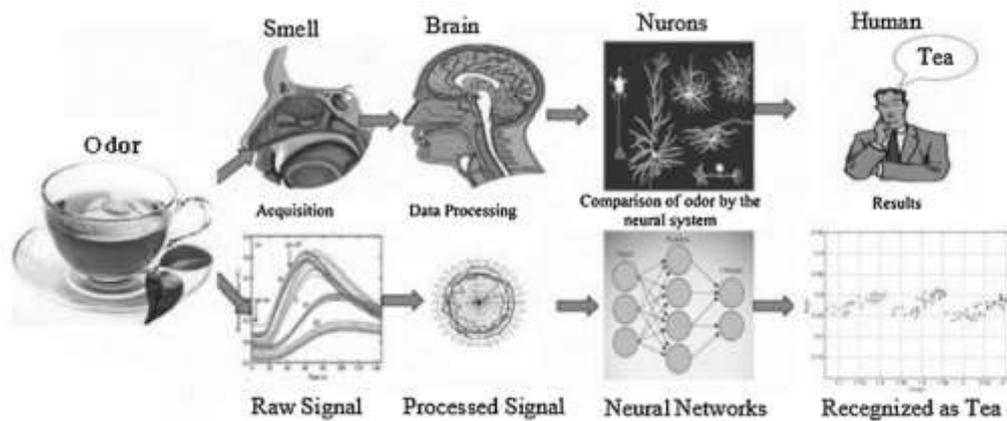
*Nota:* La figura muestra el almacenamiento de los lotes de cacao previo a su exportación.

**Arquitectura de Prototipo de Nariz Electrónica**

El objetivo principal de una nariz electrónica es emular el sentido del olfato de un ser humano para determinar el tipo de sustancia que se analiza en la Figura 13 se muestra un diagrama de bloques con una analogía del proceso de olfato biológico que se puede replicar por uno electrónico.

### Figura 13

*Comparación entre olfato biológico y electrónico. (Sajad, Saeid & Mahdi, 2015)*



*Nota:* La figura muestra el equivalente electrónico de las diferentes etapas del olfato biológico.

La sustancia para analizar son los volátiles característicos de los granos de cacao, los cuales en el artículo “*Identificación de los compuestos aromáticos en el cacao criollo de Venezuela usando micro extracción en fase sólida y cromatografía de gases*” desarrollada por Álvarez, Pérez y Boulanger (2012), se identificaron las siguientes familias de compuestos químicos:

- Aldehídos y cetonas
- Esteres
- Alcoholes
- Pirazinas
- Monóxido de carbono

Al definir los compuestos químicos a analizar se inició con el diseño del prototipo el cual se basa en el modelo propuesto en la investigación “*Desarrollo de nariz electrónica para calidad en los procesos críticos de transformación del cacao*”

desarrollada por Flórez, Vargas y Perez (2020). Donde, se define que el prototipo debe constar de tres elementos fundamentales:

- **Matriz de sensores químicos:** En base a los volátiles presentes en los granos de cacao en el proceso de control de calidad se determinó que la matriz sensorial debe estar conformada por un conjunto de 6 sensores de gases cuyas referencias se presentan en la Tabla 4.

**Tabla 4**

*Características de los sensores (Elaboración Propia)*

Sensor	Sensibilidad
TGS-2602	Gases olorosos (Amoniac y Disulfuro de hidrogeno)
TGS-826	Amoniac
TGS-822	Vapores orgánicos solventes
MQ-7	Monóxido de carbono
MQ-4	Gas natural y Metano
MQ-135	Óxido Nitroso

*Nota:* La tabla presenta el químico al cual es más sensible cada uno de los sensores a implementar.

- **Sistema de Adquisición de datos:** Para la selección del microcontrolador se toma en cuenta la capacidad de procesamiento de datos y el número de terminales E/S analógicos y digitales, en el desarrollo del proyecto se tomó a la Raspberry Pi3 como sistema de adquisición debido a su alta capacidad de procesamiento de grandes tramas de datos y adaptación de módulos para E/S analógicas y digitales. La función

principal de la tarjeta es la de recibir información proveniente del conjunto de sensores y realizar el procesamiento de dichas señales.

- **Software de Clasificación:** Con respecto al software utilizado se optó por una plataforma de desarrollo en código Python el cual es un lenguaje de programación de alto nivel que posee una amplia gama de librerías para el uso de redes neuronales y procesamiento de señales

### **Redes neuronales artificiales**

En el presente proyecto, un clasificador basado en redes neuronales asiste al control de calidad de granos de cacao, en donde la red neuronal es un conjunto de elementos simples interconectados de forma paralela por distintas capas, los cuales tienen propiedades dinámicas y adaptativas que permiten procesar información de forma similar al funcionamiento de las neuronas que se encuentran en el sistema nervioso (Samarasinghe, 2007).

### **Arquitectura de una red neuronal**

Una red neuronal está constituida por un conjunto de *perceptrones* o neuronas artificiales interconectadas que al ser distribuidas en varias capas se las conoce como *redes multicapa*. El perceptrón por sí sólo cumple una función de discriminador lineal, donde puede tomar decisiones a partir de una combinación lineal de sus entradas, mientras que las redes multicapa permiten resolver problemas que no se pueden separar linealmente.

Las redes neuronales están conformadas por una *capa de entrada* cuyas características o entradas están asignadas hacia una *capa de salida* utilizando una variación generalizada de función lineal. Además, existe una capa oculta (opcional) con una o varias neuronas interconectadas, que separan a las capas de entrada y salida (Fauset, 1994).

La capa de entrada recibe las  $i$  características de las variables de entrada  $\bar{X} = [x_1 \dots x_i]$  con sus respectivos pesos  $\bar{W} = [w_1 \dots w_i]$ , hacia el nodo de salida, en donde se efectúa la función lineal presentada en la ecuación (1).

$$\bar{W} \cdot \bar{X} = \sum_{j=1}^i w_j x_j \quad (1)$$

Además, de la notación matemática sumatoria, existen otras dos notaciones que son muy conocidas como *productorio* que se observa en la ecuación (2) y el *máximo* de las entradas indicado en la ecuación (3).

**Productorio:**

$$\bar{W} \cdot \bar{X} = \prod_{j=1}^i w_j x_j \quad (2)$$

**Máximo:**

$$\bar{W} \cdot \bar{X} = \text{Max}(w_j x_j) \quad (3)$$

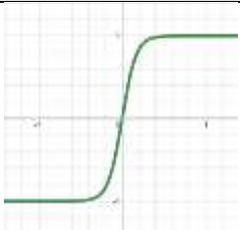
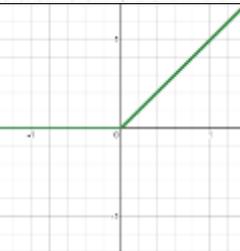
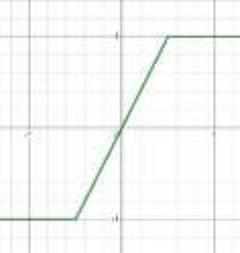
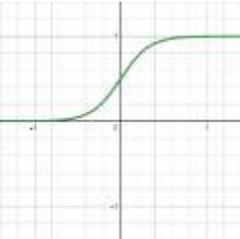
En algunas de las predicciones realizadas con redes neuronales se suele añadir una constante denominada *bias* para capturar la invarianza en la predicción cuando la media de las características no son cero. Por consiguiente, aquella variable predicha conocida como  $\hat{y}$ , puede calcularse utilizando la función signo indicada en la ecuación (4), la cual es muy utilizada en clasificadores binarios (Ketkar, 2017).

$$\hat{y} = \text{sign}\{\bar{W} \cdot \bar{X} + b\} = \text{sign}\left\{\sum_{j=1}^i w_j x_j + b\right\} \quad (4)$$

En la Ecuación (4) se puede observar a la función signo la cual se la conoce como *función de activación* que permite interpretar los valores de entrada como una clase. Las funciones de activación más utilizadas se detallan en la Tabla 5.

Tabla 5

*Funciones de Activación (Elaboración Propia)*

Función de activación	Descripción	Visualización
<b>Tangente Hiperbólica</b>	$f(x) = \frac{e^{g \cdot x} - e^{-g \cdot x}}{e^{g \cdot x} + e^{-g \cdot x}}$ <p>Los valores de <math>f(x)</math> se encuentran entre -1 y 1, la pendiente de la función de activación depende de la constante <math>g</math>.</p>	
<b>ReLU (Unidad Lineal Rectificada)</b>	$f(x) = \begin{cases} x & x \geq 0 \\ 0 & x < 0 \end{cases}$ <p>Los valores de <math>f(x)</math> siempre serán positivos, además de ser una función continua a trozos.</p>	
<b>Lineal</b>	$f(x) = \begin{cases} -1 & x \leq -\frac{1}{g} \\ g \cdot x & -\frac{1}{g} < x < \frac{1}{g} \\ 1 & x \geq \frac{1}{g} \end{cases}$ <p>Se define una función lineal durante un tramo entre <math>-\frac{1}{g}</math> y <math>\frac{1}{g}</math>, cuyos valores de <math>f(x)</math> estarán entre -1 y 1</p>	
<b>Sigmoide</b>	$f(x) = \frac{1}{1 + e^{-g \cdot x}}$ <p>Los valores de <math>f(x)</math> se encuentran entre 0 y 1, la pendiente de la función depende de la constante <math>g</math>.</p>	

*Nota:* La tabla presenta las definiciones para cada una de las funciones de activación.

Es importante escoger correctamente la función de activación, por ejemplo, si se desea un valor real a la salida de la red neuronal se puede utilizar una función de activación identidad, mientras que, si se busca predecir la probabilidad de una clase binaria, es preferible utilizar una función sigmoide para la activación (García, 2001).

El último elemento luego de la función de activación es la función de salida de una neurona, en donde las dos funciones más utilizadas son:

- Binaria: En donde se necesita rebasar un umbral para activar la salida, caso contrario permanece desactivada.
- Identidad: La salida es la misma que la entregada por la función de activación, siendo el tipo de función de salida más sencilla.

### **Entrenamiento de redes neuronales**

Dentro del diseño de un modelo de red neuronal, se debe considerar qué función de activación es la que mejor desempeño entrega para el problema. En el presente proyecto es necesario el uso de una función de activación para clasificación, por lo que cualquier función indicada en la Tabla 5 puede ser utilizada.

Una vez seleccionada la función de activación, es necesario definir los *Hiperparámetros* de la red neuronal, ya sea el número de neuronas por cada capa, el número de entradas y/o el número de *épocas* para el entrenamiento. El objetivo del entrenamiento es mejorar el desempeño de la red neuronal minimizando el error de predicción a través de una adaptación de los pesos en un proceso iterativo; cada iteración es denominada una *época* (Hush, 1989).

Al referirse a un entrenamiento de la red neuronal, también se debe hablar sobre los tipos de aprendizaje, que pueden ser tanto *supervisados* en el cual se conocen los valores de salida de la red por cada conjunto de valores de entrada, como *no supervisados* en donde no se conocen los valores de salida. Otro caso a tomar en cuenta es cuando se necesita que la red aprenda mientras se encuentra en su funcionamiento normal, también conocido como aprendizaje *online*, o en el caso en que se realiza un aprendizaje cuando el sistema no está en funcionamiento, conocido como aprendizaje *offline*, este último tipo de aprendizaje es el que se implementa para el prototipo (Galushkin, 2007).

Por consiguiente, previo al entrenamiento de la red es necesario definir los siguientes tres parámetros (Aggarwal, 2018).

- **Función de pérdida:** La forma en que la red mide su desempeño en los datos de entrenamiento. La función de pérdida se optimiza durante el entrenamiento. La función utilizada fue basada en el ajuste del error por mínimos cuadrados.
- **Optimizador:** Es un mecanismo con el que la red se actualiza a sí misma, basándose en los datos que observa y en su función de pérdida. Se utilizó el algoritmo optimizador Adam para el presente trabajo.
- **Métricas de evaluación:** En este caso se mide la exactitud, precisión, sensibilidad y puntaje F1 con la que la red clasifica los datos.

Durante el desarrollo del prototipo detallado en el Capítulo III se encuentra una explicación sobre la obtención de la función de pérdida, mecanismo optimizador y métricas de evaluación utilizadas para el proyecto.

Luego se procede con el entrenamiento y una vez finalizado, los pesos de la red neuronal permanecen fijos, con los que ya se puede implementar la red y evaluar con las métricas de desempeño seleccionadas.

## Capítulo III

### Diseño e Implementación del Prototipo

En el presente capítulo se explicará el desarrollo técnico del prototipo nariz electrónica. Se inicia con una explicación de la instrumentación en la matriz de sensado químico, luego con una descripción sobre el proceso para la adquisición de los datos y desde la obtención de la línea base hasta conformar una base de datos necesaria para el entrenamiento del sistema. Finalmente se explican a detalle las herramientas matemáticas y de software que se utilizaron durante el desarrollo del modelo.

### Diseño del Prototipo

#### Instrumentación y acondicionamiento de la matriz de sensado químico

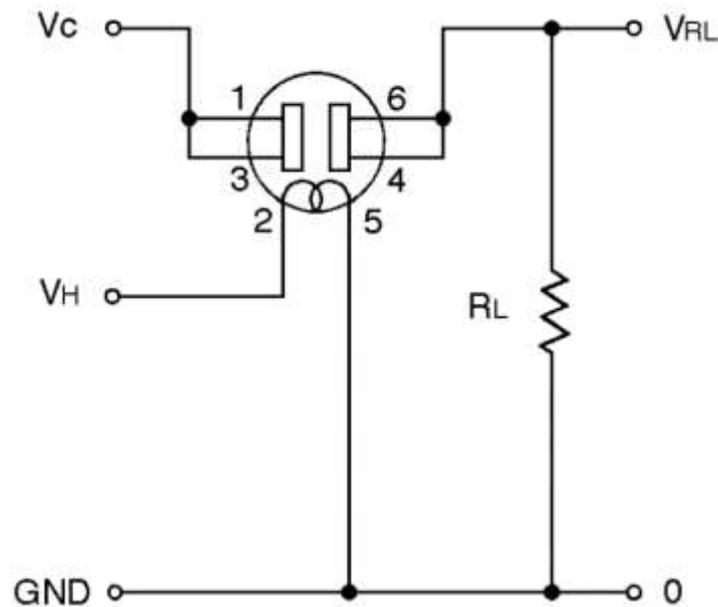
Los sensores seleccionados en el Capítulo II utilizan un Semiconductor de Óxido Metálico (MOS por sus siglas en inglés) como principal elemento para la detección de sustancias volátiles, el cual varía su conductividad dependiendo de la concentración del gas detectable en el aire. Con ayuda de un circuito electrónico como el de la Figura 14, se puede medir la diferencia de tensión entre una resistencia fija  $R_L$  y la resistencia del semiconductor (entre el nodo de  $V_C$  y  $V_{RL}$ ) que varía según la temperatura emitida por el elemento calentador “Heater” y la concentración del gas a detectar en el ambiente (Valdez & Gutiérrez, 2016). En donde:

- $V_C$ : Es el voltaje de alimentación del circuito, en este caso de 5V DC.
- $V_H$ : Es el voltaje del elemento calentador, que puede ser de 5V DC.
- $V_{RL}$ : Es el voltaje presente en la resistencia de carga, o voltaje de salida, el cual depende de  $V_C$ ,  $R_S$  y  $R_L$ .
- $R_L$ : Es la resistencia de carga variable (para la calibración del sistema), de 50 k $\Omega$ .

- $R_S$ : Resistencia del sensor según la concentración de los gases detectables.
- $R_0$ : Resistencia del sensor a 50 ppm de amoníaco (para el sensor TGS826).

### Figura 14

*Circuito de acondicionamiento (Figaro USA Inc, 2002)*



*Nota:* La figura presenta el equivalente circuito para el acondicionamiento de los sensores.

Las características típicas de sensibilidad de los sensores, como es el caso del TGS826 (detector de amoníaco), se expresan según la relación de resistencias  $R_S/R_0$  y la concentración de un determinado gas en el aire en partes por millón (ppm).

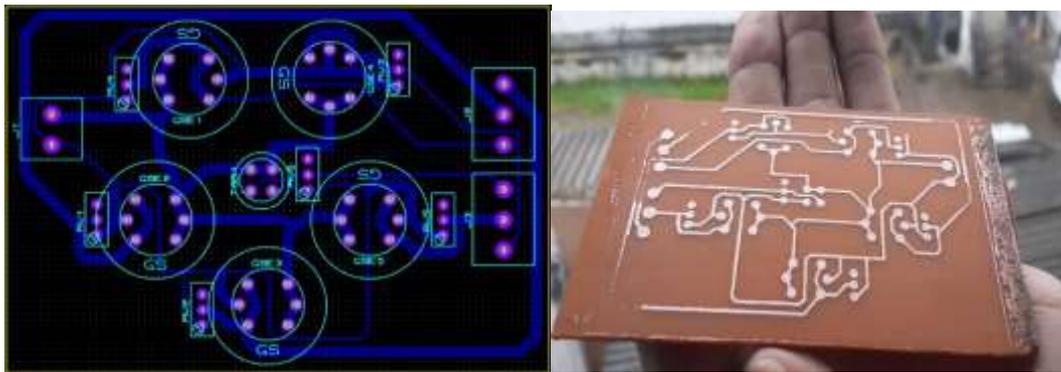
Se puede apreciar en la Figura 14, que el circuito de acondicionamiento consta de un divisor de tensión, por lo que se debe calibrar la resistencia de carga variable  $R_L$  para obtener un voltaje de salida entre 0V a 5V DC.

Al tratarse de un conjunto compuesto por seis sensores, se diseña una placa PCB como se indica en la Figura 15 para satisfacer las conexiones descritas anteriormente en la Figura 14, este diseño toma en cuenta las dimensiones de cada

uno de los sensores, así como sus puntos de energización y salidas hacia la tarjeta de adquisición.

**Figura 15**

*Diseño de placa PCB (Elaboración Propia)*



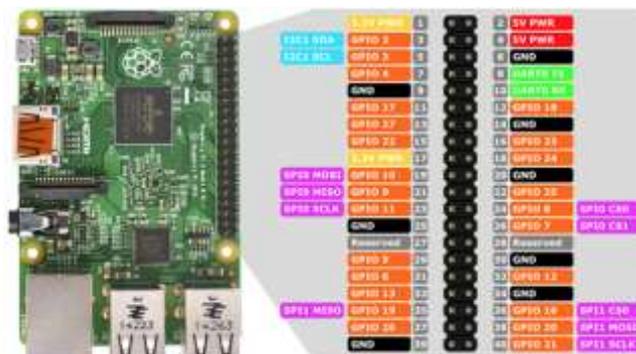
*Nota:* La figura presenta el esquemático de la placa PCB en software y la placa física a ser implementada.

### Adquisición de señales

Como se había mencionado en el Capítulo II, para la adquisición de las señales se empleó el microprocesador Raspberry Pi 3, para esto se hace uso de los pines GPIO de la tarjeta como se puede observar en la Figura 16.

**Figura 16**

*Pines GPIO Raspberry Pi3 (Microsoft Docs, 2017)*



*Nota:* La figura presenta la distribución de los 40 pines GPIO de la tarjeta Raspberry.

Cada pin se encuentra configurado internamente en la tarjeta y debe ser activado mediante la terminal de la Raspberry.

Al no poseer pines de E/S analógicos se emplea un módulo ADC, para la obtención de las 6 señales de los sensores, el cual es el circuito integrado MCP 3008 de 8 canales que envía la trama de datos mediante un protocolo de comunicación SPI, cuya interacción con la *Raspberry Pi3* fue a través del driver localizado en *adafruit-circuitpython-mcp3xxx* disponible para *Python*.

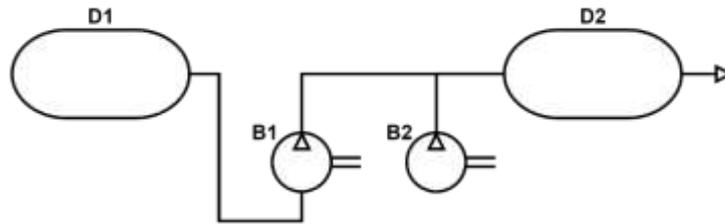
El integrado MCP 3008 tiene la capacidad de transmitir hasta 200ksps (mil muestras por segundo) con una alimentación de 5V DC, lo cual satisfizo la necesidad del muestreo de señales con 1 Hz de frecuencia. Mientras que las señales para la activación y desactivación de las bombas son salidas digitales enviadas a módulos relé que permiten el paso de energía a los motores de las bombas.

### **Diseño del sistema neumático**

El esquema neumático para el prototipo se indica en la Figura 17, el cual consta de dos depósitos: el D1 en donde se encuentran las muestras y el D2 que contiene a la matriz de sensores químicos. Mientras que las bombas se encuentran conectadas para: el bombeo de las sustancias volátiles desde D1 a D2 a través de B1 y para la limpieza con aire del depósito D2 con ayuda de la bomba B2. Cada bomba se alimenta con 12V DC y entrega un flujo de aire sin carga de 6 litros por minuto, que se ajusta con lo requerido por la cámara de sensores D2, que presenta un volumen no mayor a 1.5 litros.

**Figura 17**

*Sistema neumático de accionamiento (Elaboración Propia)*

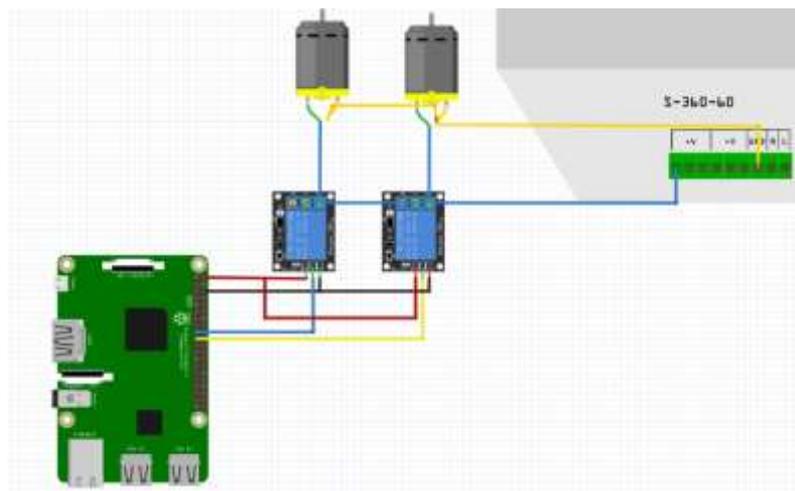


*Nota: La figura presenta el diagrama del funcionamiento de las bombas de aire empleadas.*

El sistema que constituye el accionamiento de las bombas de los motores emplea las salidas digitales de la tarjeta Raspberry las cuales mediante un programa en Python se envía una señal para la activación o desactivación de los módulos relé que permiten la energización de las bombas con 12V DC, el esquema de conexión se presenta en la Figura 18.

**Figura 18**

*Esquema de conexión bombas de aire (Elaboración Propia)*

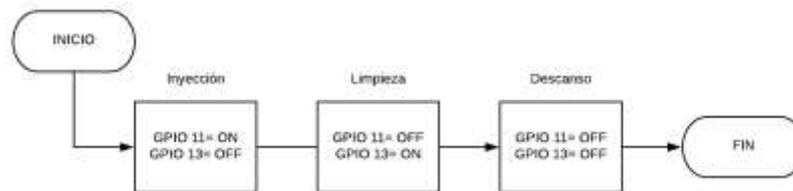


*Nota: La figura muestra el accionamiento de los actuadores correspondientes a cada bomba.*

Una vez determinado los elementos que conforman el sistema se establecen las acciones que debe realizar cada motor durante toda la toma de datos, proceso que se describe en la Figura 19, para esto se emplean las salidas digitales GPIO 11 y GPIO 13 de la tarjeta correspondientes a la bomba de inyección del volátil y la bomba de ingreso de aire limpio respectivamente, estas salidas envían señales a ambos motores para asegurar que una bomba no se encuentre en funcionamiento mientras la otra este activada.

**Figura 19**

*Accionamiento de las bombas (Elaboración Propia)*



*Nota:* La figura muestra ciclo de accionamiento de cada bomba.

### **Calibración del sistema**

Con ayuda de las exportadoras Eco-Kakao S.A. y Agrolaya S.A., se obtuvieron 4 kg. de muestras de cacao en total, ya sean sanos como el fino de aroma o el convencional CCN-51 y una mezcla de cacao con granos contaminados ya sea con moho, falta de fermentación, entre otros.

Según la norma NTE INEN 176 detallada en el Capítulo II, se tomaron 500 gramos de muestras de granos de cacao sanos ya sean fino de aroma o el convencional, y otros 500 gramos de granos contaminados con moho, impurezas y con mala fermentación, con los cuales se realiza un muestreo con una frecuencia de 1Hz utilizando distintas resistencias RL en un rango de 500 ohmios ( $\Omega$ ) a los 50 kilo ohmios ( $k\Omega$ ).

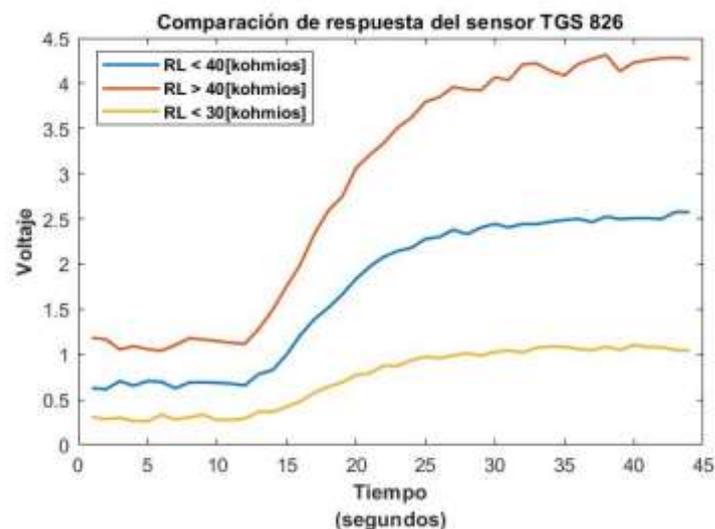
Mientras que, la calibración de la resistencia de carga  $R_L$  se efectúa observando la respuesta de los diferentes sensores, para esto se ha seleccionado al sensor TGS 826 que detecta amoníaco como ejemplo para explicar el proceso de calibración de la resistencia  $R_L$ .

En primer lugar, se fija una resistencia  $R_L$  mayor a los  $40\text{ k}\Omega$  y se comprueba el nivel de voltaje que el sistema entrega luego de estimular al sensor con los olores de los granos de cacao. Se debe tomar en cuenta no saturar los sensores y que el máximo de la señal no se aproxime a los  $5\text{V DC}$ .

Luego de esto se repite el proceso para resistencias  $R_L$  menores a los  $40\text{ k}\Omega$  y menores a los  $30\text{ k}\Omega$ , obteniendo respuestas del sensor como se indica en la Figura 20, donde se puede apreciar que la resistencia entre  $30$  y  $40\text{ k}\Omega$  presenta una mejor respuesta, por lo que se mide  $R_L$  obteniendo un valor de  $39.2\text{ k}\Omega$  y se selecciona para el circuito de acondicionamiento del sensor TGS 826.

## Figura 20

*Curva de respuesta resistencias entre  $30$  y  $40\text{ k}\Omega$  (Elaboración Propia)*

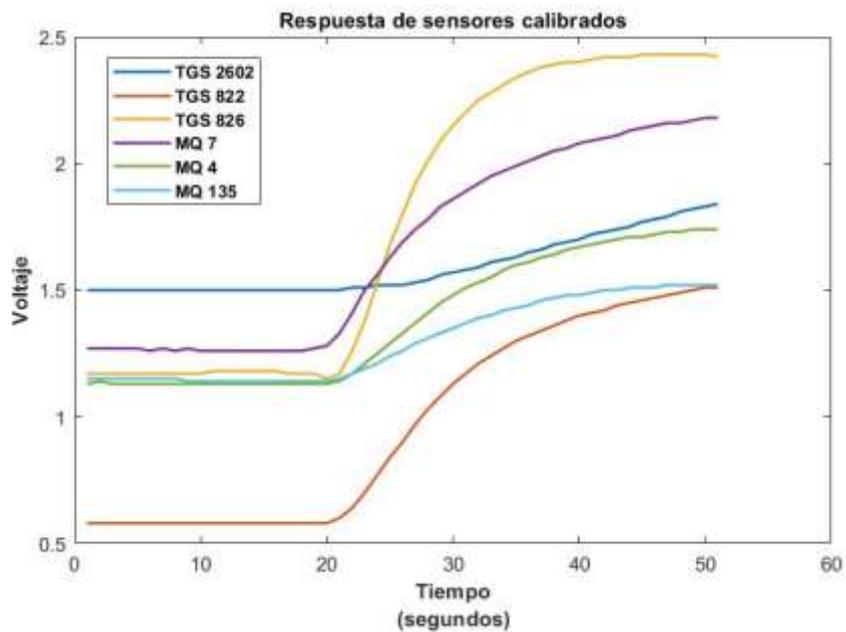


*Nota:* La figura muestra el comportamiento de los sensores frente a diferentes valores de resistencias.

El proceso de calibración se lo realiza en los demás sensores hasta llegar a respuestas como las indicadas en la Figura 21.

**Figura 21**

*Curva de funcionamiento sensores (Elaboración Propia)*



*Nota:* La figura presenta las curvas de funcionamiento con la calibración de resistencia final.

Cuyas resistencias  $R_L$  que se definieron para cada sensor son las que se indican en la Tabla 6.

**Tabla 6**

*Resistencias de calibración (Elaboración Propia.)*

Sensor	TGS 2602	TGS 826	TGS 822	MQ 4	MQ 7	MQ 135
<b>Resistencia de Carga</b>	5.67 [k $\Omega$ ]	39.2 [k $\Omega$ ]	3.17 [k $\Omega$ ]	1.34 [k $\Omega$ ]	10.54 [k $\Omega$ ]	6.18 [k $\Omega$ ]

*Nota:* En la Tabla se muestran los valores de las resistencias de carga para cada sensor implementadas en el circuito de acondicionamiento.

En la Figura 21 también se puede observar que existe un tiempo de subida hasta alcanzar el primer pico y luego una etapa de estabilización producida por la saturación. Es recomendable que los sensores no permanezcan saturados por mucho tiempo. Para evitar el efecto de saturación es necesario definir tiempos de accionamiento de las bombas de limpieza de la cámara de sensores, para la inyección de olores y para la estabilización de los sensores como se indica en la Tabla 7.

**Tabla 7**

*Tiempos de accionamientos (Elaboración Propia)*

<b>Acción</b>	<b>Tiempo</b>
Inyección del volátil	30 seg
Limpieza de la cámara de sensores	20 seg
Recuperación de los sensores	60 seg
Tiempo total del proceso de olfato artificial	110 seg

*Nota:* En la Tabla se muestran los tiempos que se manejan para cada acción, ya sea para inyección del volátil, limpieza de la cámara de sensores y para la recuperación de los sensores.

En donde:

- Inyección del volátil: Es cuando se activa la bomba de aire que transporta las sustancias volátiles de los granos de cacao desde la cámara de olores hasta la cámara de sensores. En donde se midió el tiempo en que los sensores se encontraban cerca del estado de saturación, el cual es de treinta segundos.
- Limpieza de la cámara de sensores: En esta etapa se desactiva la bomba de inyección del volátil y se enciende la bomba de inyección de aire para la limpieza de la cámara de sensores. En este caso se mide el tiempo en que la matriz de sensores detecta

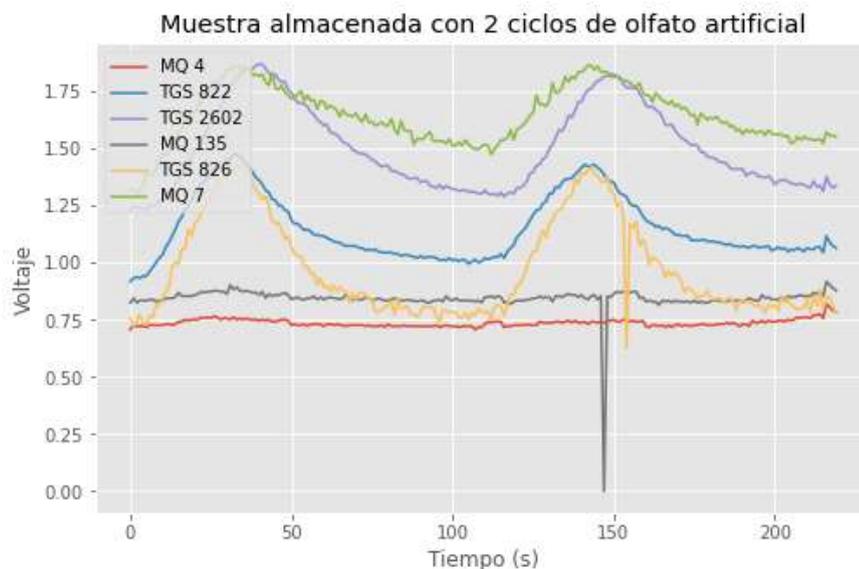
que la cámara se encuentra libre de sustancias volátiles de cacao por la limpieza con aire, con un tiempo establecido de veinte segundos.

- Recuperación de los sensores: Se desactiva la bomba de limpieza y se mide el tiempo en que la respuesta de los sensores se estabiliza en un valor constante. El tiempo establecido es de sesenta segundos.

El proceso de olfato artificial mostrado en la Tabla 7 se lo debe realizar dos veces de forma consecutiva, de tal manera que cada experimento es registrado como se indica en la Figura 22.

### Figura 22

*Proceso de olfato artificial (Elaboración Propia)*



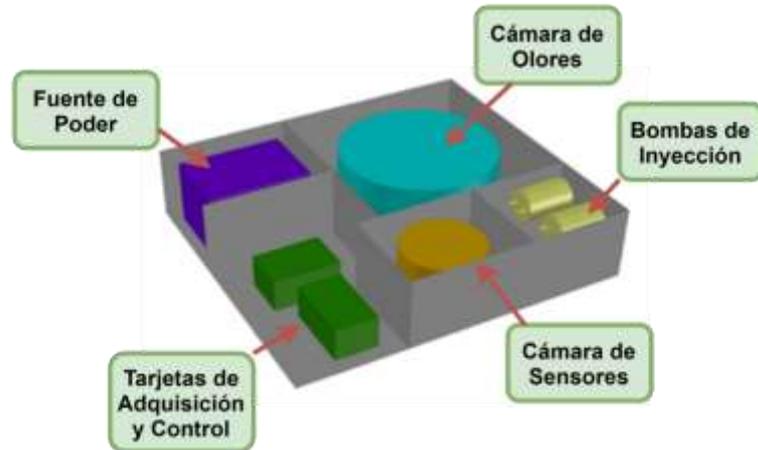
*Nota:* La figura muestra las curvas de comportamiento de cada sensor en la toma de datos de las muestras.

### Diseño funcional y estructural del prototipo

Una vez definido el sistema de adquisición, es necesario distribuir los diferentes elementos del prototipo según su funcionalidad, para lo cual se diseñó un modelo tridimensional con ayuda de AutoCAD el cual se muestra en la Figura 23.

### Figura 23

*Diseño CAD del prototipo (Elaboración Propia.)*



*Nota:* La figura muestra una propuesta inicial del prototipo.

El diseño propuesto, en la Figura 23, se divide en:

- **Cámara de olores:** Es el recipiente en donde se colocan las muestras de cacao para extraer el olor, como se aprecia en la Figura 24.

### Figura 24

*Cámara de almacenamiento de las muestras. (Elaboración Propia.)*



*Nota:* La figura muestra el recipiente empleado para almacenar las muestras de cacao.

- **Cámara de sensores:** Se conforma de un recipiente metálico que contiene la matriz de sensores, que se encuentra conectado con la cámara de olores mediante mangueras que transportan el volátil de una cámara hacia otra. En la Figura 25 se puede observar el exterior del recipiente con las resistencias  $R_L$  para cada sensor y también se aprecia el interior del recipiente con la matriz de sensores.

**Figura 25**

*Cámara de sensores (Elaboración Propia)*

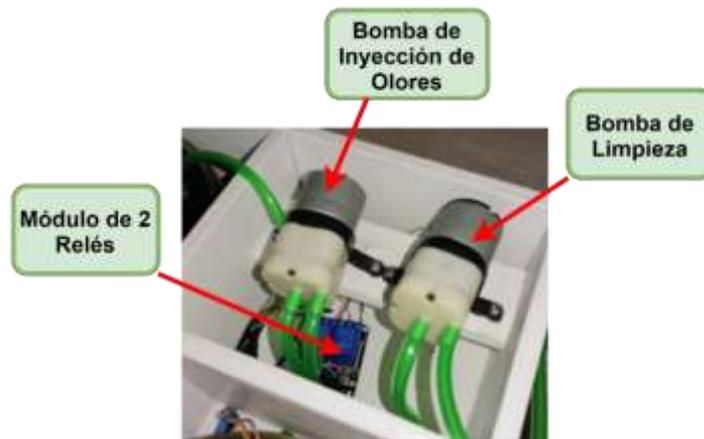


*Nota:* La figura muestra el acople de los 6 sensores al recipiente para formar la cámara de sensores.

- **Bombas de inyección:** Sección que está constituida por 2 bombas de aire, la primera transporta el volátil concentrado de la cámara de olores a la cámara de sensores, la segunda se emplea para enviar aire directamente a la cámara de sensores como acción de limpieza. El accionamiento de las bombas se encuentra controlado por la tarjeta microcontroladora Raspberry Pi3 y por un elemento pre accionador que es un módulo de 2 relés para la alimentación de las bombas con 12V DC. La implementación se puede apreciar en la Figura 26.

**Figura 26**

*Bombas de flujo de aire e inyección de olores (Elaboración Propia.)*

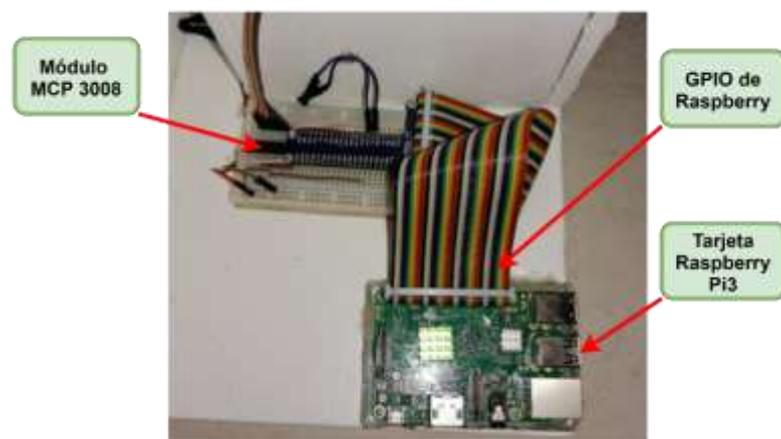


*Nota:* La figura muestra los elementos que constituyen la sección de bombas de aire.

- **Tarjeta de adquisición y control:** En la Figura 27 se presenta a la tarjeta Raspberry Pi3 y los módulos que se acoplan a la tarjeta para realizar la adquisición y procesamiento de señales.

**Figura 27**

*Sección de adquisición y control de datos. (Elaboración Propia.)*



*Nota:* La figura muestra a la tarjeta de adquisición y los módulos empleados para la adquisición de señales.

- **Fuente de Poder:** Durante el arranque del sistema de la Raspberry Pi3 se presentaba un aviso en pantalla sobre Bajo nivel de Voltaje en la Fuente de Alimentación, y al emplear diversos niveles de energía para el accionamiento de los distintos elementos que conforma el prototipo en su totalidad se utiliza una fuente de alimentación multi voltaje que se enciende de forma manual a través de un Switch manual. El sistema de la Fuente de Poder se indica en la Figura 28.

### Figura 28

*Suministro de energía del prototipo. (Elaboración Propia.)*



*Nota:* La figura muestra una fuente de energía de computador empleada para distribuir energía a todo el prototipo.

A continuación, se puede apreciar el prototipo propuesto, el cual se muestra en la Figura 29.

**Figura 29**

*Vista general del prototipo implementado (Elaboración Propia)*



*Nota:* La figura muestra una fotografía de todos los componentes que conforman el prototipo de nariz electrónica implementado.

**Creación de Base de Datos**

En la Figura 30 se muestran los 2 grupos de muestras utilizadas para la creación de la base de datos, el primer grupo son muestras de cacao sano y el segundo de cacao contaminado.

**Figura 30**

*Grupos de grano de cacao (Elaboración Propia)*



*Nota:* En la figura se muestra los diferentes tipos de granos de cacao a utilizar para la creación de la base de datos, a la izquierda el cacao sano y a la derecha el contaminado.

Este sistema de muestreo se llevó a cabo en varias etapas, que garanticen una correcta lectura por parte de los 6 sensores.

- **Concentración de olor de las muestras:** Se aíslan 100 granos de cacao en el recipiente hermético de metal en el cual se concentra el aroma por un periodo de media hora.
- **Extracción del Aroma:** Mediante el uso de las bombas se extrae el volátil concentrado y es trasladado a la cámara de olores, la cual es un recipiente hermético de metal, este proceso se lleva a cabo durante 30 segundos.
- **Limpieza de los sensores:** Una vez se captura el aroma se deja circular aire limpio en la cámara multisensorial durante 20 segundos, de este modo se asegura que no exista restos de aroma de experimentos realizados con anterioridad.
- **Recuperación de los sensores:** Durante 60 segundos se deja al sistema en reposo para que los sensores vuelvan a su estado inicial y se puede realizar una nueva medición.
- **Almacenamiento de los datos:** Cada sensor envía una trama de datos que es almacenada en un archivo en formato txt mediante un programa desarrollado en Python, de este modo se obtiene un conjunto de 1320 datos por experimento.

En conjunto la base de datos general está conformada por una matriz de 110 experimentos. Cada experimento tiene 1320 datos y 6 sensores.

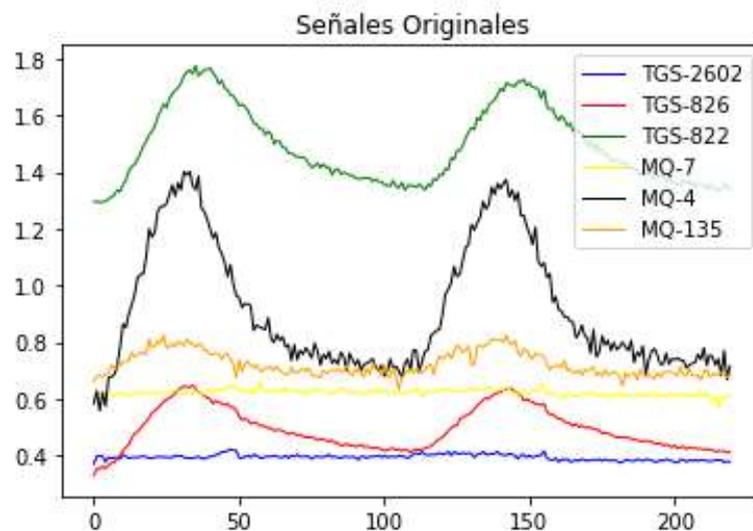
### **Preprocesado de señales**

El preprocesamiento de los datos es uno de los pasos cruciales en cualquier aplicación de aprendizaje automático, debido a que los datos en bruto pueden contener perturbaciones como el ruido o presentar valores atípicos provocados por situaciones externas al prototipo.

Para el entrenamiento del modelo de red neuronal es necesario identificar y eliminar aquellos valores atípicos ya que pueden alterar los resultados de la red. Es por ello, que se elaboró un algoritmo en Python capaz de realizar el preprocesamiento de los datos obtenidos con anterioridad, dividido en las siguientes etapas: filtrado, corrección de línea base, alineación de picos, concatenación de sensores y eliminación de outliers. En la Figura 31 se puede apreciar un ejemplo de señales en bruto.

### Figura 31

*Señales en bruto de los sensores (Elaboración Propia)*



*Nota:* La figura muestra las señales obtenidas sin ningún tipo de procesado.

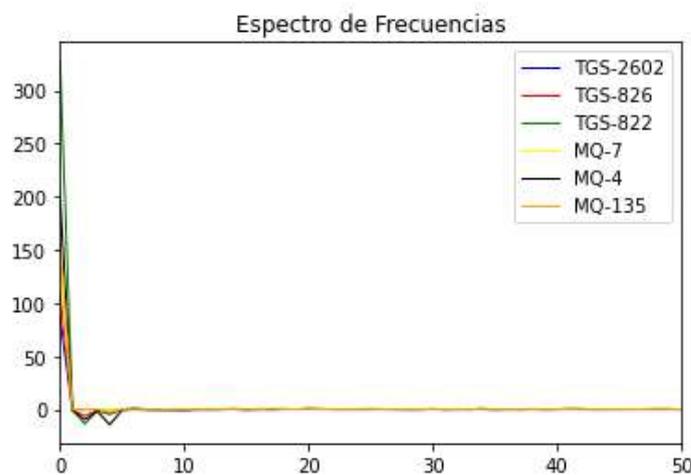
### Filtrado

Debido a las diferentes etapas en los experimentos realizados puede existir perturbaciones provenientes de señales externas o señales generadas por el mismo sistema que influyen en la toma de datos. Para determinar el tipo de filtro a utilizar se realizó un análisis espectral con el fin de determinar en qué sector del rango de frecuencias se encuentran las potencias de las señales, eliminando el ruido y perturbaciones que alteran las características de la señal.

Para el uso de las herramientas para el análisis de señales en Python se emplea la librería *numpy*, esta realiza todos los cálculos matemáticos necesarios para esta etapa, de este modo mediante el cálculo de la transformada rápida de Fourier (FFT) se obtuvo los datos de las muestras en el dominio de la frecuencia, a partir de este nuevo conjunto de datos se obtiene el espectro de potencias permitiendo determinar las frecuencias asociadas a los componentes de las señales.

**Figura 32**

*Espectro de frecuencias (Elaboración Propia)*



*Nota:* La figura muestra el comportamiento de la señal en el dominio de la frecuencia.

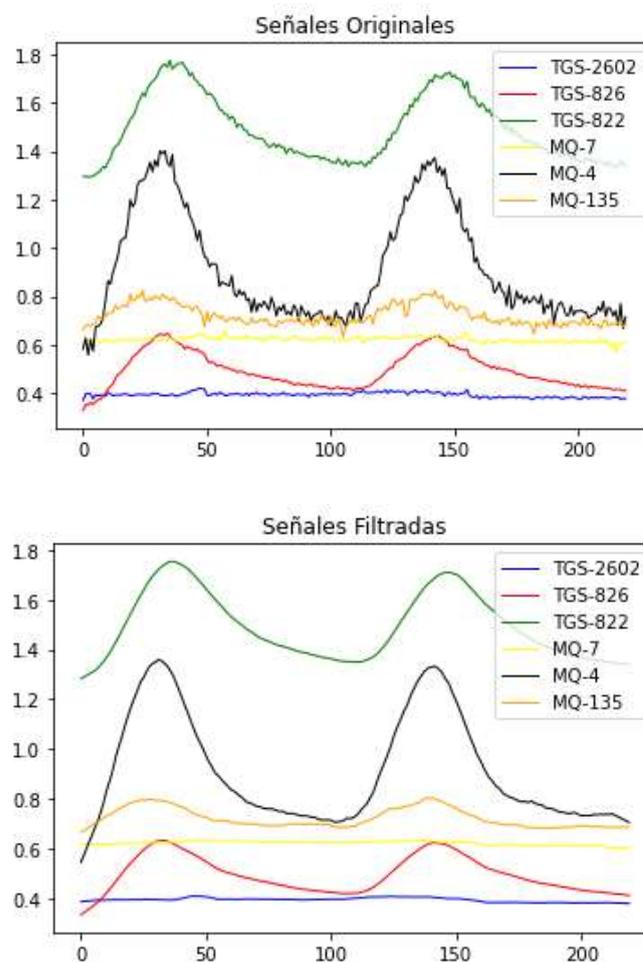
Al analizar el espectro de frecuencias se puede visualizar que la potencia de las señales se encuentra contenida en las bajas frecuencias como se muestra en la Figura 32, por esta razón se es necesario un filtro que suavice y conserve las características de la señal, en base al trabajo desarrollado en (De la Cruz, 2020), se opta por un filtro Savitzky–Golay el cual presenta un rendimiento óptimo para este tipo de experimento, ya que realiza un suavizado de la señal sin perder su forma original.

El filtrado realiza una aproximación polinómica en Python, a través del ajuste de un polinomio a un grupo de datos mediante la función `savgol_filter`, dicha función

tiene como parámetros el grado del polinomio de ajuste y la longitud de la ventana del filtro, para el desarrollo de este trabajo se tomó como grado del polinomio  $N=1$ , el tamaño de la ventana  $M=10$  con el fin de obtener un conjunto de datos representativos. Figura 33, presenta la comparación entre la señal original y la señal posterior a la aplicación del filtro.

**Figura 33**

*Rendimiento del filtro aplicado (Elaboración propia.)*



*Nota:* La figura muestra una comparativa antes y después de la aplicación del filtro.

### Corrección de Línea Base

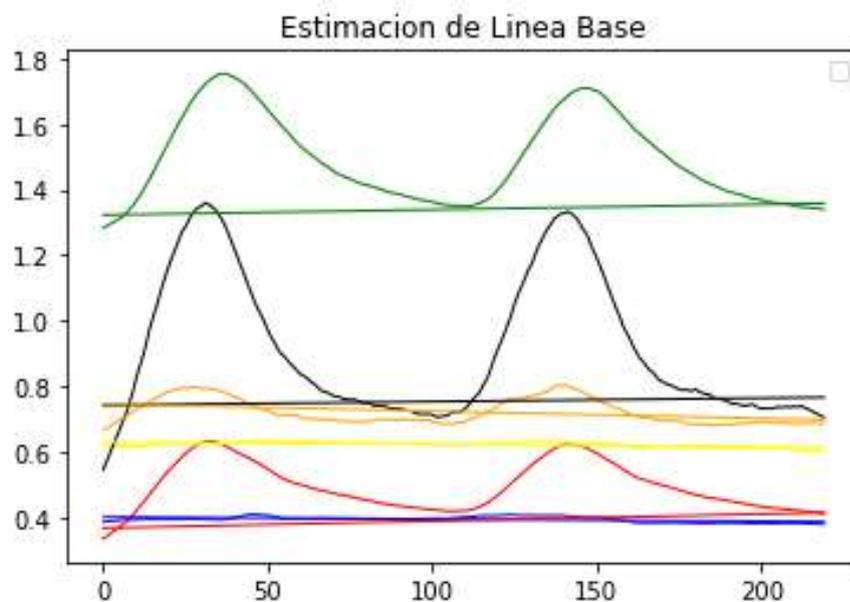
Dentro de la calibración de los sensores se observó que presenta una línea base, por esta razón es necesario una corrección para mejorar el desempeño de la red neuronal.

Para la compensación de esta línea base se realiza la resta de la señal original con un polinomio de estimación de la línea base, dicha estimación se encuentra en los puntos mínimos de cada señal, para el desarrollo de este trabajo se optó por un polinomio de grado 1 para que la línea base de todas las señales se ajuste a la línea en el origen.

Para la obtención de los coeficientes de la ecuación se empleó la función *polyfit* la cual realiza un ajuste del polinomio en los puntos mínimos de la señal, obteniendo una recta que minimice el error.

### Figura 34

*Línea base estimada por sensor. (Elaboración Propia)*

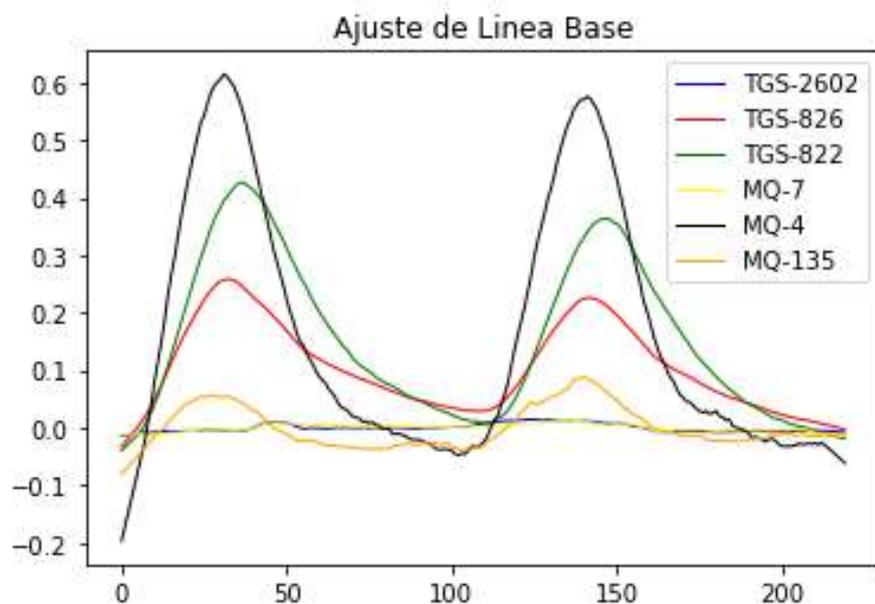


*Nota:* La figura presenta cada señal con su respectiva línea base estimada.

Los puntos a los cual debe ajustarse el polinio se obtuvieron mediante la función *np.amin*, dicha función identifica los valores mínimos en un conjunto de datos, con estos se calcula la línea base estimada como se observa en la Figura 34 y se resta de la señal original, teniendo como resultado todas las señales ajustadas en el origen, como se observa en la Figura 35.

### Figura 35

*Señales ajustadas al origen. (Elaboración Propia)*



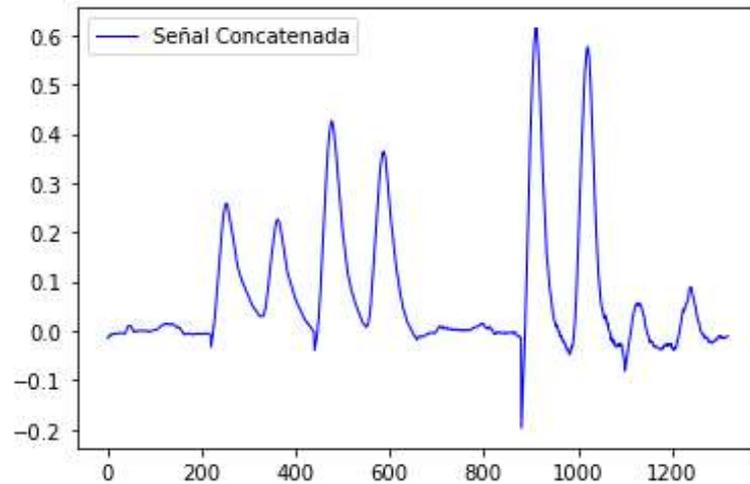
*Nota:* La figura muestra la alineación de la línea base en el origen para todas las señales obtenidas.

### Concatenación de Sensores

Para el entrenamiento de la red neuronal es necesario que el conjunto de datos se encuentre agrupado por experimento, por esta razón se unen las respuestas de cada señal en una sola trama de datos por cada experimento como se muestra en la Figura 36.

### Figura 36

Señales concatenadas de los 6 sensores. (Elaboración Propia)



*Nota:* En la Figura se observa un experimento con la concatenación de las señales entregadas por cada sensor.

### Alineación de Picos

Dentro de la etapa de extracción del aroma de las muestras el tiempo de reacción de cada sensor puede variar entre cada experimento, lo cual provoca que los valores pico de las señales varíe en relación con el tiempo, por lo tanto, se debe alinear los picos de los experimentos realizados con el fin de que los puntos de interés se encuentren ubicados en el mismo sector.

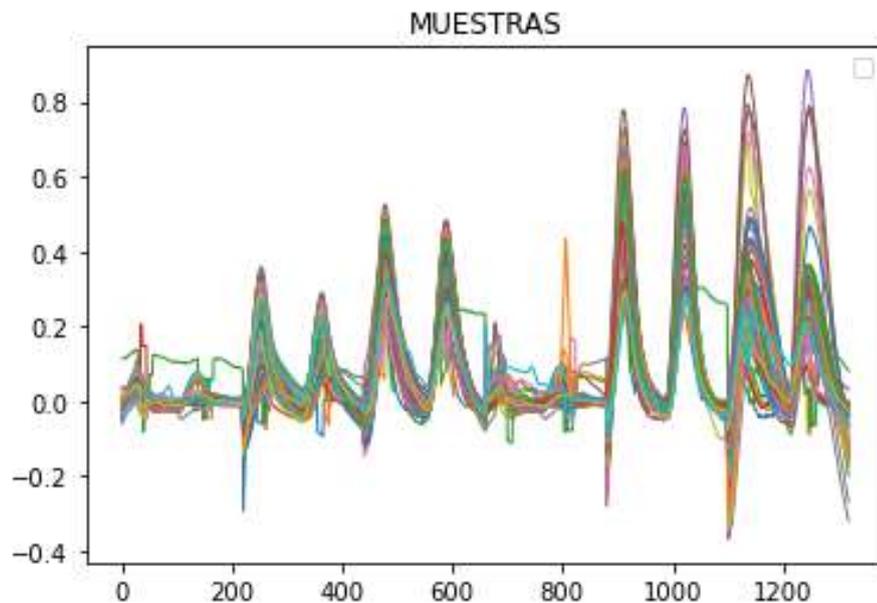
Para alinear los picos de las señales se identifican los puntos máximos de las señales y se calcula el promedio de los puntos máximos. De este modo se establece un punto para alinear el resto de las señales, si las señales se encontraban desfasadas con respecto al punto de alineación se realiza un desplazamiento a la izquierda o derecha dependiendo el desfase de la señal.

El desplazamiento se realiza mediante la función *np.roll*, la cual desplaza todo el conjunto de datos, de este modo el primer valor se convierte en el último y viceversa

cuando se rota al lado contrario. De esta manera se obtienen los resultados indicados en la Figura 37 para el conjunto de experimentos.

### Figura 37

*Resultado de la alineación de picos en las señales procesadas. (Elaboracion Propia)*



*Nota:* En la Figura se puede apreciar que todos los experimentos se encuentran alineados respecto a los picos de las señales.

### Detección y Eliminación de Outliers

Para definir un modelo de red neuronal se necesita un conjunto de datos que en su mayoría sean homogéneos ya que una señal fuera de rango puede afectar al entrenamiento y causar errores en la clasificación, para lo cual es importante determinar aquellas señales que presenten desviaciones al comportamiento normal.

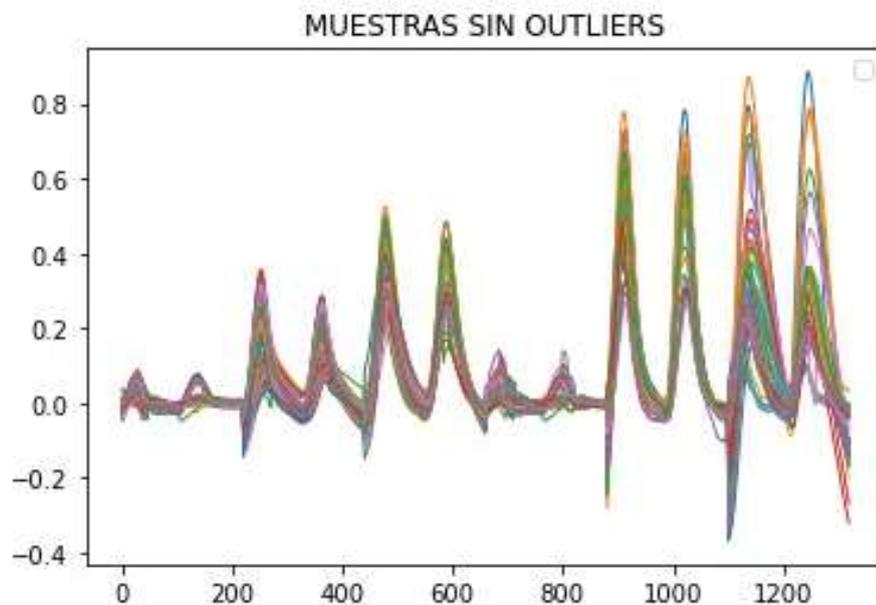
La presencia de outliers en las señales se debe a perturbaciones del sistema de adquisición ocasionadas por el ruido del ambiente, vibraciones o en el caso del integrado MCP 3008, a variaciones en el voltaje de alimentación.

Por consiguiente, se procedió a analizar por inspección visual a cada una de las señales con respecto al conjunto de datos, verificando que cada uno de los experimentos realizados no presente saturación de sensores ni perturbaciones que alteraban la forma de las señales.

Las señales que presentaron anomalías se retiraban del conjunto de datos obteniendo una base de datos definitiva tanto para el entrenamiento como para la validación, en la Figura 38 se representa el nuevo conjunto de datos de los cuales en la Tabla 8 se muestra el número de experimentos antes y después de la eliminación de outliers.

### Figura 38

*Conjunto de datos final.* (Elaboración Propia)



*Nota:* En la Figura se aprecia el conjunto de experimentos sin presencia de Outliers que puedan alterar el entrenamiento de la red neuronal.

**Tabla 8**

*Numero de datos antes y después de la eliminación de Outliers. (Elaboración Propia)*

	Experimentos Iniciales	Outliers	Experimentos Finales
<b>Experimentos de entrenamiento</b>	75	7	68
<b>Experimentos de validación</b>	35	5	30
<b>Experimentos Totales</b>	110	12	98

*Nota:* En la Tabla se muestran los experimentos analizados inicialmente, el número de experimentos donde se detectaron Outliers y los experimentos finales que ya no presentan valores atípicos.

### **Modelo de la Red Neuronal**

En el desarrollo del modelo de clasificación primero fue necesaria una división de los datos pre-procesados anteriormente en dos grupos, un grupo del 70% dedicados al entrenamiento y el 30% restante para la evaluación del modelo. Para esta separación se utilizó la función *train\_test\_split* perteneciente a la librería *sklearn* (Scikit-learn) de Python, la cual además de realizar la distribución de los datos de entrenamiento-evaluación, permite distribuirlos de forma aleatoria evitando riesgos de overfitting o underfitting (Aggarwal, 2018).

Mientras que, para construir la forma del modelo se utiliza una API de alto nivel denominada *Keras* de *Tensorflow*, para luego agrupar y entrenar el modelo que consta de una capa de entradas y una capa para la salida, a través de la clase *Sequential*, la cual permite definir cuántas neuronas, entradas, qué función de activación, función de pérdida, mecanismo de optimización y métricas se desean utilizar (Ketkar, 2017).

Por lo tanto, es necesario identificar qué funciones de pérdida, mecanismos de optimización y qué métricas de evaluación están disponibles en *Keras* para

implementar en el modelo de clasificación (Chollet, 2018). Entre las *funciones de pérdida* comúnmente utilizadas están:

- **Error cuadrático medio (*tf.keras.losses.MeanSquaredError*):** Calcula el error cuadrático medio entre los valores de salida reales y los predichos por la red neuronal. Se obtiene como se indica en la ecuación (5).

$$ECM = \frac{1}{n} \sum_{i=1}^n (\hat{Y}_i - Y_i)^2 \quad (5)$$

En donde  $\hat{Y}_i$  es cada valor predicho obtenido,  $Y_i$  es cada valor real y  $n$  es el número total de datos.

- **Error absoluto medio (*tf.keras.losses.MeanAbsoluteError*):** Calcula el error absoluto medio entre los valores de salida reales y los predichos por la red neuronal. Se calcula utilizando la ecuación (6).

$$EAM = \frac{1}{n} \sum_{i=1}^n |\hat{Y}_i - Y_i| \quad (6)$$

En donde  $\hat{Y}_i$  es cada valor predicho obtenido,  $Y_i$  es cada valor real y  $n$  es el número total de datos.

Mientras que en los *mecanismos de optimización* más utilizados se encuentran:

- **RMSprop (*tf.keras.optimizers.RMSprop*):** A lo largo del entrenamiento mantiene una media móvil de los gradientes al cuadrado para cada peso de la red neuronal, y luego divide el gradiente para la raíz de esta media.
- **Adagrad (*tf.keras.optimizers.Adadelta*):** Se parece a RMSprop, ya que también es una tasa de aprendizaje adaptativo, aunque mantiene una tasa de aprendizaje por cada parámetro basada en una suma acumulada de gradientes cuadrados. Para actualizar la tasa de aprendizaje, se divide únicamente para dicha suma acumulada.

- **Adam (*tf.keras.optimizers.Adam*):** Este tipo de optimización es un método de descenso de gradiente estocástico, el cual combina los métodos RMSprop y Adagrad. Mantiene fija su tasa de *aprendizaje* durante las adaptaciones de los pesos.

Las métricas de evaluación para modelos de clasificación más utilizadas se basan de una Matriz de Confusión (*sklearn.metrics.confusion\_matrix*) como se indica en la Tabla 9, la cual permite interpretar el desempeño de un modelo a través del conjunto de datos para *test*, en donde los valores de salida son conocidos (Ketkar, 2017).

**Tabla 9**

*Matriz de Confusión. (Ketkar, 2017)*

		Predicho	
		Positivo	Negativo
Actual	Positivo	Verdaderos Positivos	Falsos Negativos
	Negativo	Falsos Positivos	Verdaderos Negativos

*Nota:* En la Tabla se aprecia la estructura de una matriz de confusión.

Dentro de la Matriz de confusión se pueden presentar 4 distintas situaciones:

- Verdaderos Positivos (TP): Es cuando el valor predicho (de salida) por la red neuronal y el valor real o actual de la base de datos coinciden, presentando ambos su estado en Alto.
- Falsos Negativos (FN): Es cuando el valor predicho está en estado Bajo y el real se encuentra en Alto.
- Falsos Positivos (FP): Es cuando el valor predicho está en estado Alto y el real se encuentra en Bajo.

- Verdaderos Negativos (TN): Es cuando el valor predicho y el valor real coinciden, presentando ambos su estado en Bajo.

Con ayuda de la Matriz de Confusión se pueden obtener las *métricas de evaluación* que también están disponibles en la librería *sklearn* (Brownlee, 2017):

- **Exactitud (*sklearn.metrics.accuracy\_score*):** Mide cuántas predicciones se han realizado correctamente, y se calcula con la ecuación (7).

$$Exactitud = \frac{TP + TN}{TP + FN + FP + TN} \quad (7)$$

- **Precisión (*sklearn.metrics.precision\_score*):** Mide cuán preciso es el clasificador para predicciones positivas, y se calcula con la ecuación (8).

$$Precisión = \frac{TP}{TP + FP} \quad (8)$$

- **Sensibilidad (*sklearn.metrics.recall\_score*):** Mide cuán sensible es el clasificador en la detección de valores reales positivos, y se calcula con la ecuación (9).

$$Sensibilidad = \frac{TP}{TP + FN} \quad (9)$$

- **Puntaje de F1 (*sklearn.metrics.f1\_score*):** Permite representar a la precisión y sensibilidad del modelo en un solo valor, y se calcula con la ecuación (10).

$$Puntaje F1 = \frac{2 \cdot Precisión \cdot Sensibilidad}{Precisión + Sensibilidad} \quad (10)$$

Una vez definidas las métricas y todas las partes necesarias en un modelo de red neuronal, se puede definir su estructura, calcular sus pesos óptimos mediante entrenamiento y observar su desempeño, todo esto se puede apreciar en el Capítulo IV.

## Capítulo IV

### Validación, Pruebas y Resultados

En este capítulo se presentan las métricas obtenidas durante el entrenamiento y la evaluación del modelo a implementar y del funcionamiento del prototipo. Para lo cual se tabularon los resultados luego de haber modificado las funciones de activación en la capa de entrada de la red neuronal a ReLU, Función Sigmoide y Función Signo. Se incluye una métrica importante que es el tiempo de entrenamiento o tiempo en el que el ordenador tarda en procesar todas las épocas definidas. Por último, con el modelo de red neuronal entrenado e implementado, se creó una nueva base de datos con veinte experimentos para calcular métricas de evaluación del funcionamiento del prototipo utilizando validación ciega y se registran los resultados, en donde se obtiene las métricas de exactitud, precisión, sensibilidad, puntaje F1 y tiempo requerido por el ordenador en el entrenamiento.

### Métricas de evaluación del modelo de red neuronal

Es necesario detallar el ordenador que se utilizó, el cual tiene un procesador Intel Core(TM) i7-10750H de 6 núcleos y 12 hilos con una frecuencia base de 2.6GHz. Además, las muestras de granos de cacao que se utilizan a continuación pertenecen a la base de datos dividida en 70% para entrenamiento y 30% para *test*. Se implementan 1320 entradas hacia la red neuronal para cada muestra obtenida en los experimentos, y una sola salida que indica la predicción (sano o contaminado) sobre la calidad del cacao medido.

Se utiliza en primer lugar el subconjunto de datos de entrenamiento que contiene 68 experimentos y para comenzar con el entrenamiento, se deben definir en *Python* cuántas neuronas tendrá la red neuronal y el número de épocas utilizando *tf.keras.models.Model.fit*. La función de pérdida a utilizar será del *Error Cuadrático Medio*, junto con el mecanismo de optimización *Adam*.

Luego se entrena la red neuronal utilizando dos neuronas en la capa de entrada y una neurona en la capa de salida, cuyas métricas del modelo entrenado son calculadas según las predicciones obtenidas con el subconjunto de datos para *test*, que contiene treinta experimentos. En donde se escogen los tres mejores modelos entrenados, se calcula una media de las métricas y se registran en la Tabla 10.

**Tabla 10**

*Entrenamiento del modelo de red neuronal con 2 neuronas en la capa de entrada.*

(Elaboracion Propia)

<b>Nº de épocas</b>	<b>Función de Activación capa de entrada</b>	<b>Exactitud (%)</b>	<b>Precisión (%)</b>	<b>Sensibilidad (%)</b>	<b>Puntaje F1 (%)</b>	<b>Tiempo de entrenamiento (ms)</b>
<b>25</b>	ReLU	88.89	78.57	96.30	86.06	425.77
<b>50</b>	ReLU	95.56	90.48	100	94.44	525.24
<b>100</b>	ReLU	100	100	100	100	878.20
<b>25</b>	Sigmoide	76.67	78.57	80.12	72.49	405.71
<b>50</b>	Sigmoide	93.33	100	88.43	93.61	516.35
<b>100</b>	Sigmoide	95.56	92.86	90.63	94.85	820.99
<b>25</b>	Signo	68.89	33.33	100	47.39	435.42
<b>50</b>	Signo	88.89	76.19	100	86.44	577.68
<b>100</b>	Signo	96.66	92.85	100	96.20	941.58

*Nota:* Se puede observar en la Tabla que se entrena la red neuronal utilizando distintas funciones de activación en la capa de entrada y se registra la media de las métricas obtenidas.

En la Tabla 10 se puede apreciar que al implementar una función de activación ReLU en la capa de entrada, se requiere un tiempo en el entrenamiento de la red de 878 milisegundos y todas las métricas tienen un valor de 100% utilizando cien épocas. Por otro lado, las métricas obtenidas en las funciones de activación como función Sigmoide o función Signo, no alcanzan el 100% al utilizar las cien épocas y el tiempo

de entrenamiento en el caso de la función Signo es de 941 milisegundos, siendo más costoso computacionalmente a diferencia de utilizar una función ReLU.

Mientras que, al utilizar tres neuronas en la capa de entrada se obtuvieron las métricas detalladas en la Tabla 11 y al utilizar una función de activación ReLU para la capa de entrada, el sistema nuevamente presenta métricas de evaluación con valores del 100%.

**Tabla 11**

*Entrenamiento del modelo de red neuronal con 3 neuronas en la capa de entrada.*

*(Elaboracion Propia)*

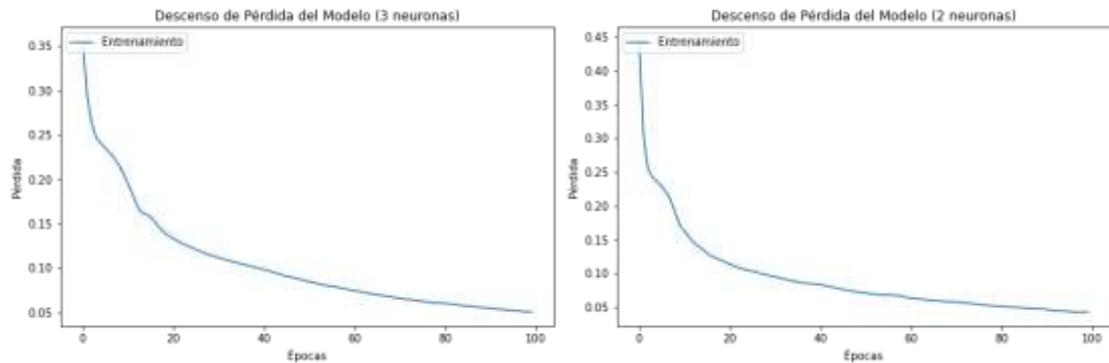
<b>Nº de épocas</b>	<b>Función de Activación capa de entrada</b>	<b>Exactitud (%)</b>	<b>Precisión (%)</b>	<b>Sensibilidad (%)</b>	<b>Puntaje F1 (%)</b>	<b>Tiempo de entrenamiento (ms)</b>
<b>25</b>	ReLU	88.89	76.19	100	85.35	450.36
<b>50</b>	ReLU	98.89	97.62	100	98.77	515.44
<b>100</b>	ReLU	100	100	100	100	616.49
<b>25</b>	Sigmoide	77.77	57.14	95.83	61.64	331.67
<b>50</b>	Sigmoide	86.66	71.42	100	75.00	410.24
<b>100</b>	Sigmoide	98.89	100	97.78	98.85	614.36
<b>25</b>	Signo	86.67	71.42	100	83.20	358.28
<b>50</b>	Signo	88.88	76.19	100	86.19	450.25
<b>100</b>	Signo	93.33	85.71	100	90.91	627.84

*Nota:* La tabla presenta valores obtenidos en el entrenamiento con 3 neuronas en la capa de entrada

Por lo que se escoge la función ReLU para la activación de la capa de entrada, mientras que el número de neuronas a utilizar se lo selecciona comparando el mejor desempeño presente en el descenso de la función de pérdida durante el proceso de entrenamiento, con un mismo número de cien épocas.

**Figura 39**

*Descenso del error cuadrático medio durante el entrenamiento del modelo con dos y tres neuronas en la capa de entrada. (Elaboración Propia)*



*Nota:* En la Figura se puede observar el descenso de la función de pérdida durante el entrenamiento del modelo. A la izquierda se observa el descenso cuando la capa de entrada contiene tres neuronas y a la derecha cuando contiene dos neuronas.

En la Figura 39 se puede apreciar que al utilizar dos o tres neuronas se encuentra un descenso de función de pérdida similar durante el entrenamiento, con la diferencia en que al manejar dos neuronas se necesitaron 597.40 milisegundos hasta llegar a un ECM (error cuadrático medio) de 0.043, mientras que con tres neuronas fueron necesarios 857.45 milisegundos hasta obtener un ECM de 0.05, por lo que se escoge utilizar dos neuronas en la capa de entrada.

### **Pruebas de desempeño funcional del prototipo**

El modelo de red neuronal debe incorporar los parámetros de la Tabla 12, los cuales ya fueron definidos en el apartado anterior.

**Tabla 12**

*Parámetros para el modelo de red neuronal a implementar. (Elaboracion Propia)*

<b>Nº de épocas</b>	<b>Nº de neuronas capa de entrada</b>	<b>Nº de neuronas capa de salida</b>	<b>Función de activación capa de entrada</b>	<b>Función de activación capa de salida</b>	<b>Función de pérdida</b>	<b>Mecanismo de optimización</b>
100	2	1	ReLU	Sigmoide	Error Cuadrático Medio (ECM)	Adam

*Nota:* La tabla muestra los parámetros con los cuales se programó la red neuronal.

A continuación, se realiza una validación ciega en donde se efectuaron veinte nuevas mediciones con el prototipo, de las cuales diez corresponden al cacao sano y las diez restantes al contaminado. El proceso de medición se efectuó en seis pasos.

- Se colocaron veinte muestras de cacao con 100 granos cada una dentro de la cámara de olores, ya sean sanos o contaminados. Donde reposan por un tiempo de 5 minutos para que se concentre el aroma dentro de la cámara como se observa en la Figura 40.

**Figura 40**

*Granos de cacao en la cámara de olores. (Elaboracion Propia)*



*Nota:* La figura muestra una la colocación de las muestras de granos de cacao en la cámara de olores para la obtención de las señales.

- En la Figura 41 se puede observar que se conectó la *Raspberry Pi3* a su fuente de alimentación y se encendió la Fuente de Poder Multi-Voltaje. Luego se debe esperar un tiempo de diez minutos para que el sistema se encienda y que se calienten los sensores químicos.

### Figura 41

*Conexión de prototipo a las fuentes de alimentación. (Elaboracion Propia)*



*Nota:* La figura muestra la energización del sistema a través de la fuente de poder multi-voltaje.

- Dentro de la interfaz proyectada de la tarjeta *Raspberry Pi3*, se ejecuta el código para la adquisición de señales desarrollado en *Python* y se presenta un menú en la terminal como el indicado en la Figura 42, en donde se selecciona la opción de *Encendido* con ayuda del teclado para comenzar el proceso de olfato de la muestra de cacao.

## Figura 42

*Menú de selección para comenzar el proceso de olfato. (Elaboracion Propia)*

```
Shell x
Python 3.7.3 (/usr/bin/python3)
>>> %Run sensores6.py

[1] Encendido
[0] Salir
Ingresar Opcion:1
```

*Nota:* La figura muestra las acciones a realizar dentro del proceso de adquisición de señales.

- En la Figura 43 se puede observar que durante la ejecución del programa se imprime en la terminal de *Python* los niveles de voltaje que entrega la matriz de seis sensores químicos con una actualización de 1Hz, que es la frecuencia de muestreo de las señales.

## Figura 43

*Niveles de voltaje de los sensores químicos durante el funcionamiento del prototipo. (Elaboracion Propia)*

```
Shell x
0.301, 0.301, 1.003, 0.815, 0.999, 0.783
0.364, 0.361, 1.650, 0.812, 0.902, 0.770
0.371, 0.354, 1.669, 0.815, 0.951, 0.764
0.364, 0.364, 1.647, 0.796, 0.941, 0.748
0.367, 0.364, 1.650, 0.793, 0.947, 0.783
0.364, 0.361, 1.650, 0.809, 0.954, 0.786
0.361, 0.364, 1.650, 0.815, 0.935, 0.761
0.361, 0.354, 1.627, 0.793, 0.938, 0.767
0.374, 0.364, 1.653, 0.799, 0.951, 0.770
0.361, 0.351, 1.650, 0.806, 0.960, 0.732
0.361, 0.371, 1.673, 0.806, 0.967, 0.783
```

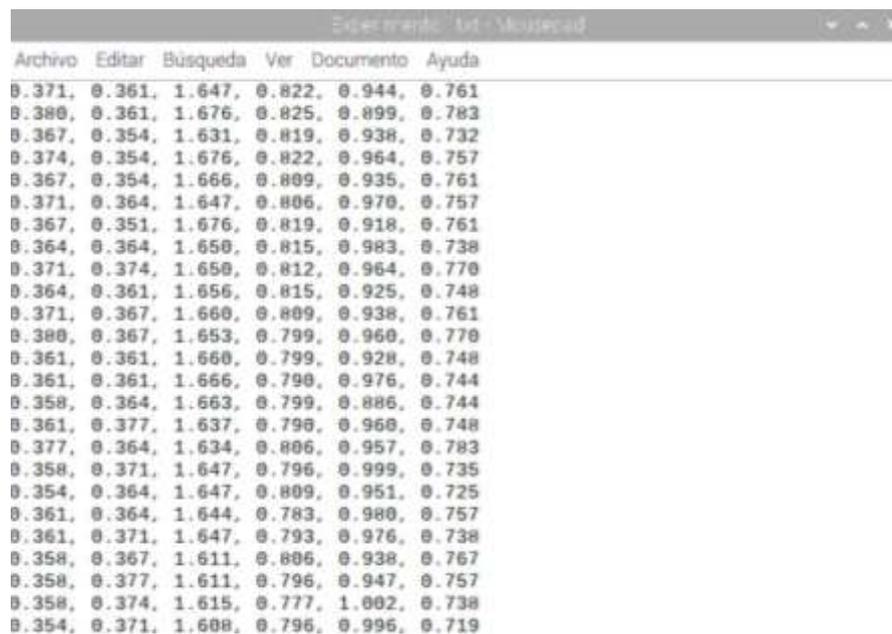
*Nota:* La figura muestra los valores que se obtienen en el transcurso del proceso de adquisición de señales.

- Una vez finalizado el proceso de olfato, se almacenan las señales medidas en un archivo .txt, como se observa en la Figura 44. Luego los datos se envían al ordenador para su preprocesamiento e ingreso al modelo de red neuronal.

#### Figura 44

Archivo .txt con las mediciones realizadas por cada sensor a las muestras de cacao.

(Elaboracion Propia)



Archivo	Editar	Búsqueda	Ver	Documento	Ayuda
0.371,	0.361,	1.647,	0.822,	0.944,	0.761
0.380,	0.361,	1.676,	0.825,	0.899,	0.783
0.367,	0.354,	1.631,	0.819,	0.938,	0.732
0.374,	0.354,	1.676,	0.822,	0.964,	0.757
0.367,	0.354,	1.666,	0.809,	0.935,	0.761
0.371,	0.364,	1.647,	0.806,	0.970,	0.757
0.367,	0.351,	1.676,	0.819,	0.918,	0.761
0.364,	0.364,	1.650,	0.815,	0.983,	0.738
0.371,	0.374,	1.650,	0.812,	0.964,	0.770
0.364,	0.361,	1.656,	0.815,	0.925,	0.748
0.371,	0.367,	1.660,	0.809,	0.938,	0.761
0.380,	0.367,	1.653,	0.799,	0.960,	0.770
0.361,	0.361,	1.660,	0.799,	0.928,	0.748
0.361,	0.361,	1.666,	0.790,	0.976,	0.744
0.358,	0.364,	1.663,	0.799,	0.886,	0.744
0.361,	0.377,	1.637,	0.790,	0.960,	0.748
0.377,	0.364,	1.634,	0.806,	0.957,	0.783
0.358,	0.371,	1.647,	0.796,	0.999,	0.735
0.354,	0.364,	1.647,	0.809,	0.951,	0.725
0.361,	0.364,	1.644,	0.783,	0.980,	0.757
0.361,	0.371,	1.647,	0.793,	0.976,	0.738
0.358,	0.367,	1.611,	0.806,	0.938,	0.767
0.358,	0.377,	1.611,	0.796,	0.947,	0.757
0.358,	0.374,	1.615,	0.777,	1.002,	0.738
0.354,	0.371,	1.608,	0.796,	0.996,	0.719

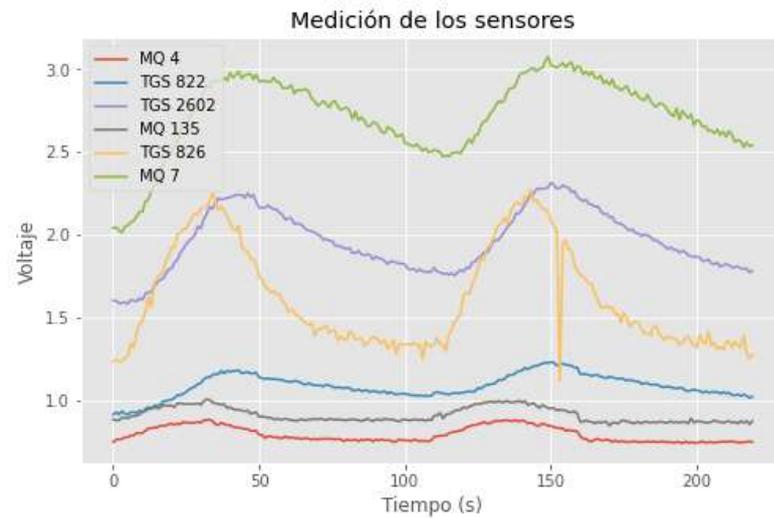
*Nota:* La figura muestra el almacenamiento de los datos en un archivo clasificado por número de experimento.

- Finalmente, en el ordenador se ejecuta el código desarrollado en *Python* para la predicción de la muestra de cacao medida. En donde se puede apreciar: la gráfica con cada medición realizada en el prototipo que se indica en la Figura 45, la gráfica de cada señal preprocesada como la que se observa en Figura 46 y el resultado de cada predicción que se presenta en la Tabla 13.

**Figura 45**

*Señales medidas por los sensores por cada muestra de granos de cacao.*

*(Elaboracion Propia)*

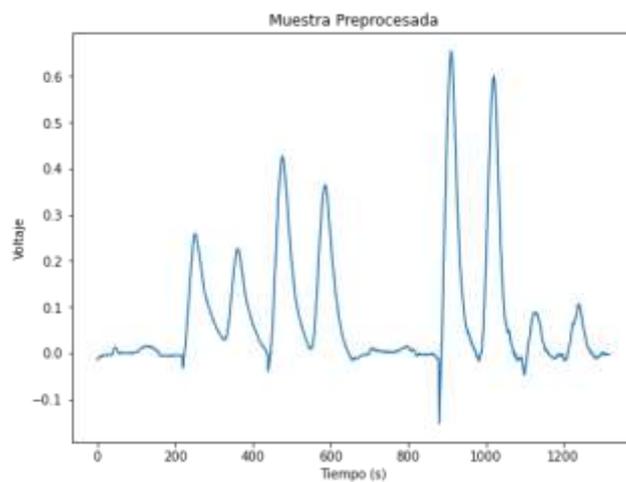


*Nota:* La figura muestra las gráficas de las señales por cada uno de los sensores para visualizar su comportamiento.

**Figura 46**

*Resultado del preprocesamiento de señales para una muestra de granos de cacao.*

*(Elaboracion Propia)*



*Nota:* La figura muestra la gráfica del conjunto de datos finales a analizar.

**Tabla 13**

*Comparación del Valor Real esperado y el Valor Predicho por el prototipo.*

*(Elaboracion Propia)*

<b>Nº</b>	<b>VALOR</b>	<b>SALIDA DEL</b>	<b>PREDICCIÓN</b>
<b>MUESTRA</b>	<b>REAL</b>	<b>PROTOTIPO</b>	<b>(<math>\geq 0.5</math>)</b>
<b>1</b>	0	0.027	0
<b>2</b>	0	0.305	0
<b>3</b>	0	0.011	0
<b>4</b>	0	0.016	0
<b>5</b>	0	0.258	0
<b>6</b>	0	0.318	0
<b>7</b>	0	0.228	0
<b>8</b>	0	0.011	0
<b>9</b>	0	0.027	0
<b>10</b>	0	0.411	0
<b>11</b>	1	0.524	1
<b>12</b>	1	0.954	1
<b>13</b>	1	0.479	0
<b>14</b>	1	0.987	1
<b>15</b>	1	0.928	1
<b>16</b>	1	0.849	1
<b>17</b>	1	0.742	1
<b>18</b>	1	0.956	1
<b>19</b>	1	0.989	1
<b>20</b>	1	0.997	1

*Nota:* La tabla muestra los resultados de la predicción de por la red neuronal, en donde la muestra número 13 presenta un error en el valor esperado.

En la Tabla 13 el Valor Real indica si la muestra de cacao medida pertenece a la clase de cacao sano "1" o caso contrario no pertenece a la clase y se considera como cacao contaminado "0". En la columna de Salida del Prototipo se puede observar el resultado de la predicción del prototipo basado en una función Sigmoide, cuyo valor determina si la muestra pertenece a la clase de cacao sano cuando es

mayor o igual a 0.5 y en caso de pertenecer se registra con "1" en la columna de Predicción.

Una vez obtenidos los veinte resultados de predicción del prototipo, se registran en una Matriz de Confusión presentada en la Tabla 14.

**Tabla 14**

*Matriz de confusión obtenida en las pruebas del prototipo. (Elaboracion Propia)*

		Predicho	
		Contaminado	Sano
Real	Contaminado	TP = 10	FN = 0
	Sano	FP = 1	TN = 9

*Nota:* La tabla muestra los valores para determinar la efectividad del prototipo desarrollado.

Utilizando los valores registrados en la Matriz de Confusión de la Tabla 14, se calculan las métricas del prototipo con ayuda de las ecuaciones (7), (8), (9) y (10).

- **Exactitud:**

$$\frac{TP + TN}{TP + FN + FP + TN} = \frac{10 + 9}{10 + 0 + 1 + 9} = 0.95$$

- **Precisión:**

$$\frac{TP}{TP + FP} = \frac{10}{10 + 1} = 0.91$$

- **Sensibilidad:**

$$\frac{TP}{TP + FN} = \frac{10}{10 + 0} = 1$$

- **Puntaje de F1:**

$$\frac{2 \cdot \textit{Precisión} \cdot \textit{Sensibilidad}}{\textit{Precisión} + \textit{Sensibilidad}} = \frac{2 \cdot 1 \cdot 0.91}{1 + 0.91} = 0.95$$

**Análisis de Resultados**

En la Matriz de Confusión de la Tabla 14 se puede apreciar la distribución de las predicciones realizadas por el prototipo, en donde se detectó la presencia de un Falso Positivo, es decir, una de las 20 predicciones fue incorrecta, al momento de no aprobar una muestra de granos de cacao como sano.

Por lo que las métricas obtenidas indican lo siguiente:

- Exactitud de 95%: El prototipo fue capaz de clasificar correctamente la mayoría de las mediciones, aunque el concepto de exactitud suele presentar confusiones o engaños cuando se tratan de conjuntos de datos desbalanceados, pero en este caso se realizó un conjunto con la mitad de mediciones con granos de cacao sanos y la otra mitad con granos contaminados, por lo que la métrica de exactitud para el prototipo es fiable e indica que en una de cada veinte mediciones, el sistema presenta un error de clasificación.
- Precisión de 91%: Es decir, el prototipo cometió un error en la predicción de los verdaderos valores negativos, debido a que una muestra de cacao sano fue clasificada como uno contaminado, esto implica un reprocesado del lote de donde se analizaron las muestras desde la etapa de fermentación. Garantizando que todo lo clasificado como sano sea verdadero en el proceso de control de calidad.
- Sensibilidad de 100%: El prototipo acertó todas las predicciones verdaderas positivas ya que ningún cacao contaminado lo clasificó como sano, lo que brinda confianza al momento de realizar un proceso de control de calidad, en donde nuevamente todo lo que sea clasificado como un cacao sano sea verdadero.

- Puntaje F1 de 95%: Previamente se pudo apreciar que al calcular la precisión en la ecuación (8) se consideran los falsos positivos y para la sensibilidad en la ecuación (9) se utilizan los falsos negativos, mientras que el puntaje F1 representa ambas métricas y resulta útil para cuando se desea un primer indicio si existió cualquier tipo de error de predicción respecto a los verdaderos positivos.

En este caso el puntaje F1 es del 95% lo que significa que el prototipo presenta errores de predicción en un 5% respecto a las diez muestras clasificadas correctamente como contaminadas.

Se puede concluir que el prototipo tiene un alto índice de exactitud para la clasificación de muestras de granos de cacao. Su exactitud es del 95%, debido a que, en el proceso de pruebas de funcionamiento con una muestra total de veinte granos, falló en una. Por lo tanto, esto permite que el prototipo sea confiable como una ayuda en los procesos de control de calidad del cacao, favoreciendo a la mejora de los procesos industriales y la maximización de su producción.

## Capítulo V

### Conclusiones y Recomendaciones

#### Conclusiones

Se diseñó e implementó un prototipo que permitió extraer los olores del cacao como soporte del proceso de control de calidad. Los olores, luego fueron transportados hacia la cámara de sensores en donde se realizó la medición y almacenamiento de datos dentro de la tarjeta Raspberry Pi3. Luego de esto, los datos medidos se enviaron a un ordenador en donde se procesaron las señales y se efectuó la clasificación de los granos de cacao utilizando un modelo de red neuronal, el cual consta con dos neuronas en la capa de entrada y una neurona en la capa de salida. Después utilizando validación ciega, se calcularon las métricas de funcionamiento del prototipo en donde se encontró una exactitud de predicción del 95% indicando que una de cada veinte predicciones son incorrectas, además el sistema presenta una precisión del 91% lo que significa la existencia de falsas clasificaciones de cacao contaminado, mientras que la sensibilidad es de 100% indicando que ninguna de las muestras contaminadas se clasificaron como sanas, es decir, el sistema es sensible ante pequeñas cantidades de cacao contaminado dentro de una muestra de cacao sano, lo cual garantiza que la calidad del cacao clasificado como sano sea verdadero.

Los sensores químicos implementados en el prototipo midieron las sustancias volátiles presentes en los granos de cacao como amoníaco, monóxido de carbono, metano, óxido nitroso, vapores orgánicos solventes, entre otros, en donde los sensores TGS 2602 y TGS 826 presentaron un patrón de respuesta mayor en la detección de granos de cacao sanos en relación a la respuesta obtenida con granos contaminados. Mientras que en los sensores TGS 822, MQ 7, MQ 4 y MQ 135 se observó un patrón de respuesta mayor ante los estímulos con olores de granos

contaminados. Por lo que la cámara de sensores es capaz de obtener huellas olfativas diferentes según la calidad del cacao.

Se desarrolló un algoritmo en *Python* dentro de la tarjeta *Raspberry Pi3*, que permitió la comunicación con el conversor análogo-digital MCP 3008 utilizando el protocolo SPI y se almacenaron las mediciones en archivos de tipo *.txt* por cada uno de los experimentos realizados.

El sistema de adquisición de señales logró medir en total 130 muestras de cacao sano y contaminado, de los cuales, con ayuda del preprocesamiento de señales, se encontraron doce muestras que presentaban *outliers* debido a que el integrado MCP 3008 al utilizar tecnología CMOS es sensible a variaciones en la alimentación y a perturbaciones externas como el ruido o vibraciones, lo que altera el comportamiento natural de las señales.

Para el entrenamiento de la red neuronal implementada en el prototipo se utilizaron 68 muestras de la base de datos preprocesados, utilizando dos neuronas y una función de activación ReLU para la capa de entrada. Luego de esto, se obtuvo un modelo de red que sirvió de herramienta para la clasificación de los granos de cacao, en donde el 5% de las veinte mediciones obtenidas en la implementación se clasificaron incorrectamente.

En definitiva, la implementación del prototipo es viable como asistente en los procesos de control de calidad para cacao de exportación, debido a que identificó con éxito el 95% de las muestras de granos entregados por las empresas ECO-KAKAO S.A. y de AGROLAYA S.A. El prototipo cuenta con una estructura compacta de fácil instalación y transporte, además incluye un software que muestra los resultados de la clasificación del cacao con sus respectivas gráficas.

## Recomendaciones

Para la etapa de adquisición de datos se aconseja mejorar el sistema de inyección de volátiles, dado que al emplearse un control de encendido y apagado de bombas este genera mayores perturbaciones en el sistema, por consiguiente, se debe considerar utilizar un sistema automatizado controlado por electroválvulas.

Al implementar una red neuronal para la clasificación entre 2 grupos determinados el número de experimentos empleados en el desarrollo de este proyecto fue de 110, con el cual se obtuvieron resultados óptimos, no obstante, se recomienda ampliar el rango de muestras para la base de datos, con un número mayor de experimentos también se aconseja emplear métodos de reducción de dimensiones lo cual mejoraría el tiempo de entrenamiento y desempeño de la red neuronal. En base a esta premisa al utilizar conjuntos de datos más grandes se puede optar por diferentes tipos de arquitecturas de red neuronal.

Al hablar de los datos almacenados se recomienda que estos presenten un formato de datos que pueda ser rápidamente interpretado en Python para su preprocesamiento, en el caso de este proyecto los datos de los sensores se almacenaban en un archivo .txt el cual debía ser modificado para la etapa de preprocesamiento y entrenamiento.

La etapa de entrenamiento y validación se llevó a cabo en un computador externo al prototipo, para futuros trabajos se aconseja embeber todo el sistema en la tarjeta de adquisición, de esta manera se reduce el número de equipos necesarios para la implementación del prototipo.

## Referencias

- Aggarwal, C. (2018). *Neural Networks and Deep Learning*. New York: Springer International Publishing AG.
- Benedetti, S., Mannino, S., Sabatini, A. G., & Luigi, M. G. (2004). Electronic nose and neural network use for the classification of honey. *Apidologie*, 397-402.
- Brownlee, J. (2017). *Deep Learning With Python: Develop Deep Learning Models on Theano and Tensorflow Using Keras*. N/A: Machine Learning Mastery.
- Brudzewski, K., Osowski, S., & Pawlowski, W. (2012). Metal oxide sensor arrays for detection of explosives at sub-parts-per million concentration levels by the differential electronic nose. *Sensors and Actuators B: Chemical*, 528-533.
- Chollet, F. (2018). *Deep Learning with Python*. New York: Manning Publications Co.
- Distante, C., Ancona, N., & Siciliano, P. (2003). Support vector machines for olfactory signals recognition. *Sensors and Actuators B: Chemical*, 30-39.
- Fauset, L. (1994). *Fundamentals of Neural Networks: Architectures, Algorithms, And Applications*. New Jersey: Prentice-Hall.
- Figaro USA Inc. (2002). *Product information for TGS826*. USA.
- Flórez, A., Vargas, J., & Harold, P. (2020). Análisis de componentes principales utilizando python para identificar clúster asociados a muestras de cacao seco sano e infectado con monilia en Norte de Santander. *Aibi*, 16-22.
- Galushkin, A. (2007). *Neural Networks Theory*. Berlin: Springer-Verlag Berlin Heidelberg.
- García, D. (2001). Redes Neuronales: Conceptos Básicos y Aplicaciones. *Academia*, 0-55.
- Gardner, J., & Barlett, P. (1999). *Electronic Noses: Principles and Applications*. Oxford: Oxford University Press.
- Hush, D. (1989). Classification With Neural Networks: A Performance Analysis. *IEEE Xplore*, 277-280.

- Ketkar, N. (2017). *Deep Learning With Python: A Hands-On Introduction*. Bangalore: APRESS.
- Lebrun, M., Plotto, A., Goodner, K., Ducamp, M.-N., & Baldwin, E. (2008). Discrimination of mango fruit maturity by volatiles using the electronic nose and gas chromatography. *Postharvest Biology and Technology*, 122-131.
- Microsoft Docs. (2017, Agosto 28). *Documentación de Windows para IoT*. Retrieved from Raspberry Pi 2 & 3 asignaciones de patillas: <https://docs.microsoft.com/es-es/windows/iot-core/learn-about-hardware/pinmappings/pinmappingsrpi>
- Moreno, I., Caballero, R., Galán, R., Matía, F., & Jiménez, A. (2009). La Nariz Electrónica: Estado del Arte. *Revista Iberoamericana de Automática e Informática Industrial RIAI*, 76-91.
- Nagle, T. H., Gutierrez, O., & Schiffman, S. (1998). The how and why of electronic noses. *IEEE Spectrum*, 22-31.
- Persaud, K., & Dodd, G. (1982). Analysis of discrimination mechanisms in the mammalian olfactory system using a model nose. *Nature*, 352-355.
- Persaud, K., Pisanelli, A. M., Evans, P., & Travers, P. (2006). Monitoring urinary tract infections and bacterial vaginosis. *Sensors and Actuators B: Chemical*, 116-120.
- Plaza, M. (2016). *Industria de Cacao*. Guayaquil: ESPAE - ESPOL.
- Quiñónez, K., Salvatierra, D., Cadena, M., & Baque, L. (2020). Análisis de los niveles de automatización de los procesos de producción de cacao. *Journal of Business and entrepreneurial*, 8-23.
- Rodríguez, M. L. (2001). Análisis sensorial de alimentos: nariz electrónica. *Anales de la Real Sociedad Española de Química*, 10-18.
- Rosales, C., & Reyes, O. (2016). Clasificación de Datos de Olor de Café provenientes de una Nariz Electrónica Utilizando Redes Neuronales. *ResearchGate GmbH*, 1-6.

- Samarasinghe, S. (2007). *Neural Networks for Applied Sciences and Engineering: From Fundamental to Complex Pattern Recognition*. Boca Ratón: Auerbach Publications.
- Servicio Ecuatoriano de Normalización. (1995). NTE INEN 177 Granos de cacao. Muestreo. Ecuador.
- Servicio Ecuatoriano de Normalización. (2006). NTE INEN 176 Granos de cacao. Requisitos. Ecuador.
- Servicio Ecuatoriano de Normalización. (1987). NTE INEN 175 Granos de cacao. Ensayo de Corte. Ecuador.
- Valdez, L., & Gutiérrez, J. (2016). Chocolate Classification by an Electronic Nose with Pressure Controlled Generated Stimulation. *MDPI*, 0-12.