



UNIVERSIDAD TÉCNICA DE AMBATO
FACULTAD DE INGENIERÍA EN SISTEMAS
ELECTRÓNICA E INDUSTRIAL
Carrera de Ingeniería en Electrónica y Comunicaciones

TEMA:

“PROTOTIPO PARA ANÁLISIS DE ALIMENTOS SÓLIDOS USANDO ESPECTROSCOPIA DE INFRARROJO CERCANO”

Proyecto de Graduación. Modalidad: Proyecto de Investigación, presentado previo la obtención del título de Ingeniero en Electrónica y Comunicaciones.

SUBLÍNEA DE INVESTIGACIÓN: Procesamiento digital de señales e imágenes

AUTOR: Luis Alfredo Ramírez Freire

TUTOR: Ing. Mario García.

AMBATO- ECUADOR

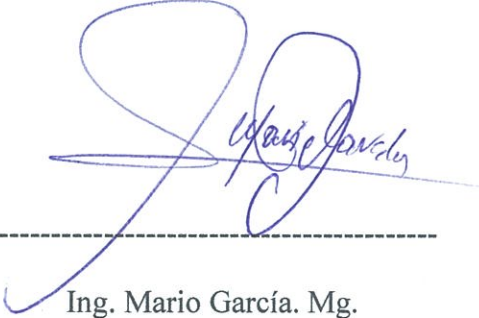
Junio 2018

APROBACIÓN DEL TUTOR

En mi calidad de tutor del Trabajo de Investigación sobre el tema: “PROTOTIPO PARA ANÁLISIS DE ALIMENTOS SÓLIDOS USANDO ESPECTROSCOPIA DE INFRARROJO CERCANO” del señor Luis Alfredo Ramírez Freire estudiante de la Carrera de Ingeniería en Electrónica y Telecomunicaciones, de la Facultad de Ingeniería en Sistemas, Electrónica e Industrial, de la Universidad Técnica de Ambato, considero que el informe investigativo reúne los requisitos suficientes para que continúe con los trámites y consiguiente aprobación de conformidad con el numeral 7.2 de los Lineamientos Generales para la aplicación de Instructivos de las Modalidades de Titulación de las Facultades de la Universidad Técnica de Ambato.

Ambato junio, 2018

EL TUTOR



Ing. Mario García. Mg.

AUTORÍA

El presente Proyecto de Investigación titulado: “PROTOTIPO PARA ANÁLISIS DE ALIMENTOS SÓLIDOS USANDO ESPECTROSCOPIA DE INFRARROJO CERCANO”, es absolutamente original, auténtico y personal, en tal virtud, el contenido, efectos legales y académicos que se desprenden del mismo son de exclusiva responsabilidad del autor.

Ambato junio, 2018



Luis Alfredo Ramírez Freire

CC: 1803873247

DERECHOS DE AUTOR

Autorizo a la Universidad Técnica de Ambato, para que haga uso de este Trabajo de Titulación como un documento disponible para la lectura, consulta y procesos de investigación.

Cedo los derechos de mi Trabajo de Titulación, con fines de difusión pública, además autorizo su reproducción dentro de las regulaciones de la Universidad.

Ambato junio, 2018



Luis Alfredo Ramírez Freire

CC: 1803873247

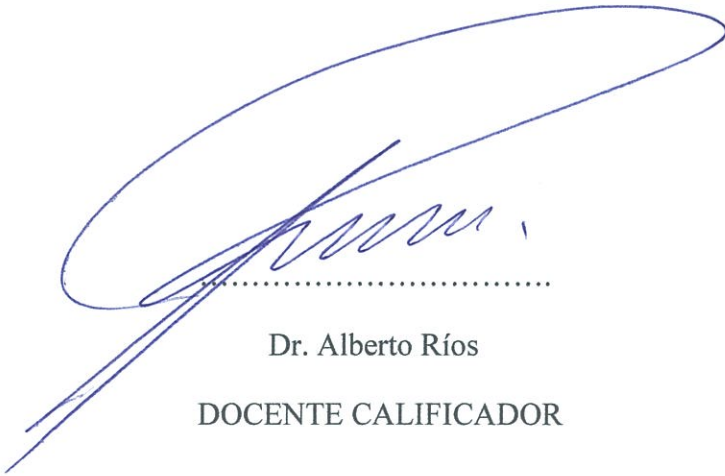
APROBACIÓN DE LA COMISIÓN CALIFICADORA

La Comisión Calificadora del presente trabajo conformada por los señores docentes Dr. Alberto Ríos e Ing. Marco Jurado revisó y aprobó el Informe Final del Proyecto de Investigación titulado **“PROTOTIPO PARA ANÁLISIS DE ALIMENTOS SÓLIDOS USANDO ESPECTROSCOPIA DE INFRARROJO CERCANO”**, presentado por el señor Luis Alfredo Ramírez Freire de acuerdo al numeral 9.1 de los Lineamientos Generales para la aplicación de Instructivos de las Modalidades de Titulación de las Facultades de la Universidad Técnica de Ambato.



.....
Ing. Mg. Elsa Pilar Urrutia

PRESIDENTA DEL TRIBUNAL



.....
Dr. Alberto Ríos

DOCENTE CALIFICADOR



.....
Ing. Marco Jurado

DOCENTE CALIFICADOR

DEDICATORIA:

A mis padres y de manera especial a mi madre, que es la mujer que con su infinito amor y tenacidad para sacar a sus hijos adelante ha estado apoyándome en este trajinar que representa la vida.

Luis Alfredo Ramírez Freire

AGRADECIMIENTO:

Agradezco a la universidad que me brindó los conocimientos necesarios para llevar a cabo este trabajo de investigación. Así como a mi tutor Ing. Mario García que con su guía me apoyó para la consecución de este trabajo de investigación

Luis Alfredo Ramírez Freire

ÍNDICE GENERAL

APROBACIÓN DEL TUTOR.....	ii
AUTORÍA.....	iii
DERECHOS DE AUTOR	iv
APROBACIÓN DE LA COMISIÓN CALIFICADORA	v
DEDICATORIA:	vi
AGRADECIMIENTO:	vii
ÍNDICE GENERAL.....	viii
ÍNDICE DE FIGURAS.....	xii
ÍNDICE DE TABLAS	xiv
RESUMEN.....	xv
ABSTRACT	xvi
INTRODUCCIÓN	xvii
CAPÍTULO 1	1
1.1 TEMA.....	1
1.2 PLANTEAMIENTO DEL PROBLEMA.....	1
1.3 DELIMITACIÓN.....	5
1.4 JUSTIFICACIÓN.....	6
1.5 OBJETIVOS.....	7
1.5.1 Objetivo General:	7
1.5.2 Objetivos Específicos:.....	7
CAPÍTULO 2	8
2.1 ANTECEDENTES INVESTIGATIVOS.....	8
2.2 FUNDAMENTACIÓN TEÓRICA	11
2.2.1 Teoría de Espectroscopía NIR	11

2.2.2	Reflectancia NIR.....	13
2.2.3	Transmisión NIR.....	14
2.2.4	Teoría de espectroscopía vibracional	14
	Interacción entre la luz y la materia física	14
	Oscilador Armónico Simple.....	15
	El oscilador armónico simple es un sistema formado por un objeto que pende de un resorte, dicho objeto produce un movimiento oscilatorio, tendiendo al equilibrio y describiendo movimientos que se repiten cada cierto período de tiempo. [19].....	15
	Oscilador Armónico como descriptor del NIR	15
2.3	CALIBRACIÓN DE UN INSTRUMENTO	17
2.4	CALIBRACIONES	17
2.5	ANÁLISIS DE COMPONENTES PRINCIPALES (PCA)	19
2.5.1	Descripción del método.....	21
2.6	REGRESIÓN POR MÍNIMOS CUADRADOS PARCIALES	23
2.7	TARJETAS PARA ADQUISICIÓN DE DATOS	24
2.7.1	Arduino Tian	26
2.7.2	Raspberry Pi 3	26
2.7.3	BeagleBoard.....	27
2.7.4	ASUS Tinkerboard.....	28
2.7.5	UDOO	28
2.8	LENGUAJES DE PROGRAMACIÓN PARA ANÁLISIS DE DATOS	29
2.8.1	Lenguaje de programación R	29
	Características y ventajas de la programación en R.....	29
2.8.2	Lenguaje de programación PYTHON.....	30
	Características y ventajas de la programación en PYTHON	30

2.8.3	XLSTAT para Microsoft Excel.....	31
2.9	PROPUESTA DE SOLUCIÓN.....	32
CAPÍTULO 3		33
3.1	MODALIDAD DE LA INVESTIGACIÓN.....	33
	Investigación Aplicada.....	33
	Investigación Bibliográfica	33
	Investigación Experimental.....	34
3.2	PROCESAMIENTO Y ANÁLISIS DE DATOS.....	34
3.3	DESARROLLO DEL PROYECTO.....	35
CAPÍTULO 4		36
4.1	INTRODUCCIÓN.....	36
4.2	ANÁLISIS DE FACTIBILIDAD	36
	4.2.1 Factibilidad Técnica	36
	4.2.2 Factibilidad Económica.....	37
	4.2.3 Factibilidad Bibliográfica.....	37
4.3	SISTEMA NIR	37
	4.3.1 Región espectral NIR	37
	4.3.2 Métodos analíticos NIR.....	39
	Análisis de los componentes principales (PCA)	39
	Mínimos parciales cuadrados (PLS)	40
	4.3.3 Estándares	41
4.4	DISEÑO DEL PROTOTIPO NIR.....	43
	4.4.1 Sensor de Infrarrojo Cercano	44
	Sensor AS7263.....	45
	4.4.2 Placa Qwicc para Arduino	47

4.4.3	Protocolo I2C	47
4.5	SISTEMA DE MONITOREO.....	49
4.5.1	Sistema de procesamiento de información.....	50
4.5.2	Comunicación con la placa Raspberry Pi 3.....	52
4.5.3	Configuración de Software	52
4.5.4	Diseño de la Interfaz Gráfica	53
4.6	DISEÑO E IMPLEMENTACIÓN DEL SERVIDOR	55
4.6.1	Instalación del Servidor LAMP	55
4.6.2	Base de Datos	57
4.6.3	Integración con Python	58
4.7	ANÁLISIS DE RESULTADOS	59
4.8	ANÁLISIS ECONÓMICO DEL PROYECTO.....	70
CAPÍTULO 5		72
5.1	CONCLUSIONES.....	72
5.2	RECOMENDACIONES	73
BIBLIOGRAFÍA		74
ANEXO A.....		81
ANEXO A.....		87
ANEXO C		91
ANEXO D.....		95

ÍNDICE DE FIGURAS

Figura 1.1 Gráfico temperatura vs. tiempo de un proceso de digestión [6].....	3
Figura 2.1 Onda típica propagándose a través del espacio [16].....	12
Figura 2.2 Espectro electromagnético. [16].....	12
Figura 2.3 El enlace vibratorio entre los átomos de carbono (C) e hidrógeno (H) absorbe ondas NIR de una longitud de onda particular mientras otras ondas son reflejadas [17]	13
Figura 2.4 Rayos NIR absorbidos por algunos enlaces, reflejados por otros (Reflectancia NIR) o transmitidos a través de una muestra (Transmisión NIR generalmente nombrada NIT). [8].....	14
Figura 2.5 Distribución de probabilidad conjunta de dos variables aleatorias x_1 y x_2	20
Figura 2.6 Nueva distribución de probabilidad conjunta de las variables tras un procesamiento PCA.....	21
Figura 2.7 Arduino Tian [14].....	26
Figura 2.8 Raspberry Pi 3 [15].....	27
Figura 2.9 BeagleBoard [10].....	27
Figura 2.10 ASUS Tinker Board [11].....	28
Figura 2.11 UDOO Quad/Dual [16].....	29
Figura 4.1 Picos en las bandas de absorción [20].....	38
Figura 4.2 Curvas de observación del prototipo NIR.....	39
Figura 4.3 PCA de las muestras.....	40
Figura 4.4 PLS del canal R.....	41
Figura 4.5 Diagrama de Bloques del Prototipo NIR.....	43
Figura 4.6 Diagrama de bloques del chip [32].....	46
Figura 4.7 Placa con AS7263 de Sparkfun.....	46
Figura 4.8 Placa Qwicc para Arduino [33].....	47
Figura 4.9 Programa en arduino para la lectura del sensor.....	48
Figura 4.10 Monitor serie de Arduino con lecturas del AS7263.....	49
Figura 4.11 Componentes principales de las muestras.....	58
Figura 4.12 Gráfico de sedimentación.....	60
Figura 4.13 Variables (ejes F1 y F2: 99,98 %).....	62
Figura 4.14 Observaciones (ejes F1 y F2: 99,98 %).....	65

Figura 4.15 Biplot (ejes F1 y F2: 99,98 %)65

ÍNDICE DE TABLAS

Tabla 1.1 Precios de reactivos de grado analítico	4
Tabla 2.1 Tipos de tarjetas de adquisición	25
Tabla 4.1 Longitudes de onda en nm y número de ondas en algunas bandas de infrarrojo cercano de compuestos orgánicos y compuestos sulfúricos [20].....	38
Tabla 4.2 Selección de estándares internacionales que usan NIR [20].....	42
Tabla 4.3 Comparación de sensores NIR	44
Tabla 4.4 Comparación entre placas Raspberry.....	50
Tabla 4.5 Características de la Raspberry Pi 3.....	51
Tabla 4.6 Estadísticos descriptivos (Datos cuantitativos).....	59
Tabla 4.7 Valores propios	59
Tabla 4.8 Vectores propios:	60
Tabla 4.9 Cargas factoriales:.....	61
Tabla 4.10 Correlaciones entre las variables y los factores:	61
Tabla 4.11 Contribuciones de las variables (%):	62
Tabla 4.12 Cosenos cuadrados de las variables:	63
Tabla 4.13 Puntuaciones factoriales:.....	63
Tabla 4.14 Contribuciones de las observaciones (%):	66
Tabla 4.15 Cosenos cuadrados de las observaciones:.....	67
Tabla 4.16 Análisis del costo	70

RESUMEN

En la presente investigación se realiza un estudio con el objetivo de diseñar e implementar un prototipo que permita el análisis de alimentos sólidos, basado en espectroscopía de infrarrojo cercano, que permita determinar el contenido porcentual de proteínas, grasa y humedad. Esta técnica usa el principio de reflexión de la luz en el espectro de la luz no visible. Las radiaciones electromagnéticas que tienen una determinada longitud de onda parten desde una fuente de luz hasta un material puesto en el punto de observación y hacen que los enlaces atómicos de las sustancias presentes vibren de tal manera que una cierta cantidad de energía de los fotones sea absorbida por dichos enlaces y el resto rebote hacia un arreglo de diodos que se encarga de filtrar esas ondas y basado en la cantidad de energía que ingresa al arreglo, calcular mediante métodos estadísticos, la cantidad de proteína, grasa y humedad presentes en la muestra.

Para este trabajo se ha usado hardware y software libres, su uso permite abaratar el costo del prototipo propuesto para análisis de alimentos, basado en espectroscopía de infrarrojo cercano, descrito en el presente trabajo de investigación. En el hardware se han usado circuitos de reciente lanzamiento en lo que respecta al sensor de NIR (reflectancia de infrarrojo cercano) como es el AS7263, para la parte de control se ha usado el microcontrolador Arduino y para el sistema de análisis de datos una Raspberry Pi 3.

El software, por otro lado, se ha implementado en lenguaje Python, el mismo que está siendo muy utilizado en lo que respecta al análisis de datos (big data), para la aplicación de los métodos estadísticos que son necesarios para la determinación de los valores de humedad, proteína y grasa.

En el análisis de los resultados se hace uso de la herramienta estadística XLSTAT que es ofrecida como un complemento del paquete de Microsoft Excel.

Palabras clave: Arduino, Big data, espectroscopia, infrarrojo, Python, PCA, PLSR, XLSTAT.

ABSTRACT

In the present investigation a study is made for the design of a prototype for the analysis of the quality of food using near infrared spectroscopy, this technique uses the principle of reflection of the light in the spectrum of non-visible light. Waves that have a certain wavelength come out from a light source to a material placed at the point of observation and make atomic bonds of the present substances to vibrate in such a way that a certain amount of energy of the photons is absorbed by them and the rest bounce to a diode array that is responsible for filtering those waves and then calculate, by statistical methods, the amount of protein, fat and moisture present in the sample.

For this work, free hardware and software have been used, making it possible to reduce the costs of the prototype described here. In the hardware, recently launched circuit, AS7263, has been used as near infrared reflectance sensor. For the control part the Arduino microcontroller has been used and for the data analysis system a Raspberry Pi 3.

The software, on the other hand, has been implemented in the Python language, the same one that is being widely used in the analysis of data (big data), for applying the statistical methods that are necessary for the determination of moisture, protein and fat values.

Keywords: Arduino, Big data, near infrared spectroscopy, statistical methods, Python.

INTRODUCCIÓN

En el presente proyecto de investigación se diseña e implementa un prototipo para el análisis de alimentos mediante la técnica de espectroscopía de infrarrojo cercano, basado en una tarjeta de adquisición de datos para la interpretación de las salidas del sensor de infrarrojos, además de la creación de una interfaz de monitoreo donde se observa los resultados arrojados de las variables sensadas.

A continuación, se hace una descripción de los capítulos de este proyecto de investigación:

Primer Capítulo: Se justifica la necesidad del desarrollo de la investigación para la implementación de un prototipo basado en la reflectancia de infrarrojo cercano para el análisis de alimentos sólidos, explicando las causas y consecuencias que lo originan, y por último los objetivos de la investigación.

Segundo Capítulo: Consta de los antecedentes sobre investigaciones realizadas del tema prototipo propuesto, así como sobre el uso de sistemas de espectroscopía de infrarrojo cercano para el análisis de distintas sustancias en alimentos y productos farmacéuticos. Así mismo, se desarrolla el marco teórico basado en el tema de investigación.

Tercer Capítulo: Especifica las diferentes técnicas de investigación a las que se debió recurrir, la forma de obtención y análisis de la información. Así mismo se presenta el proceso para el desarrollo del proyecto.

Cuarto Capítulo: Se describe, de una manera detallada, la construcción del prototipo y las pruebas de funcionamiento del sistema completo.

Quinto Capítulo: Se muestran las conclusiones y recomendaciones obtenidas durante el desarrollo del proyecto de investigación y construcción del Prototipo para el análisis de alimentos mediante la técnica de espectroscopía de infrarrojo cercano.

EL PROBLEMA

1.1 TEMA

Prototipo para análisis de alimentos sólidos usando espectroscopia de infrarrojo cercano

1.2 PLANTEAMIENTO DEL PROBLEMA

Desde el punto de vista de la bromatología, que es el estudio de los alimentos bajo una consideración científica, se debe abarcar el análisis químico, físico, nutritivo, higiénico, entre otros de los alimentos. [1] Tomando en cuenta la necesidad de identificar diferentes compuestos tales como: proteínas, carbohidratos, etc.; se considera la importancia de todos estos estudios y sus resultados para garantizar a la sociedad alimentos con contenidos de sustancias que favorezcan su alimentación.

Se ha encaminado a la búsqueda de diferentes formas de caracterización, debido a que los métodos cuantitativos y cualitativos tradicionales son complejos, demandan tiempo, mano de obra y son costosos; basado en el tarifario de ensayos y servicios del Instituto Ecuatoriano de Normalización, el análisis de proteína en alimentos tiene un costo de 33.00 dólares; humedad a presión atmosférica normal en alimentos, un costo de 11.00 dólares; y de grasas en alimentos un costo de 27.50 dólares. Cabe señalar que estos costos son por muestras. [2] [3] Los costos antes mencionados se transfieren a los ciudadanos, pero con la presente investigación se busca que éstos sean los mínimos posibles.

En las industrias: alimenticias, farmacéuticas y petroquímica, se utilizan los métodos químicos y químico-biológicos tradicionales para analizar muestras de sus productos y determinar su calidad.

En la rama alimenticia se usa el análisis proximal o de Weende, que fue desarrollado en 1860, el que es el más conocido y usado para la determinación de la calidad nutritiva parcial de un alimento. [4] Con éste se obtiene seis principios nutritivos, o grupos, dentro de los cuales podrían estar incluidos los siguientes compuestos:

- Humedad: agua y compuestos volátiles.
- Ceniza: materia inorgánica en general.
- Proteína bruta: proteínas, péptidos, aminoácidos, bases nitrogenadas, amidas, nitrógeno vitamínico.
- Extracto etéreo: grasas, ceras, resinas, lípidos complejos, pigmentos, vitaminas liposolubles.
- Fibra bruta: celulosa, hemicelulosa, lignina insoluble.
- Extracto libre de nitrógeno (ELN): almidón, azúcares, pectinas, pigmentos, vitaminas hidrosolubles.

Sólo para citar un ejemplo en la industria alimenticia para la determinación de la proteína bruta (PB), se utiliza un método que data de 1883 y se llama el método Kjeldahl, el que consiste en hervir una muestra en ácido sulfúrico concentrado y mediante métodos químicos determinar la cantidad de proteína presente en la muestra [5]. Para realizarlo se requiere los siguientes implementos: batería digestora y matraces Kjeldahl, aparato Kjeldahl de destilación y valoración, ácido sulfúrico concentrado, catalizador, solución de hidróxido sódico (30%), ácido bórico con indicador y ácido clorhídrico valorado (0,1 N). En la figura 1.1 se muestra el gráfico de la temperatura en una batería digestora con respecto al tiempo.

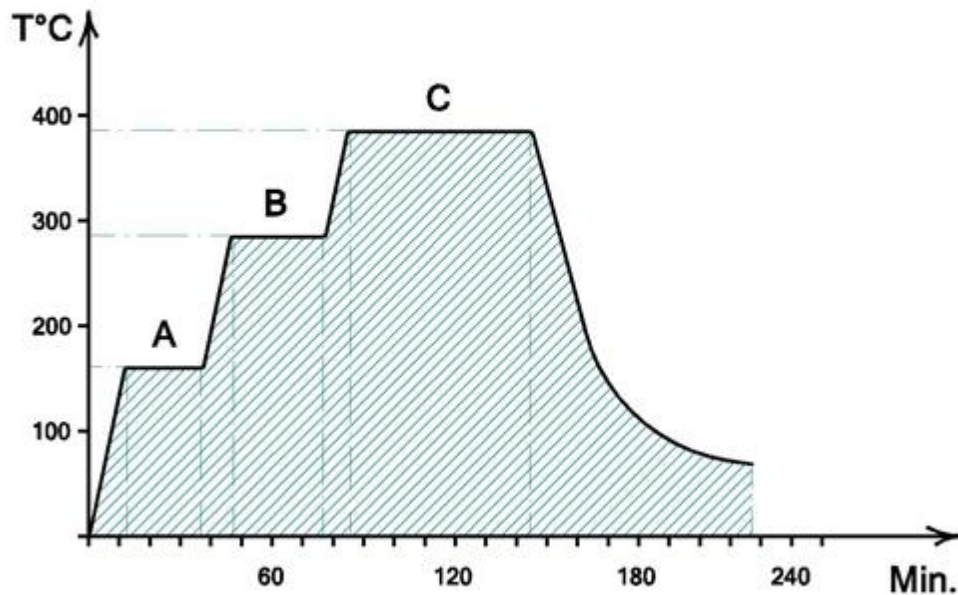


Figura 0.1 Gráfico temperatura vs. tiempo de un proceso de digestión [6]

Por la cantidad de insumos y el proceso como tal, se podría hablar de que se tiene una pérdida de tiempo que se puede ahorrar utilizando otro tipo de análisis.

El consumo de reactivos es un problema adicional y obviamente no aislado, éste conlleva a que se incrementen los desperdicios en los análisis, acompañando a este proceso un aumento en el costo inherente a los reactivos, además se generan costos energéticos, de agua, y por supuesto de personas capacitadas en las labores operativas asociadas a este proceso de análisis. Adicionalmente puede haber costos elevados que obedecen a reactivos importados que son de grado analítico. [7] En la tabla 1.1 se puede observar el precio de algunos de los reactivos de grado analítico empleados en el análisis de muestras:

Tabla 0.1 Precios de reactivos de grado analítico

Descripción	Presentación	Precio
Hidróxido de sodio 97%	FR 500 GR	\$10.17
Alcohol metílico 99.8%	FR 5000 ML	\$29.42
Metil Etil Cetona	FR 1000 ML	\$29.31
Tolueno	FR 1000 ML	\$13.45
Molibdato de Amonio	FR 1000 GR	\$227.57

Fuente: Investigador [8]

Desde la década de los setenta se ha venido utilizando a nivel mundial la espectroscopia como una alternativa a los métodos químicos y químico-biológicos tradicionales, con un muy buen potencial para obtener estimaciones seguras y muy rápidas de la composición química de algunos sistemas biológicos, la interacción de la radiación con las moléculas se controla mediante la medición de las reflectancias dentro del espectro visible (VIS) y del infrarrojo (IR). [9]

Los componentes bioquímicos tales como pigmentos, proteínas, carbohidratos, aceites, agua, la morfología y la estructura de un alimento influyen en la tendencia para absorber, transmitir o reflejar la radiación de diferentes longitudes de onda. Cuando la luz incide sobre una muestra, los fotones viajan a través de ésta. Dichos fotones son reflejados o absorbidos por algunos enlaces covalentes que se acoplan con la frecuencia o longitud de onda exacta de la radiación lumínica. Al absorber energía, los enlaces de las moléculas vibran en dos formas: se extienden aumentando la distancia interatómica a lo largo del eje entre dos átomos (frecuencias más altas), o se doblan (frecuencias más bajas) lo que produce un cambio en el ángulo del enlace entre los átomos.

La energía total reflejada por la muestra es la suma de la reflexión especular (superficial o forma de espejo) más la reflexión difusa que es la temporalmente absorbida por la muestra y sólo ésta proporciona información útil acerca de su composición.

Dado que la industria alimenticia en nuestro país se maneja con rigurosas normativas tanto nacionales como internacionales, y que los reactivos necesarios para los análisis deben ser

importados, la técnica de infrarrojo cercano (NIR) es la mejor alternativa para suprimir dichos costos, porque ésta es una técnica rápida, no invasiva, ni destructiva y, lo más, exacta, siempre y cuando se sigan los parámetros adecuados de calibración del equipo. [7]

La Facultad de Ingeniería en Alimentos de la Universidad Técnica de Ambato dispone de un equipo NIR, el cual ha sido calibrado para varios proyectos enfocados a los cereales, subproductos de cereales y alimentos balanceados que disminuyen el tiempo de respuesta en toda la línea de alimentos. Sin embargo, esta máquina no se encuentra totalmente operativa, porque no existe el personal capacitado para su uso. Las calibraciones que se encuentran en el equipo fueron realizadas por investigadores que ya no laboran en la institución.

Considerando que la técnica de espectroscopía por infrarrojo cercano es más rápida y no necesita de reactivos extras, se desea aplicar este método de análisis en la presente investigación, pero utilización hardware y software libres.

1.3 DELIMITACIÓN

Universidad:	Universidad Técnica de Ambato
Facultad:	Ingeniería en Sistemas, Electrónica e Industrial
Área Académica de la Carrera:	Comunicaciones
Línea de Investigación:	Sistemas de control
Sublínea:	Procesamiento digital de señales e imágenes

Delimitación Espacial:

El presente trabajo de investigación se realizó en la Universidad Técnica de Ambato.

Delimitación Temporal:

El trabajo se lo realizó en la ciudad de Ambato, durante el período septiembre 2017 – febrero 2018, de acuerdo a lo establecido en el Reglamento de Graduación para obtener el Título Terminal de Tercer Nivel de Universidad Técnica de Ambato.

1.4 JUSTIFICACIÓN

A través de esta investigación se busca implementar un sistema que sea capaz de analizar las propiedades de los alimentos sólidos para determinar los contenidos proteínicos, las grasas y la humedad presentes en ellos a través de la absorción y reflexión de las diferentes longitudes de onda de los alimentos, en el espectro del infrarrojo cercano entre los 600-2500nm.

Además, se busca que la técnica NIR, al ser ejecutada con software y hardware libre, sea de mayor utilización en las industrias, de manera particular en la alimenticia, para garantizar que los productos elaborados lleguen al consumidor con excelente calidad, en lo concerniente al porcentaje de proteína, grasa y humedad presente, y éste se beneficie sabiendo que lo que va a consumir contribuirá a una mejor alimentación, y por lo tanto una mejor salud.

Con más frecuencia, lo que la sociedad demanda es una alta calidad de los productos que consume, desde los más novedosos aparatos tecnológicos hasta los productos más típicos de los supermercados. Sin embargo, la calidad de un producto no es una preocupación exclusiva de los consumidores; las empresas que los producen también están interesadas en esa alta calidad ya que de ello dependerán factores como su balance económico, su posición en los mercados o la percepción que los consumidores tienen de ella, siendo ellos los primeros beneficiarios.

Otros beneficiarios son los estudiantes de Ingeniería en Alimentos, porque pueden experimentar con esta técnica innovadora para realizar sus prácticas.

Con el uso de este prototipo se interrelacionan la Ingeniería Electrónica con otros campos; en este caso especial con la Ingeniería en Alimentos, llevando lo aprendido en las aulas al campo práctico.

En el presente proyecto se usarán dispositivos electrónicos tales como un espectrómetro digital en el rango de infrarrojo cercano, una placa controladora, entre otros, y software de código abierto, lo que representa una gran ventaja tanto económica como de implementación.

1.5 OBJETIVOS

1.5.1 Objetivo General:

Implementar un prototipo de análisis de alimentos sólidos con espectrometría de infrarrojo cercano.

1.5.2 Objetivos Específicos:

- Realizar un estudio de la absorción y reflexión de las diferentes longitudes de onda, del espectro infrarrojo, en los alimentos.
- Aplicar al diseño del prototipo las normativas vigentes para análisis de alimentos con la técnica de espectroscopia de infrarrojo cercano.
- Diseñar el prototipo NIR, usando software y hardware libre.

MARCO TEÓRICO

1.6 ANTECEDENTES INVESTIGATIVOS

De la búsqueda en diferentes repositorios relacionados con el conocimiento humano, sobre trabajos con temas de investigación relacionados con el uso de espectroscopía de infrarrojo cercano, se considera importante citar algunos de ellos que servirán como soporte y pauta para la presente investigación.

Jhon Fredy Ocampo, en el año 2017, en Manizales Colombia, en su trabajo **“Estandarización de las curvas de calibración, por la metodología NIR y la química húmeda en las materias primas y carnes frías para la optimización de las respuestas de análisis”**, usa técnicas clásicas como el análisis de cloruros basado en el principio de precipitación con nitrato de plata, y el cromato de potasio como indicador final y lo compara con un equipo NIR para obtener las desviaciones en los resultados. De lo que obtiene que un análisis bromatológico completo tarda cerca de seis horas hábiles, mientras que con el equipo NIR, dicho análisis se reduce a un minuto. [7]

Otro trabajo realizado por HanjongYou, Younsik Kim, JaeHyung Lee y SunwoongChoi, en Corea del Sur, en el año 2016, titulado **“Classification of food powders using handheld NIR spectrometer”**, utiliza un espectómetro NIR de la marca LinkSquare, que es un espectómetro basado en silicio para obtener sus propios datos y así detectar el tipo de polvos alimenticios, utilizados comúnmente, pero visualmente indistinguibles, tales

como azúcar, sal, crema, harina, maíz, arroz, fréjol y papa. Como resultado de su investigación se pudo distinguir, con una alta probabilidad, los compuestos antes mencionados, además, se deduce que es posible identificar también compuestos orgánicos genéticamente modificados, [10]

Weiran Song, Hui Wang, Paul Maguire, Omar Nibouche, en Reino Unido, en el año 2017, en su trabajo con el tema **“Differentiation of organic and non-organic apples using near infrared reflectance spectroscopy – a pattern recognition approach”**, se hace uso de un equipo NIRS, que es un elemento pequeño que usa ondas de infrarrojo cercano para la determinación de manzanas orgánicas y no orgánicas creando curvas para así tener un patrón de reconocimiento. Como resultado se obtiene que se puede hacer un reajuste de curvas usando otros métodos para así llegar a tener un valor de reconocimiento cercano al 86%. [11]

Anna Peguero Gutiérrez, en Barcelona, España, en el año 2010, en su trabajo con el tema **“La espectroscopía NIR en la de terminación de propiedades físicas y composición química de intermedios de producción y productos acabados”**, desarrolla nuevas metodologías de análisis basadas en la espectroscopía NIR con el objetivo de otorgar un valor añadido al producto final, tanto en la industria del cemento como en la industria farmacéutica. En la primera se han obtenido como resultado dos perfiles diferentes de distribución de partículas con diferente números de puntos en la curva (número de tamices utilizados en la determinación) construyendo dos modelos de calibración para cada curva para obtener la mejor calidad de ajuste en términos de coeficiente de correlación entre la curva de referencia y la predicha en el modelo quimiométrico. En el campo farmacéutico se ha determinado que numerosos factores físicos y químicos influyen de manera importante en la calidad de un fármaco, y para asegurar la calidad del producto final es necesario realizar un completo y amplio control de ellos durante el proceso, desarrollando una nueva metodología para incorporar la variabilidad del proceso productivo reflejada en el espectro NIR al conjunto de calibración. [2]

La Excelentísima Diputación Provincial de Pontevedra, con su Servicio Agrario y el Laboratorio Mouriscade, en la provincia de Galicia, España, en el año 2016, realizó un

estudio titulado **“Aplicación de la espectroscopía NIR al análisis de algunos parámetros nutricionales en habas”**, utilizando 121 muestras de habas de Galicia procedentes de cosechas propias, del procesado industrial y al granel, sometieron estas muestras a una identificación por la metodología NIR; obteniendo como resultados que la espectrofotometría de infrarrojo cercano es una técnica adecuada para la determinación de humedad y proteína en habas sin tratamiento previo de la muestra. Además, la espectrofotometría de infrarrojo cercano no es una técnica adecuada, según este estudio, para la determinación de cenizas y materia grasa en habas. [12]

Elisângela Serenato Madalozzo, Elenise Sauer, Noemi Nagata, el año 2013, en su trabajo **“Determination of fat, protein and moisture in ricotta cheese by near infrared spectroscopy and multivariate calibration”**, haciendo uso de un analizador de espectroscopía de infrarrojo cercano (FEMTO, Model NIR 900 PLS) determinan la cantidad de proteína, humedad y grasa en una muestra de queso ricota. Para lograrlo utilizan el análisis de componentes principales (PCA) y la regresión por mínimos cuadrados (PLSR). Del estudio se deduce que usando los métodos de análisis de multivariable pueden obtener valores de medida con un error mucho más bajo que con otros métodos. [13]

F. Estienne, L. Pasti, V. Centner, B. Walczak, F. Despagne, D. Jouan Rimbaud, O.E. de Noord, D.L. Massart, el año 2002, en su publicación titulada **“A comparison of multivariate calibration techniques applied to experimental NIR data sets Part II. Predictive ability under extrapolation conditions”**, se realiza un estudio del desempeño de diferentes técnicas de calibración para instrumentos NIR y da los aspectos de calibración cuando se encuentra con set de datos que no presentan homogeneidad y en los que aparecen variables irrelevantes para la predicción. El estudio concluye que cada método debe ser usado de acuerdo a la calidad del set de datos que se obtenga y viendo cuál es la particularidad encontrada en cada uno de ellos. [14]

Haiyan Cen y Yong He, en el año 2007, en su trabajo con el nombre **“Theory and application of near infrared reflectance spectroscopy in determination of food quality”**, muestran la necesidad de buscar la calidad en los alimentos mediante el uso de la técnica de espectroscopía de infrarrojo cercano (NIRS) obteniendo las cantidad de

grasa, proteínas y otros compuesto presentes en los alimentos, ya que éstos están directamente relacionados con la calidad de vida y la necesidad de mejorar la alimentación de las personas. En ese trabajo se hacen los análisis necesarios para determinar la cantidad de energía absorbida por los enlaces atómicos de las moléculas de los compuestos presentes en los alimentos, además se hace un cuadro de referencia para saber qué compuestos se pueden detectar en diferentes alimentos mediante el uso de la tecnología NIRS. [15]

1.7 FUNDAMENTACIÓN TEÓRICA

1.7.1 Teoría de Espectroscopía NIR

La palabra “espectroscopia” se deriva de la raíz latina spectrum cuyo significado es apariencia o imagen y la palabra griega skopia que significa ver, por lo que se refiere a la observación del espectro electromagnético en un determinado rango [16]. Esta definición es más bien descriptiva de la medición espectroscópica en sí misma. Por ejemplo: visualizar una leve imagen procedente de una muestra. La tecnología NIR involucra luz interactuando con un material, en el que una radiación electromagnética se presenta en forma de ondas. La longitud de onda es la distancia entre los dos picos o puntos altos, y se indica con el símbolo λ de la figura 2.1. La longitud de onda en el espectro NIR se mide normalmente en nanómetros (nm) donde $1\text{nm} = 10^{-9}\text{ m}$ o $1000\text{nm} = .001\text{mm}$.

La parte del espectro visible al ojo humano se extiende desde aproximadamente 400nm a 800nm, mientras que el espectro infrarrojo se extiende aproximadamente desde los 2,500nm a 25,000nm. El infrarrojo cercano es considerado la parte del espectro situada entre la región visible y la región infrarroja. El rango de longitudes de onda que la NIR cubre, está entre 750nm a 2,600nm como se muestra en la figura 2.2.

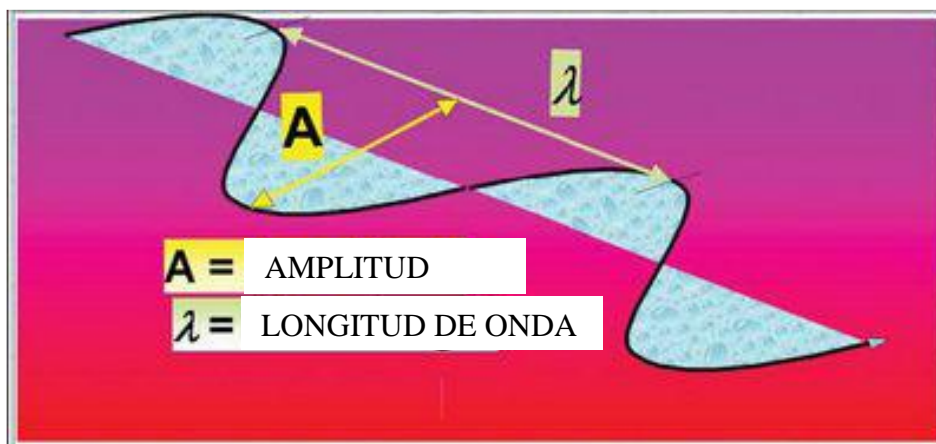


Figura 0.1 Onda típica propagándose a través del espacio [16]

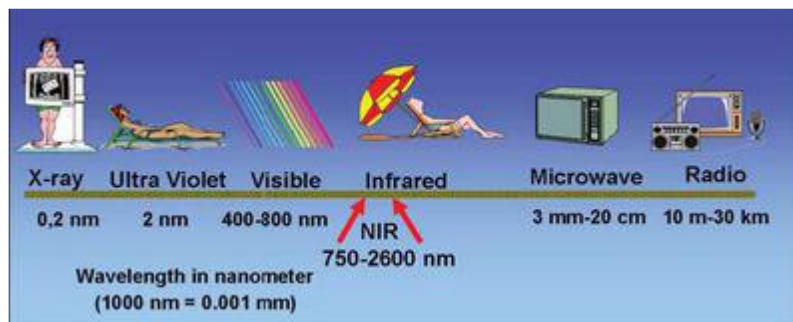


Figura 0.2 Espectro electromagnético. [16]

Aprovechando las vibraciones que se producen en los enlaces atómicos al ser éstos estimulados con rayos de la zona del infrarrojo cercano, se pueden obtener ciertos beneficios analíticos, que serán detallados en el presente documento.

En los alimentos y otras sustancias, las moléculas son grupos de átomos, que se combinan para formar compuestos químicos. Por ejemplo, el metano contiene un átomo de carbono (C) y cuatro átomos de hidrógeno (H).

Las uniones específicas entre los átomos vibran a una cierta frecuencia, y cada tipo de estas uniones químicas dentro de una muestra absorberán rayos NIR de una longitud de onda específica, mientras todas las demás longitudes de onda son reflejadas como se muestra en la figura 2.3

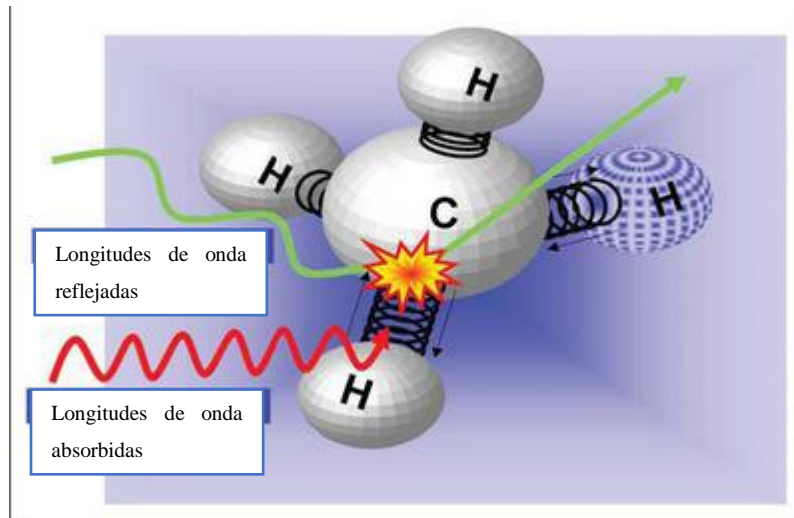


Figura 0.3 El enlace vibratorio entre los átomos de carbono (C) e hidrógeno (H) absorbe ondas NIR de una longitud de onda particular mientras otras ondas son reflejadas [17]

1.7.2 Reflectancia NIR

En la práctica, la muestra a ser analizada es bombardeada con rayos NIR de diferentes longitudes de onda como se ilustra en la figura 2.4. Por cada longitud de onda, algunos de los rayos serán entonces absorbidos por uniones químicas específicas. Al mismo tiempo, otros rayos serán diseminados y reflejados por otras uniones químicas. Este proceso es comúnmente descrito como Reflectancia NIR. Por el contrario, algunos de los rayos pasarán a través de la muestra, lo cual es denominado Transmisión NIR (a menudo referida como NIT).

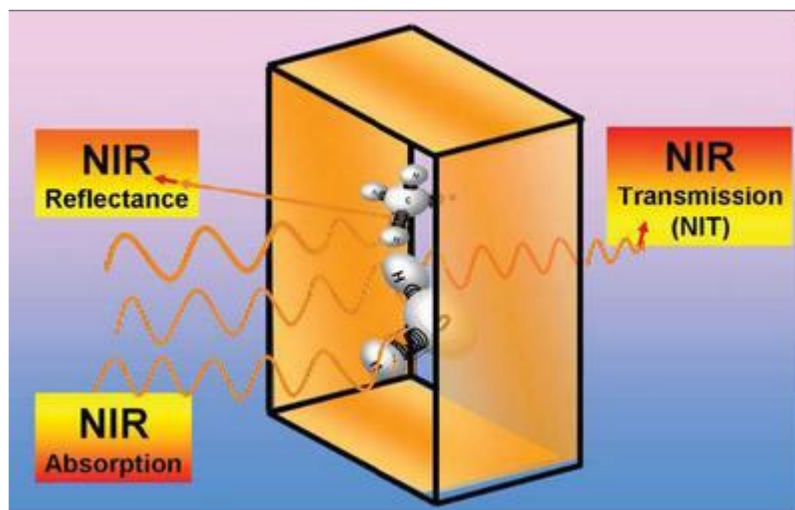


Figura 0.4 Rayos NIR absorbidos por algunos enlaces, reflejados por otros (Reflectancia NIR) o transmitidos a través de una muestra (Transmisión NIR generalmente nombrada NIT). [8]

1.7.3 Transmisión NIR

La Transmisión NIR, donde el detector está ubicado detrás de la muestra, es ideal para líquidos transparentes y algunos productos que no son demasiado densos ópticamente. Productos tales como muestras de girasol, canola y tierra son ópticamente tan densos que permiten que sólo una porción de rayos pase a través de la muestra, por lo tanto, en estas condiciones se considera que mediciones confiables son muy difíciles de obtener por NIT.

1.7.4 Teoría de espectroscopía vibracional

Interacción entre la luz y la materia física

La espectroscopía utiliza la luz, en su calidad de radiación electromagnética, para analizar materiales mediante la descripción de la energía que es transferida entre la luz y la materia [18]. La energía de un fotón, partícula fundamental de la luz, es definida como se muestra en la ecuación (1):

$$E = h \cdot \nu = \frac{h \cdot c}{\lambda} = h \cdot c \cdot \tilde{\nu} \quad (1)$$

$$\text{Con } \tilde{\nu} = \frac{c}{\lambda} \quad (2)$$

$$\tilde{\nu}[cm^{-1}] = \frac{10^7}{\lambda} \quad (3)$$

donde;

h – constante de Planck en $m^2 \cdot kg/s$

ν – frecuencia de la luz en s^{-1}

c – velocidad de la luz en m/s

λ – longitud de onda de la luz en m

$\tilde{\nu}$ – número de ondas en m^{-1}

Generalmente, los números de onda son presentados en cm^{-1} . Para convertir la longitud de onda en número de ondas, el recíproco de la longitud de onda en nm es multiplicado por 10^7 como se muestra en la ecuación (3). La ecuación (1) muestra que mientras más corta es la longitud de onda, más alto será el número de ondas y también más alta la energía del fotón.

El espectro electromagnético está dividido en varias regiones, cada una de las cuales puede servir para realizar una o más técnica de análisis espectroscópico midiendo la cantidad de energía que retorna después de haber excitado ciertos enlaces moleculares en frecuencias vibracionales.

Oscilador Armónico Simple

El oscilador armónico simple es un sistema formado por un objeto que pende de un resorte, dicho objeto produce un movimiento oscilatorio, tendiendo al equilibrio y describiendo movimientos que se repiten cada cierto período de tiempo. [19]

Oscilador Armónico como descriptor del NIR

El clásico modelo físico del oscilador armónico simple puede ser usado para describir las vibraciones moleculares como el más simple sistema vibratorio [20]. Dos masas vibratorias conectadas por un resorte con una fuerza constante dada es capaz de producir cambios en la distancia intranuclear. Mediante una combinación entre la Ley de Hook y

la Ley de Fuerza de Newton, la frecuencia vibratoria vendrá dada por la ecuación (4) como se muestra a continuación:

$$\nu_{vib} = \frac{1}{2\pi} \sqrt{\frac{k}{\mu}} \quad (4)$$

donde;

ν – frecuencia de enlace de la banda de absorción

k – constante de fuerza del enlace

μ – masa reducida de los átomos enlazados

La energía y la frecuencia de vibración son una función de la fuerza del enlace. Entonces, las frecuencias vibratorias son la clave que contienen información sobre la estructura del compuesto a investigarse. Ésta es la base para la aplicación de la espectroscopía vibracional.

De acuerdo a la Ley de Hook, la energía potencial V del oscilador armónico es una función cuadrática del desplazamiento de la vibración de los átomos. La función parabólica que se muestra en la ecuación (4) es simétrica con respecto a la longitud del enlace en equilibrio r_e :

$$V = \frac{1}{2} k(r - r_e)^2 = \frac{1}{2} kx^2 \quad (5)$$

donde;

V – energía potencial

r – distancia internuclear

r_e – distancia internuclear en equilibrio

x – desplazamiento de los átomos vibratorios

Las consideraciones de la mecánica cuántica aplicadas a los sistemas químicos moleculares: la teoría del enlace de valencia (EV) y la teoría del orbital molecular (OM) [21], que hacen uso de la ecuación de Schrödinger, muestran que la energía vibracional

para el oscilador armónico no tiene niveles energéticos continuos, sino, únicamente niveles energéticos discretos como se muestra en la ecuación (6):

$$E_{vib} = hv \left(v + \frac{1}{2} \right) \quad (6)$$

donde;

v – número cuántico vibracional (0, 1, 2, ...)

El número cuántico vibracional puede tener únicamente valores enteros. Sustituyendo la ecuación (4) en la ecuación (6) obtenemos la energía vibracional que viene dada por la ecuación (7) como se muestra a continuación:

$$E_{vib} = \left(v + \frac{1}{2} \right) \frac{h}{2\pi} \sqrt{\frac{k}{\mu}} \quad (7)$$

De acuerdo con la ecuación (7), sólo son permitidas transiciones entre niveles energéticos cercanos y equidistantes cuando se aplica el modelo oscilador armónico. La distribución de las moléculas a través de los niveles vibracionales permitidos sigue la distribución de Boltzman. Debido a esto, a temperatura ambiente la mayoría de las moléculas permanecen en el nivel vibracional $v=0$.

1.8 CALIBRACIÓN DE UN INSTRUMENTO

La calibración de un instrumento es el proceso de comparar dos valores: uno, el medido por el instrumento; y otro, una medida patrón.

Los resultados de esta comparación estable una relación entre las medidas suministradas por el instrumento y aquellas tomadas como patrón, dando de esta manera un porcentaje de error y es una característica de la precisión del instrumento de medida. [22]

1.9 CALIBRACIONES

El instrumento NIR no se “calibra” como un balance donde las lecturas son ajustadas meramente hacia arriba o abajo a un valor estándar.

El instrumento debe ser capacitado para reconocer diferentes productos y elementos. Este proceso de entrenamiento se llama procedimiento de calibración, y en este punto yace el secreto del éxito de esta revolucionaria tecnología.

Para el entrenamiento, un número de muestras son analizadas por métodos químico-analíticos tradicionales para determinar la composición real de ellas. Cada una de éstas es colocada luego en el instrumento NIR, y se obtienen los valores de reflectancia de las diferentes longitudes de onda. Con la ayuda de un microcomputador y un eficiente software químico-métrico, la combinación de los resultados analíticos y los valores de reflectancia son transformados a las constantes de calibración. Este software es tan poderoso que debe tenerse gran cuidado ya que no presenta simplemente una solución estadística, sino que realmente suministra una solución científica que puede ser verificada.

Para desarrollar una nueva calibración o aún mantener las ya existentes, es importante primeramente procurarse físicamente de un grupo ideal de muestras de cada producto. El conjunto elegido de muestras debe incluir aquéllas que representen tantas variaciones de los componentes analíticos y nutritivos con las que puedan contarse. Este grupo debe, idealmente, contener también muestras representando la variación natural que pueda darse. Esto incluye variedades de cultivos, áreas de crecimiento, condiciones de crecimiento, y temporadas de crecimiento. Dersjant-Li y Peisker [17], recientemente enfatizaron la gran variación en composición entre muestras de soya recolectadas de diferentes países o aún de distintas áreas dentro de un mismo país. Una vez procurado el grupo de muestras que cubra la mayor parte de las combinaciones, la mayoría de los programas de software de calibración tienen una herramienta, la cual entonces ayuda a seleccionar además un sub-muestreo para preparar la calibración.

Más aún, las calibraciones secundarias, realizadas con instrumentos NIR, raramente representarían un verdadero reflejo de las muestras de áreas locales, y usualmente necesitan bastante ajuste. Esto puede hacerse agregando un número de muestras cuidadosamente seleccionadas de un producto específico local, a los datos existentes de calibración.

1.10 ANÁLISIS DE COMPONENTES PRINCIPALES (PCA)

El Análisis de Componentes Principales (PCA), también llamado transformada de Karhunen-Loève, es una técnica estadística para el tratamiento de datos de forma compacta que elimina la redundancia de los datos originales. Se trata pues de obtener una mejor representación posible de los datos. Las técnicas PCA encuentran numerosas aplicaciones en diversas materias que requieren el tratamiento de datos (especialmente para la compresión). En este caso, será una herramienta especialmente útil para el procesado previo de datos y la aplicación de técnicas de Análisis de Componentes Independientes (ICA).

Dado un vector aleatorio x con n componentes, del cual se conoce una muestra $x(1), x(2), \dots, x(T)$, el objetivo de PCA consiste en tratar de eliminar la redundancia existente entre cada una de sus componentes de forma que se obtenga un nuevo vector y con m componentes, tal que $n \geq m$. Con esto se consigue un vector de menor dimensión, sin redundancia entre sus componentes.

En PCA, no se tiene en cuenta el tipo de función de densidad que poseen los vectores. Sólo se supone que se pueden estimar los estadísticos de primer y segundo orden de la muestra. PCA trata de buscar un nuevo sistema de ejes coordenados formado por una base ortogonal para representar al vector x de forma que se maximiza la varianza de las proyecciones del vector en dicho sistema. Así se conseguirá que las componentes de x en el nuevo sistema, no estén correlacionadas.

Para ilustrar el funcionamiento de PCA, en la figura 2.5 se representa la distribución de probabilidad conjunta de dos variables aleatorias x_1 y x_2 , obtenidas cada una de ellas a partir de la combinación lineal de otras dos variables aleatorias s_1 y s_2 independientes entre sí.

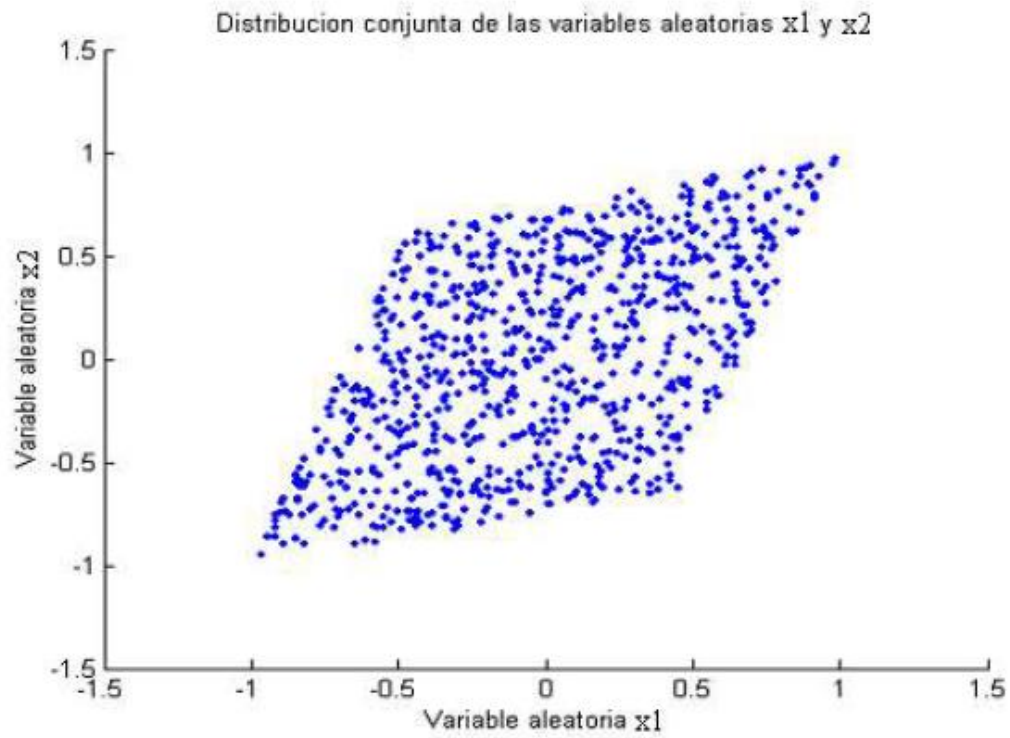


Figura 0.5 Distribución de probabilidad conjunta de dos variables aleatorias x_1 y x_2

Si a las variables aleatorias x_1 y x_2 se les aplica un procesamiento PCA, las nuevas variables que se obtienen no estarán correlacionadas entre sí y tendrán una distribución de probabilidad conjunta como se ilustra en la figura 2.6

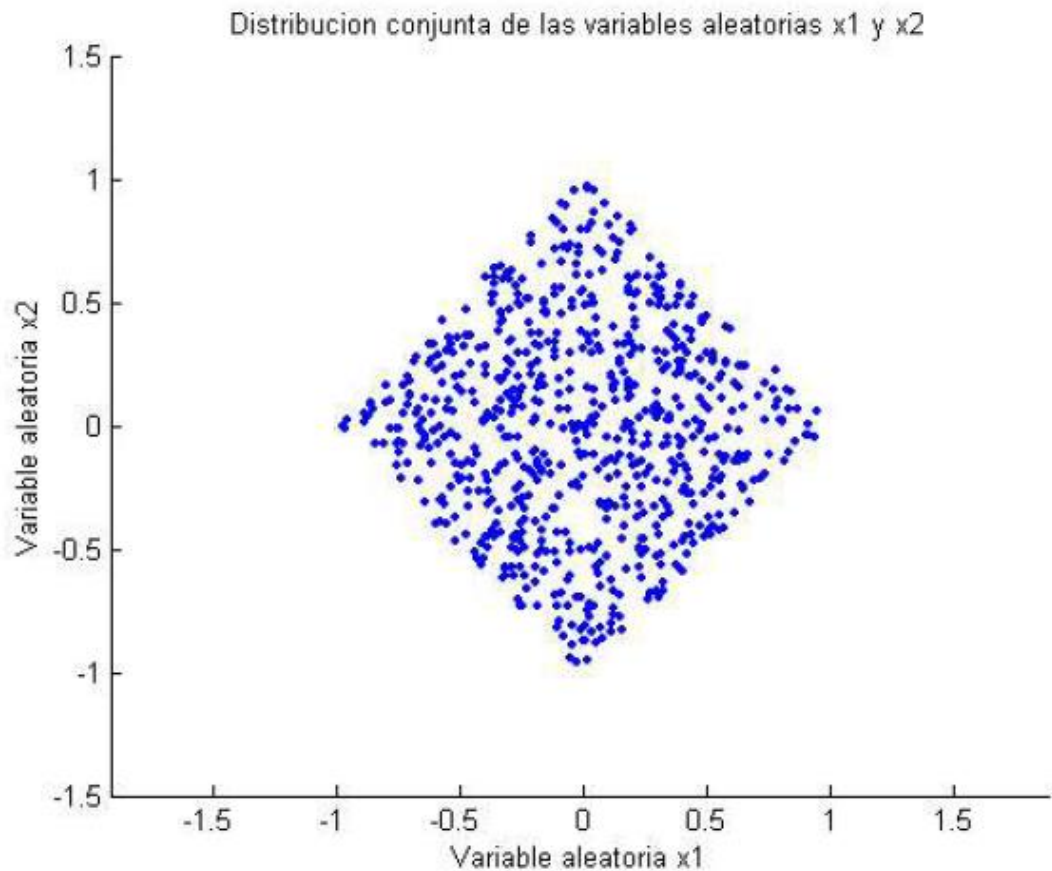


Figura 0.6 Nueva distribución de probabilidad conjunta de las variables tras un procesamiento PCA.

Tal como puede verse en la figura 2.6 se muestran las nuevas variables obtenidas, que estarán no correlacionadas entre sí y su función de distribución conjunta estará contenida en un cuadrado. Las técnicas PCA persiguen objetivos distintos que las técnicas ICA. Sin embargo, se demostrará que PCA puede ser útil como un procesamiento previo al uso de técnicas ICA. También pueden usarse para blanquear o decorrelacionar los datos antes de aplicarles un determinado procesado.

1.10.1 Descripción del método

En PCA se trata de buscar las componentes principales de las observaciones que consisten en los autovalores de la matriz de covarianzas de éstas. Para poder aplicar el método, se

debe centrar el vector x . Para ello, simplemente, se sustrae la media tal como se muestra en la ecuación (8):

$$x = x - E\{x\} \quad (8)$$

Se denota como y la nueva representación de los datos. Se trata de un vector de dimensión m . En general, se cumplirá que $m \leq n$ siendo n la dimensión de los datos originales. La nueva representación de los datos tras aplicar las técnicas PCA vendrá dada por la ecuación (9):

$$y = w^T x \quad (9)$$

Donde w es una matriz cuyas filas son vectores de dimensión n , tales que maximizan la varianza entre las componentes del vector y . Para dicha matriz sus columnas serán ortogonales entre sí, como se muestra en la ecuación (10):

$$w_i^T w_j = \delta_{ij} \quad (10)$$

Puede demostrarse, que las columnas de la matriz w coinciden con una base ortogonal expandida a partir de los autovectores e_1, e_2, \dots, e_n de la matriz de covarianzas, C_X , del vector x . Dichos autovectores serán distintos entre sí y se ordenarán en función del valor de sus autovalores $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m$. La mayor parte de la información estará contenida en el primer autovector e_1 . La cantidad de información contenida en cada autovalor será mayor cuanto mayor sea su autovalor asociado, de forma que en las primeras componentes se concentrará la mayor parte de la información. Al valor expresado en la ecuación (11) se le denomina *componente principal* de x , y es la que contiene más información de todas.

$$y_1 = e_1^T x \quad (11)$$

En el caso de que $m < n$, se estaría cometiendo un error en la representación de los datos. Como ventaja se comprime la información contenida en el vector inicial.

El error cuadrático mínimo que se puede cometer se dará cuando se tome una base ortogonal expandida a partir de los m autovectores e_1, e_2, \dots, e_m de la matriz de covarianzas.

Dicho error vendrá dado por la ecuación (12):

$$J_{MSE}^{PCA} = \sum_{i=m+1}^n d_i \tag{12}$$

Si $m = n$ no se cometerá error alguno en la representación de los datos en la nueva base formada por los autovectores de la matriz de covarianzas de los datos. Este caso será de especial interés, ya que al aplicar un análisis PCA a los datos, se conseguirá decorrelacionarlos.

1.11 REGRESIÓN POR MÍNIMOS CUADRADOS PARCIALES

Históricamente, el primer tipo de algoritmo PLS es un método alternativo para el cálculo de componentes principales. Este procedimiento se extendió casi inmediatamente a una serie de procedimientos entre los que había una versión para el cálculo de correlaciones canónicas.

Bajo el nombre de Niles, abreviatura de “no lineal por mínimos cuadrados iterativos” ,Wold, en el año 1966, presenta un resumen de ejemplos resueltos mediante procedimientos iterativos basados en modelos de regresiones de mínimos cuadrados. Curiosamente, estos trabajos iniciales contenían los elementos matemáticos fundamentales de todos los métodos PLS siguientes: cálculo de las componentes de datos como sumas ponderadas de variables, obtenidos operacionalmente a través de los pasos de regresiones de mínimos cuadrados. No mucho tiempo después de su presentación, se sustituyó el término “Niles” por “NIPALS” (no lineal Iterative Partial Least Squares), en consecuencia, se cambiará procedimientos Niles por procedimientos NIPALS [23].

En PLS, a diferencia de Componentes Principales, los datos de entrada además de la matriz de variables predictoras X de orden $n \times p$ (donde n representa al número de muestras y p al número de variables), deben contener una matriz de variables dependientes Y de orden $n \times q$ (donde n representa al número de muestras y q al número de variables) con el fin de predecir los valores de Y a partir de X para nuevas observaciones.

La regresión con proyección consiste en encontrar una nueva matriz de variables, reduciendo la dimensión de las variables originales, ésta se llama matriz de variables latentes. Esta nueva matriz tendrá un máximo de varianza para que todas las variables tengan la misma importancia en el sistema, es decir, que tenga una media de cero y una desviación estándar de uno. [24]

1.12 TARJETAS PARA ADQUISICIÓN DE DATOS

Las tarjetas de prototipado actuales en el mercado tales como la Raspberry o la tarjeta Arduino entre las más sobresalientes, más allá de ser compactas ofrecen sin duda grandes prestaciones para proyectos de desarrollo en el área de control, automatización y ejecución de programas para diversos diseños electrónicos. El uso de estos dispositivos como base para la adquisición de señales, diseño y programación, ofrece una amplia gama de posibilidades, claro está, considerando los requerimientos que se presenten en cada proyecto.

Existen diferentes modelos de tarjetas que presentan diversas características, la potencia de cada una de ellas radica en la configuración interna que poseen en cuanto a la capacidad de procesamiento, memoria y la flexibilidad de adaptación a sistemas externos, lo que señala la pauta para la elección entre una específica que ofrezca las funciones que mejor se adapten a la necesidad de diseño.

Tabla 0.1 Tipos de tarjetas de adquisición

Tipos	Características	Aplicaciones
Plug-in	<p>Poseen alta velocidad que va desde los 100 KHz a 1 GHz.</p> <p>Amplio abanico de funciones: A/D, D/A. E/S digitales. Contadores, temporizadores.</p>	<p>Son diseñadas para ser insertadas en los slots internos de una computadora.</p> <p>Actualmente se usan para el bus PCI, aunque existen versiones como: EISA, IBM Micro Channel, etc.</p>
Sistemas externos de adquisición de datos	<p>Poseen mayor velocidad y versatilidad, ya que cuentan con un entorno más protegido del ruido.</p> <p>Cuentan con las arquitecturas:</p> <ul style="list-style-type: none"> • PXI (PCI extensión for Instrumentation) • VISA (Virtual Instrumentation Software Architecture) • VXI (VME extension for Instrumentation) 	<p>Actualmente son equipos autónomos orientados a procesos de control en tiempo real, en proyectos con alta sensibilidad para bajos niveles de tensión en las entradas.</p>
Sistemas híbridos de adquisición de datos	<p>Utilizan una combinación de tarjetas de adquisición de datos e instrumentación específica de test</p>	<p>Manejo de varias señales simultáneamente.</p> <p>Sistemas donde se requieran módulos específicos para análisis especiales</p>

Fuente: Investigador [16]

La adquisición de datos es un requisito fundamental para procesar correctamente las señales eléctricas del sensor y de esta forma realizar un mejor tratamiento de los datos y tener salidas adecuadamente ajustadas.

Entre algunas de las placas de desarrollo para proyectos relacionados con aplicaciones relacionadas generalmente con IoT, robótica, dispositivos portátiles, entre otras, que presentan altas prestaciones en cuanto a la capacidad de sus microcontroladores son:

1.12.1 Arduino Tian

Basado en la filosofía de Arduino, pero considerando que más bien es una placa tipo SBC combinada con las propiedades de una placa Arduino, esta placa esta creada para el desarrollo de aplicaciones tipo IoT. Puede correr un sistema operativo GNU/Linux, además, posee un módulo Wi-Fi que facilita enormemente el trabajo junto con su procesador ARM Cortex M0 de 32 bits y un Qualcomm Atheros AR9342, acompañados de 64MB de RAM y 4 GB de memoria tipo flash.



Figura 0.7 Arduino Tian [14]

1.12.2 Raspberry Pi 3

La Raspberry Pi 3 posee un procesador Quad Core 1.2GHz Broadcom BCM2837 de 64 bits, con 1 GB de RAM, Bluetooth v4.1 de bajo consumo incorporado, además de un adaptador LAN inalámbrico BCM43438. Posee un puerto de conexión ethernet RJ45.

Los 4 puertos USB 2.0 permiten adherir dispositivos para un mejor trabajo. El conector GPIO de 40 pines permite la realización de experimentos electrónicos.

Dispone de salida de 4 polos para audio estéreo y video compuesto, 1 salida HDMI, puerto CSI para cámara, puerto DSI para pantalla táctil, puerto micro SD y un conector micro USB para alimentación. [25]



Figura 0.8 Raspberry Pi 3 [15]

1.12.3 BeagleBoard

Este modelo producido por la compañía Texas Instruments en asociación con DigiKey se presenta como la competencia para Raspberry, debido a su arquitectura que se compone de un SoC OMAP3530, basado en ARM Cortex A8, con un DSP TMS320C64x y una GPU PowerVR SGX530 de Imagination Technologies. Esta placa incluye además una memoria RAM de 256 MB, ranura SD/MMC, USB, RS-232 y jacks. [26]



Figura 0.9 BeagleBoard [10]

1.12.4 ASUS Tinkerboard

Se trata de una placa que ha conseguido superar las prestaciones de la Raspberry Pi 3 en los benchmarks a los que se ha sometido, en gran parte a su chip RockChip RK3288 Quad-Core de 1.8 GHz. Además de poseer gran capacidad de memoria RAM con 2GB, ofrece una tarjeta de red con Gigabit LAN y conexión Bluetooth 4.0 +EDR y soporte de 192K/24 bits para audio entre sus características más relevantes. [26] [27]



Figura 0.10 ASUS Tinker Board [11]

1.12.5 UDOO

Esta potente placa desarrollada como proyecto de la plataforma crowdfunding Kickstarter, es el resultado de sumar 4 Raspberries Pis y una placa Arduino Due, Se trata definitivamente de una de las opciones más complejas y de mayor rendimiento en comparación con las demás tarjetas con las que compite, principalmente gracias a su procesador i.MX6 de Freescale, basado en ARM Cortex-A9 quadcore de 1 GHz, GPU PowerVR, 1GB de RAM DDR3, capaz de manejar un sistema operativo Android o distribuciones Linux.

Por otro lado, incorpora un microcontrolador Atmel SAM3X8E ARM Cortex-M3 como el Arduino Due, tiene 76 GPIOs, HDMI y LDVS + Touch, Ethernet, Wi-Fi integrado, miniUSB, conexiones para audio, SATA, conexión para cámara, micro SD y alimentación. [26]



Figura 0.11 UDOO Quad/Dual [16]

1.13 LENGUAJES DE PROGRAMACIÓN PARA ANÁLISIS DE DATOS

El análisis de grandes cantidades de datos requiere del conocimiento y dominio de lenguajes de programación. Existe en el mundo del analista de datos, que es la persona que se desempeña en este campo, una gran cantidad de lenguajes dedicados a esta tarea. En la presente sección se tratarán algunos lenguajes dedicados a esta tarea.

1.13.1 Lenguaje de programación R

R es un entorno y un lenguaje de programación enfocado en el análisis estadístico de los más utilizados en el campo de la minería de datos que pueden aplicarse a gran variedad de disciplinas.

Este lenguaje forma parte de GNU y se distribuye bajo licencia GNU GPL.

Características y ventajas de la programación en R

El lenguaje de programación R es un proyecto colaborativo y abierto, los desarrolladores pueden descargar el código de forma gratuita y modificarlo para incluir mejoras.

Es un lenguaje interpretado, es decir, funciona mediante comandos.

R proporciona una amplia gama de herramientas estadísticas que incluyen análisis de datos y generación de gráficos. Este lenguaje tiene capacidad de generar gráficos de alta calidad. Estas características lo convierten en una potente herramienta de cálculo.

Gracias a este lenguaje de programación los científicos de datos pueden manejar grandes volúmenes de datos.

Puede integrarse con distintas bases de datos.

Una de las ventajas más importantes de R es que funciona con diferentes tipos de hardware y software (Windows, Unix, Linux...)

El lenguaje R ofrece la posibilidad de cargar bibliotecas y paquetes con diversas funcionalidades lo que permite a los usuarios extender su configuración básica.

La comunidad en torno a R es muy activa por lo que es sencillo encontrar soluciones rápidamente a los problemas que los usuarios se puedan encontrar. [28]

1.13.2 Lenguaje de programación PYTHON

Python es un lenguaje que se usa en varias áreas de tecnología: web, redes, procesamiento de datos, inteligencia artificial, etc. y ha sido desarrollado teniendo en mente a los programadores principiantes pues es muy sencillo familiarizarse con la sintaxis del lenguaje. Gracias al uso de expresiones comunes, Python requiere menos líneas de código para realizar tareas básicas. En promedio, el código escrito en Python es de tres a cinco veces más corto que Java y cinco a 10 veces más corto que C++.

Adicionalmente, Python tiene una librería estándar que permite ejecutar otras funciones y tareas más complejas con mayor facilidad que otros lenguajes. Para desarrollo en últimas tendencias como machine learning, Python es el lenguaje más popular del mundo. Python es un lenguaje perfecto para cualquier programador principiante que esté decidido a aprender por cuenta propia por la facilidad de su sintaxis. [29]

Características y ventajas de la programación en PYTHON

Python es un lenguaje de programación interpretado, por lo que funciona en cualquier tipo de sistema que integre su interpretador. A parte de esta ventaja, Python ofrece dialectos como el ya conocido Jython, que se utiliza para escribir en Java.

Python no sólo es multiplataforma y multiparadigma, sino que también sirve para desarrollar cualquier tipo de vía, como por ejemplo web o móvil. Para que esto se lleve a cabo, este lenguaje de programación cuenta con frameworks de gran calibre, los cuales

auxilian desde el desarrollo web, hasta el desarrollo de juegos o algoritmos científicos de cálculos avanzados.

Si se habla de la licencia que posee, ésta es Python Software Foundation License, licencia muy parecida a la de GPL, pero encontrando la excepción de que se pueden distribuir los binarios del lenguaje sin tener que anexar las fuentes.

Python se encuentra en multitud de aplicaciones y servicios que son usados habitualmente. Ostenta una gran lista de usuarios como Google, YouTube o Facebook, los cuales utilizan este lenguaje de programación. Poco a poco Python ha ganado territorio y se ha convertido en uno de los lenguajes más solicitados y, sobretodo, más esenciales del momento. Esto demuestra que programar en Python es la opción más viable y efectiva que hay ahora mismo en el mercado.

Si se habla de programación orientada a objetos, se puede decir que el usuario se encontrará ante un paradigma que propone modelar todo en función a clases y a objetos, el cual nos ofrece un uso de conceptos de cohesión, polimorfismo, herencia, abstracción y mucho más.

Este paradigma de programación se utiliza para tratar el rápido aumento en el tamaño y la complejidad de los sistemas de software, y facilitar la modificación de esos grandes y complicados sistemas a lo largo del tiempo. [30]

1.13.3 XLSTAT para Microsoft Excel

XLSTAT es el programa líder en análisis de datos y soluciones estadísticas para Microsoft Excel.

El complemento XLSTAT de análisis estadístico ofrece una amplia variedad de funciones para mejorar las capacidades de análisis de Excel, por lo que es la herramienta ideal para sus necesidades diarias estadísticas y de análisis de datos.

El software de análisis estadístico de XLSTAT es compatible con todas las versiones de Excel desde la versión 2003 a la versión 2016 (versión 2011, 2016 para Mac), y es compatible con los sistemas Windows Vista hasta Windows 10, así como con los sistemas Mac basados en Intel y PowerPC.

Debido a que es rápido, fiable, fácil de instalar y de usar, XLSTAT ha crecido hasta convertirse en uno de los paquetes de software estadísticos más utilizados en el mercado. Hoy en día, XLSTAT es utilizado por más de 30.000 clientes, empresas y universidades, grandes y pequeños, en más de 200 países de todo el mundo.

1.14 PROPUESTA DE SOLUCIÓN

La implementación de un sistema basado en espectroscopía de infrarrojo cercano, usando análisis de señales, y software-hardware libre, permitirá obtener a bajo costo, características y cualidades únicas en alimentos sólidos. Este tipo de equipo puede estar al alcance de las personas que se dedican a la elaboración de alimentos, así como de pequeños laboratorios, reduciendo significativamente los tiempos y los costos de loa análisis.

METODOLOGÍA

1.15 MODALIDAD DE LA INVESTIGACIÓN

En el presente trabajo de investigación, se aplicaron las siguientes modalidades:

Investigación Aplicada

En la ejecución de la implementación del prototipo de análisis de alimentos basado en espectroscopía de infrarrojo cercano se hizo uso de investigación aplicada para verificar cuáles son los requerimientos para un sistema que sea capaz de analizar la cantidad de proteínas, grasas y humedad en alimentos, con el propósito de obtener datos y generar reportes que satisfagan los objetivos de la investigación.

Investigación Bibliográfica

La investigación tuvo modalidad bibliográfica al hacer una búsqueda de las alternativas para desarrollar el modelo óptimo del prototipo para realizar la medición de parámetros propuestos, basado en componentes de bajo costo, utilizando además software libre, mediante el uso de artículos científicos, libros técnicos, entre otros.

Para este tipo de investigación se recurrirá, principalmente, al internet para obtener la información, los conocimientos y las técnicas actualizadas sobre el tema desarrollado; también se recurrirá a la biblioteca para hacer usos de los libros más relevantes que aporten a la investigación.

Investigación Experimental

Esta investigación se realizó desarrollando ensayos de prueba y error, utilizando los estándares que definen las líneas guías para aplicación de la espectroscopía de infrarrojo cercano. Es decir, hubo una serie de pruebas metódicas que llevaron a la culminación satisfactoria de la implementación del prototipo.

1.16 PROCESAMIENTO Y ANÁLISIS DE DATOS

Una vez realizada la recolección de datos a través de la investigación de campo y experimental se realiza una fase esencial para toda investigación, que es la clasificación de los datos referentes a cada variable del proyecto y su presentación conjunta. Para esto se sigue un procedimiento de cuatro pasos para su análisis.

- Validación y Edición
- Codificación
- Introducción de datos
- Tabulación y análisis estadísticos.

Después de aplicar correctamente los instrumentos se procedió a establecer la validez de los datos mediante gráficos estadísticos apropiados los cuáles mostrarán las características en cuanto a los parámetros que se deben tomar en consideración junto con las variables de la investigación. Al determinar los requerimientos específicos del sistema de análisis de alimentos sólidos, usando el prototipo y con base en el marco teórico, se ejecutará un análisis integral de toda la información, enfocando las soluciones conforme a los objetivos planteados.

Como siguiente paso se realizará un estudio de las ventajas y desventajas que se presentan con respecto al desarrollo de este proyecto, considerando cada uno de los datos técnicos obtenidos de libros, revistas y artículos científicos, de este modo se buscará proponer la mejor solución para el proyecto.

1.17 DESARROLLO DEL PROYECTO

Para llevar a cabo el presente proyecto se realizaron las siguientes actividades:

1. Estudiar del espectro electromagnético para tener presente la banda de frecuencia de los rayos infrarrojos y definir cuál es la diferencia entre infrarrojo cercano e infrarrojo lejano.
2. Analizar la interacción entre los alimentos y sus componentes y las longitudes de onda que se encuentran en el espectro del infrarrojo cercano.
3. Realizar una búsqueda de los estándares que rigen la técnica de espectroscopía de infrarrojo cercano para alimentos.
4. Seleccionar elementos electrónicos y ópticos que permitan hacer una espectroscopía de infrarrojo cercano, dichos elementos deben ser hardware libre.
5. Diseñar un algoritmo para el análisis de las señales obtenidas de los diferentes dispositivos electrónicos y ópticos.
6. Realizar un análisis matemático de las señales, y que éste se ajuste a las normativas vigentes para el uso de NIR en alimentos.
7. Realizar mediciones que permitan cotejar los datos, y de ser necesario, realizar un reajuste del algoritmo.
8. Diseñar una interfaz de usuario que sea de fácil manejo para éste, con software libre.
9. Diseñar la carcasa que llevará la electrónica del dispositivo.
10. Implementar el prototipo y verificar la operación de éste, realizando análisis con espectroscopía.
11. Elaborar el informe del proyecto.

DESARROLLO DE LA PROPUESTA

1.18 INTRODUCCIÓN

El presente proyecto se ha desarrollado con el propósito de realizar un análisis del contenido de diferentes compuestos (proteínas, grasas, humedad) en alimentos sólidos, por medio de un sensor espectral de infrarrojo cercano, método que es no destructivo, rápido, confiable y de bajo costo, a diferencia de los métodos tradicionales.

Con lo expuesto en el párrafo anterior, no hizo falta el uso de ningún reactivo, como en los métodos tradicionales de bromatología, una parte que se quería resolver.

Para obtener los valores porcentuales de los compuestos presentes en las muestras se utilizan métodos estadísticos sugeridos en la norma ISO 12009:2017 que da lineamientos para la determinación, mediante espectroscopía de infrarrojo cercano, de los constituyentes en los alimentos.

1.19 ANÁLISIS DE FACTIBILIDAD

1.19.1 Factibilidad Técnica

Los elementos electrónicos que se utilizarán en el proyecto son de fácil adquisición, pues se encuentran a la venta en el mercado, lo que hace que su desarrollo tenga la factibilidad técnica requerida.

1.19.2 Factibilidad Económica

El investigador financiará, con sus recursos, el costo del prototipo y todos los dispositivos electrónicos y demás implementos necesarios para la consecución del proyecto, motivo de investigación; por lo tanto, es económicamente viable.

1.19.3 Factibilidad Bibliográfica

Las fuentes de información para el desarrollo del proyecto se encuentran disponibles en documentos de internet y artículos académicos relacionados con el tema de investigación.

Además se recurrió al apoyo de libros que proporcionaron una orientación sobre análisis quimiométricos y la especificación de los requerimientos que debe tener un sistema que realice mediciones usando la tecnología de espectroscopía.

1.20 SISTEMA NIR

Como se ha explicado en el marco teórico, un espectrómetro NIR usa la reflectancia de las ondas electromagnéticas en el espectro del infrarrojo cercano. Es una técnica no destructiva y rápida cuya aplicación está aumentando rápidamente en lo que se refiere a la evaluación de la calidad de los productos alimenticios. El instrumento está basado en la medición de la reflectancia difusa cubriendo la región de las longitudes de onda comprendidas entre los 610 nm hasta los 860 nm.

1.20.1 Región espectral NIR

La región espectral NIR abarca las longitudes de onda desde los 600 nm hasta los 2500 nm, es decir, desde la zona de terminación de la luz visible hasta el principio de la región espectral fundamental del infrarrojo (IR). [20]

Las absorciones más captables que se producen en las bandas de absorción de la región del infrarrojo cercano están relacionadas con los armónicos y las combinaciones de las bandas de la vibración molecular fundamental de los grupos funcionales C-H, N-H, O-H y S-H, que se los puede observar en la zona espectral del IR medio, como se muestra en

la figura 4.1 y en la tabla 4.1. Así, la mayoría de las especies químicas y bioquímicas muestran una única banda de absorción en la región espectral del infrarrojo cercano que pueden ser usadas para análisis tanto cuantitativos como cualitativos. [20]

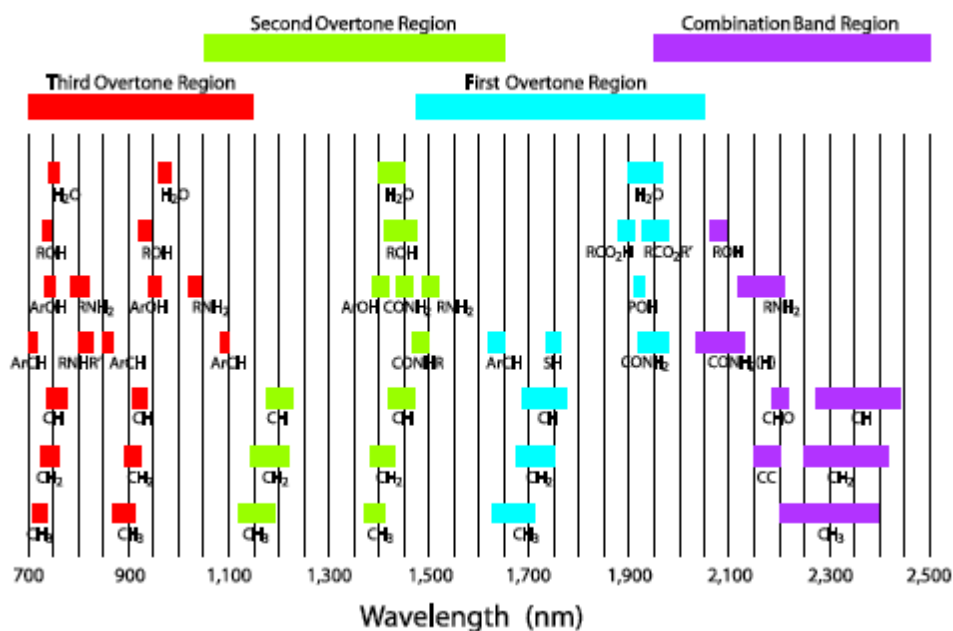


Figura 0.1 Picos en las bandas de absorción [20]

Tabla 0.1 Longitudes de onda en nm y número de ondas en algunas bandas de infrarrojo cercano de compuestos orgánicos y compuestos sulfúricos [20]

Wavelength [nm]	Wavenumber [cm ⁻¹]	Assignment
2500	4000	combination S-H stretching
2200–2460	4545–4065	combination C-H stretching
2000–2200	5000–4545	combination N-H stretching; combination O-H stretching
1620–1800	6173–5556	first overtone C-H stretching
1400–1600	7143–6250	first overtone N-H stretching; first overtone O-H stretching
1300–1420	7692–7042	combination C-H stretching
1100–1225	9091–8163	second overtone C-H stretching
1020–1060	9804–9434	combination S=O stretching
950–1100	10526–9091	second overtone N-H stretching; second overtone O-H stretching
850–950	11765–10526	third overtone C-H stretching
775–850	12903–11765	third overtone N-H stretching
600–700	16667–14286	combination C-S stretching
450–550	22222–18182	combination S-S stretching

Las bandas de absorción en el espectro NIR son típicamente entre 10 a 100 veces más débiles que la correspondiente fundamental en el IR medio. Esta naturaleza débil de la absorción NIR permite ventajosamente un análisis directo de las bandas absorbentes más fuertes. En la figura 4.2 se muestran las curvas de medición de tres muestras.

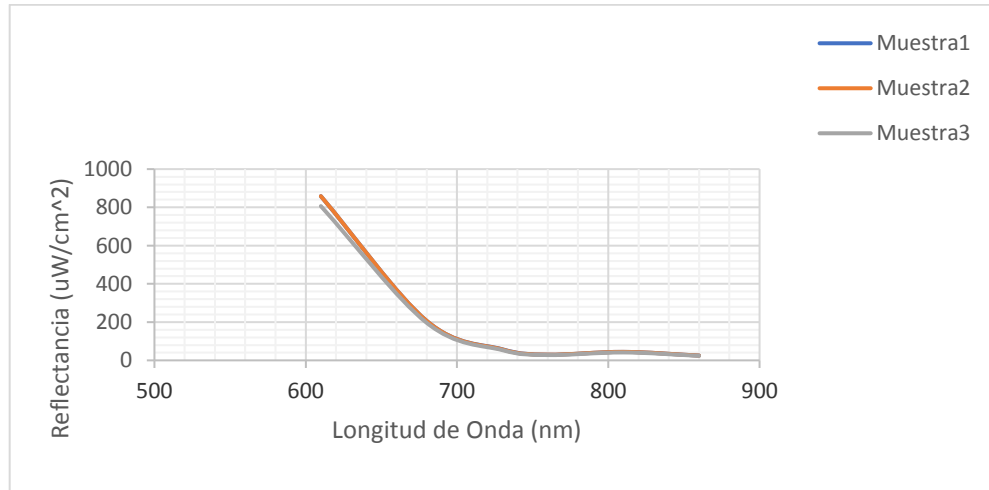


Figura 0.2 Curvas de observación del prototipo NIR

Fuente: Investigador

1.20.2 Métodos analíticos NIR

Los instrumentos NIR deben ser calibrados con métodos estadísticos de análisis para procurar que las muestras medidas arrojen valores satisfactorios, para esto se han utilizado en este prototipo dos métodos sugeridos en la norma ISO 12009:2017, que serán explicados a continuación.

Análisis de los componentes principales (PCA)

Es una forma de compresión de datos, para un conjunto de muestras. Este método funciona únicamente con los datos espectrales para hallar los componentes principales (factores) de acuerdo con una regla que expresa que cada componente principal (PC) la máxima variación en los datos a cada momento en que alguno no esté relacionado con otro PC.

El primer componente aporta una medida de la variabilidad en los datos originales, su efecto es luego retirado de la variable medida y así se van consiguiendo más PC, sin embargo, los primeros componentes principales son los que tendrán la mayor información por lo que la cantidad de datos a ser considerada es mucho menor a los datos originales [31]. En la figura 4.3 se muestra un ejemplo de PCA realizado utilizando la herramienta XLSTAT.

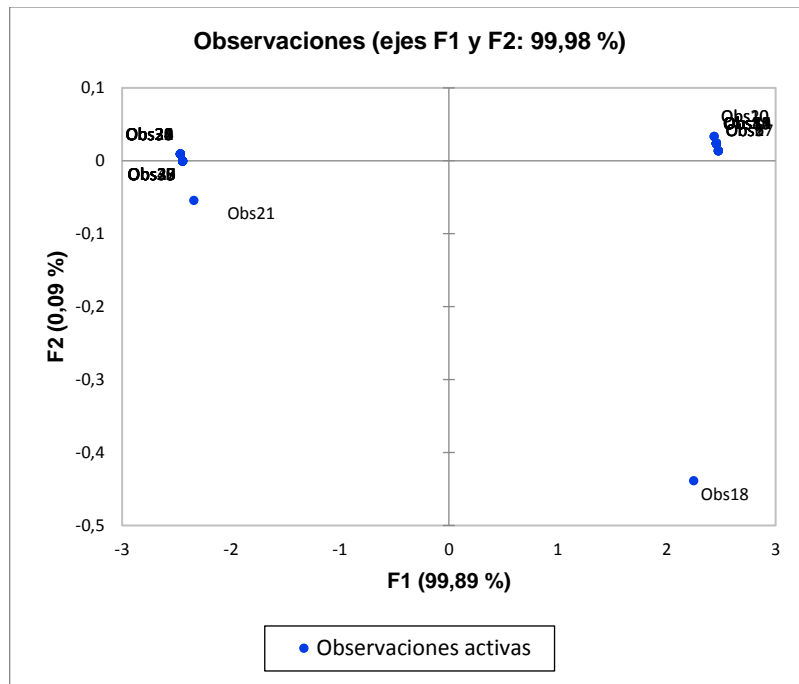


Figura 0.3 PCA de las muestras

Fuente: Investigador

Mínimos parciales cuadrados (PLS)

La regresión por mínimos cuadrados parciales es una técnica analítica en la cual se comprimen los datos. La regla para PLS es que cada factor que se usa maximiza la covarianza entre el valor de los datos predichos y las posibles combinaciones lineales de los datos de referencia [31]. En la figura 4.4 se muestra el análisis PLS para una longitud de onda del modelo.

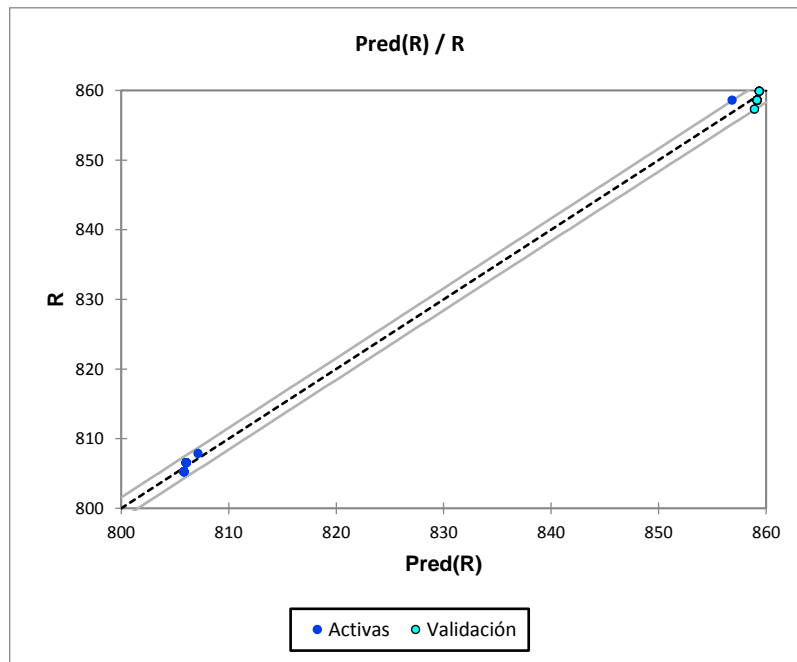


Figura 0.4 PLS del canal R

Fuente: Investigador

1.20.3 Estándares

La espectroscopía de infrarrojo cercano es cada vez más utilizada alrededor del mundo; cientos de millones de pruebas se llevan a cabo cada día por lo que las instituciones encargadas de los estándares a nivel mundial se preocupan de realizar documentos técnicos que faciliten y den soluciones para la traducción de los resultados obtenidos en formatos ampliamente aceptados.

Mientras aparezcan mayores estándares que rijan los sistemas NIR mayores serán también las herramientas de diagnóstico para aplicar modelos predictivos. En la tabla 4.2 se muestra un resumen de algunos de los estándares conocidos.

Tabla 0.2 Selección de estándares internacionales que usan NIR [20]

Estándar	Título	Industria
ASTM D 6122	Práctica estándar para la validación del rendimiento de sistemas analizadores multivariantes en línea, en línea y basados en espectrofotómetro de infrarrojo	Instrumentación
ASTM E 1655	Prácticas estándar para análisis infrarrojo cuantitativo multivariante	General
ISO 11151	Láser y equipos relacionados con el láser: óptica estándar componentes - parte 1: componentes para el UV, visible y rangos espectrales cercanos al infrarrojo	Instrumentación
ISO/TS 10868	Nanotecnologías: caracterización de nanotubos de carbono de una sola pared que utilizan espectroscopía de absorción ultravioleta-visible-infrarrojo cercano - visible (UV-Vis-NIR)	Nanotecnología
ISO 12099:2017	Alimentos para animales, cereales y productos de cereales molidos - Directrices para la aplicación de espectrometría infrarroja cercana	Alimentos

Fuente: Investigador[13]

En este proyecto se ha seguido los lineamientos de la normativa ISO 120009:2017 que rige el uso de espectroscopía por reflectancia de infrarrojo cercano en el análisis de alimentos, además se eligió esta norma para futuros trabajos relacionados con alimentos en la Facultad de Ingeniería en Alimentos.

1.21 DISEÑO DEL PROTOTIPO NIR

El proyecto se desarrolló en tres fases, y en cada una de ellas se determinaron los requerimientos técnicos. En la primera fase, la muestra de los alimentos es colocada debajo del sensor, éste emite un rayo de luz hacia el alimento y recibe y analiza la reflectancia emitida por el alimento. Así mismo el micro controlador realiza la lectura del sensor de infrarrojo cercano.

En la segunda fase, los datos enviados por la placa controladora Arduino Uno son recibidos por la placa Raspberry PI y almacenados en una base de datos montada en la misma placa.

En la tercera fase, la aplicación creada por el investigador, extrae los parámetros de los canales almacenados de la base de datos montada en la placa y los analiza. Este programa se encarga de obtener los porcentajes de proteína, humedad y grasa presentes en la muestra de los alimentos.

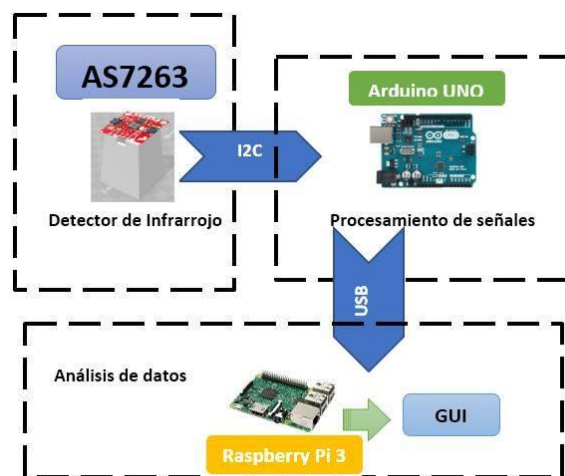


Figura 0.5 Diagrama de Bloques del Prototipo NIR

Fuente: Investigador

1.21.1 Sensor de Infrarrojo Cercano

El sensor infrarrojo es la parte más importante del proyecto ya que de éste dependerá el resto de etapas.

De la revisión en sitios especializados en la venta de este tipo de productos se observó que el costo aumentaba conforme el rango de detección tendía más hacia la frontera del infrarrojo medio. Como este proyecto pretende ser de bajo costo, se decidió por un circuito integrado de nuevo lanzamiento que fue el que menor costo representaba.

El circuito integrado escogido es el AS7263 de la compañía ams, que representa una opción de bajo costo. En la tabla 4.3 se muestran los datos del modelo de cámara o sensor NIR, con las longitudes de onda en las que trabajan y con su respectivo precio.

Tabla 0.3 Comparación de sensores NIR

Parte	Longitud de onda(nm)	Precio
1500-1600nm NIR CCD USB 2.0 Camera	1500-1600	\$3190.42
1500-1600nm NIR CCD USB 2.0 Camera (Recertified 05-P)	1500-1600	\$2552.34
1/2" CCD NIR (1460-1600nm) Analog Camera, EIA	1410-1600	\$2295.00
SparkFun Spectral Sensor Breakout - AS7263 NIR (Qwiic)	600-860	\$24.95

Fuente: Investigador [13]

Como se observa en la tabla anterior, los costos de las cámaras son mucho mayores que el costo del sensor, pero como se explicó anteriormente los armónicos que más información aportan son los que están en las longitudes más altas por lo que este dispositivo diseñado tiende a dar mayores complejidades en la detección de los

compuestos que se quiere analizar. Al trabajar con un rango muy estrecho en la banda del infrarrojo cercano las complicaciones que se pueden presentar son mayores, pero mucha fidelidad en la confianza de los datos al escoger el sensor AS7263, se puede ver mejorada por el desarrollo de un algoritmo que sustente las carencias del circuito integrado ya que el objetivo de esta investigación es el de producir un NIR a bajo costo.

Sensor AS7263

El AS7263 es un espectrómetro digital de 6 canales para la identificación espectral en el IR cercano (NIR). Está constituido por 6 filtros ópticos independientes para los cuales cada respuesta espectral está definida en las longitudes del espectro de infrarrojo cercano que oscila entre los 600 nm hasta los 870 nm, con una anchura a media altura (full-width half-max FWHM) de 20 nm [32]. Tiene un driver LED que permite programar la corriente para realizar aplicaciones de un obturador electrónico.

Además, el AS7263 está provisto de filtros gaussianos en el estándar CMOS y cuenta con una ranura que permite controlar la entrada de luz al arreglo del sensor.

El acceso a los datos espectrales y al control del chip se consigue por medio del uso del protocolo de comunicaciones I2C.

Las aplicaciones del AS7263 incluyen:

- Autenticación de productos.
- Validación de documentos y notas bancarias
- Análisis químico.
- Seguridad de alimentos y bebidas.
- En la figura 4.6 se muestra el diagrama de bloques para el funcionamiento del chip para la identificación espectral.

Este sensor espectral de ams AG presenta a los usuarios un campo de estudio que está despegando y es la espectroscopía. Ahora es más sencillo medir y caracterizar cuánto difiere de un material a otro la absorción y la reflectancia de las diferentes longitudes de onda del espectro infrarrojo cercano.

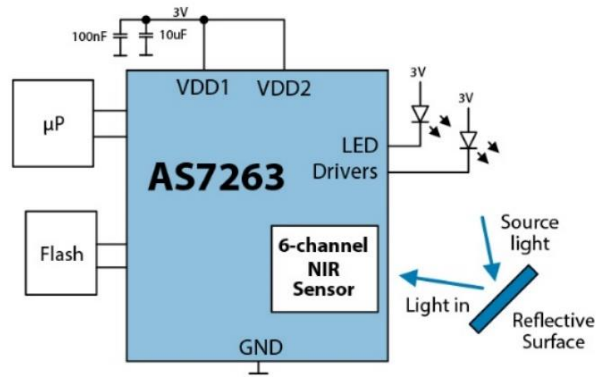


Figura 0.6 Diagrama de bloques del chip [32]

En la figura 4.7 se muestra la placa usada para la lectura de la reflectancia NIR.

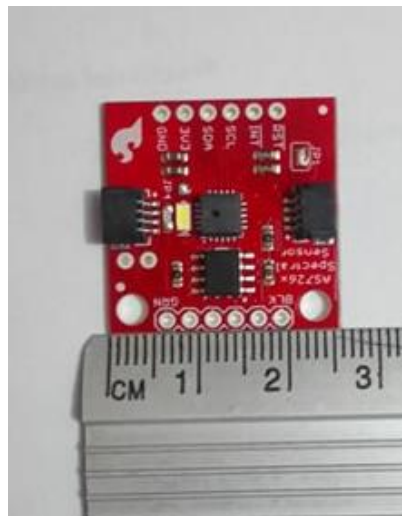


Figura 0.7 Placa con AS7263 de Sparkfun

Fuente: Investigador

La placa SparkFun Spectral Sensor Breakout - AS7263 NIR (Qwiic), facilita la intercomunicación con la placa Arduino mediante el protocolo de comunicación I2C.

1.21.2 Placa Qwicc para Arduino

La placa Qwicc tiene 4 puertos de comunicación todos ellos integrados en el mismo bus I2C. La conversión a niveles lógicos para las líneas SDA y SCL está dada por la misma placa, por lo que el usuario no debe preocuparse si se trabaja con dispositivos de 5V o de 3.3 V. [33]

Además, la creación de prototipos es más fácil ya que la placa proporciona un área exclusiva para la conexión de diferentes dispositivos. En la figura 4.8 se muestra la placa.

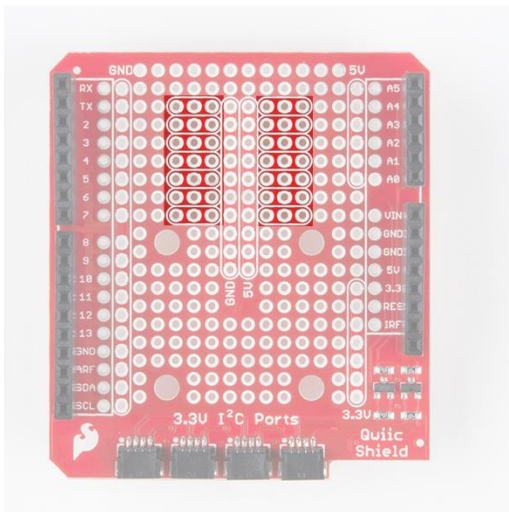


Figura 0.8 Placa Qwicc para Arduino [33]

1.21.3 Protocolo I2C

I2C significa Circuito Interintegrado (Por sus siglas en Inglés Inter-Integrated Circuit), es un protocolo de comunicación serial desarrollado por Phillips Semiconductors en 1982 para una serie de chips de la compañía [33]. La especificación original permitía únicamente comunicaciones de 100 kHz, y proveía direccionamiento de máximo 7 bits, limitando de esta manera el número de dispositivos conectados al bus a un máximo de 112.

En 1992 fueron publicadas las primeras especificaciones accesibles al público y se añadió un modo rápido que permitía una comunicación de hasta 400 kHz así como también un espacio de direccionamiento expandido de 10 bits.

En este proyecto se utilizó la comunicación I2C entre el AS7263 y la placa microcontroladora Arduino UNO haciendo uso de las librerías y manipulando éstas para obtener la lectura de los registros enviando un flag caracter (carácter de bandera) desde la interfaz creada en lenguaje Python.

El Arduino fue programado con el uso de Arduino IDE y la programación se muestra en la figura 4.9:

```
1
2
3 #include "AS726X.h"
4 AS726X sensor;
5 void setup() {
6     // Se establece la conexión serial
7     sensor.begin();
8
9 }
10
11 void loop() {
12     if (Serial.available() > 0) {
13         if (Serial.read() == 107) {
14             sensor.takeMeasurementsWithBulb();
15             sensor.printMeasurements();
16         }
17     }
18 }
19 }
20
21 }
22
23 }
```

Figura 0.9 Programa en arduino para la lectura del sensor

Fuente: Investigador

La librería "AS726X.h" con los cambios desarrollados se muestran en el Anexo A.

Las medidas que se obtienen son de los canales R, S, T, U, V y W que corresponden a las longitudes de onda de 610nm, 680nm, 730nm, 760nm, 810nm y 860nm, respectivamente. Además, se tiene un canal extra para la lectura de la temperatura, en grados centígrados, del chip AS7263. La figura 4.10 muestra la lectura obtenida vista en el monitor serie a un baudrate de 115200 baudios.

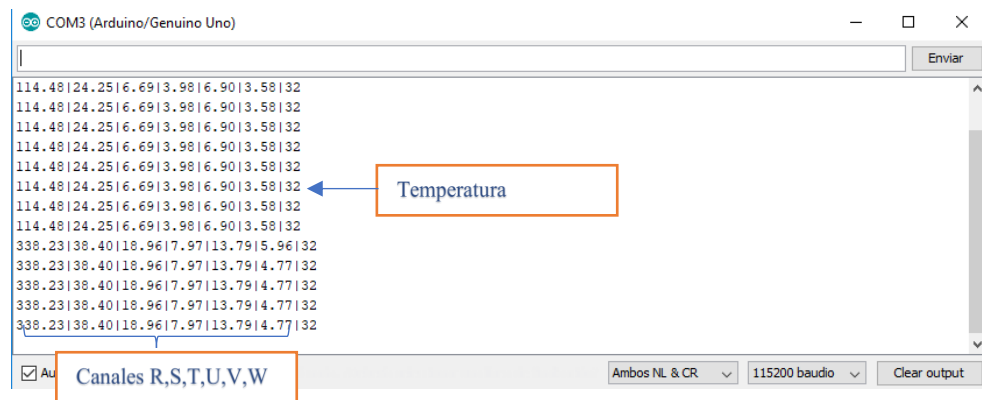


Figura 0.10 Monitor serie de Arduino con lecturas del AS7263

Fuente: Investigador

La etapa de medición, dentro del proyecto, comprende secuencias claves en lo que se refiere a la adquisición y sensado de las variables ópticas para su posterior procesado y análisis. Por ello es importante realizar el diseño de este proceso con cada uno de los parámetros que se van a manejar.

Una vez que se ha desarrollado el procesamiento de los datos en cada etapa del sistema de medición, se manejará la información acerca de los parámetros ópticos obtenidos, considerando las características relevantes. A partir de esto se efectuará la comunicación con el sistema de monitoreo para la visualización local de los datos, mediante una interfaz, en donde se podrá verificar completamente el conjunto de datos generados y con éstos obtener reportes porcentuales de las variables que se desean obtener.

1.22 SISTEMA DE MONITOREO

Con los datos que han sido procesados por el Arduino Uno es factible desarrollar una interfaz gráfica de usuario con software libre, que maneje las ventanas necesarias para registrar los parámetros ópticos que se obtienen como parte del proceso de análisis para determinar la calidad de los alimentos.

1.22.1 Sistema de procesamiento de información

La comunicación se realiza con la placa de desarrollo Raspberry Pi 3, esto permite manejar distintos lenguajes de programación para desarrollar los scripts. que se pueden integrar con estas plataformas basadas en arquitecturas ARM.




Con respecto a otras versiones de la placa Raspberry, este modelo presenta características notables relativas a la capacidad en memoria RAM, procesador y módulos compatibles para utilizar esta placa como un sistema de cómputo completo que permita desarrollar el procesamiento de los datos que se obtienen en el sistema de medición, de igual manera para implementar una interfaz gráfica de usuario que muestre en detalle el estado del sistema.

Además, para poder integrar un sistema de base de datos, se ha visto factible utilizar el lenguaje de programación Python, que permite realizar la escritura de los datos en la base, así como proporcionar el formato adecuado para poder visualizarlos en tiempo real, en una tabla dentro de la aplicación de usuario que se maneja con la pantalla táctil de la Raspberry.

Raspberry PI 3

La Raspberry Pi 3 es una computadora de una sola placa capaz de ejecutar sistemas operativos Android o Linux. Cuenta con un microcontrolador compatible con Arduino integrado para la creación de prototipos a través de diseño con software libre. Se ha elegido utilizar esta minicomputadora considerando las capacidades que posee, principalmente por su procesador y fácil integración de hardware a través de la configuración para entradas y salidas de los pines diseñados para este propósito

Tabla 0.4 Comparación entre placas Raspberry


	Raspberry Pi	Raspberry Pi 2	Raspberry Pi 3
Modelo			

CPU	700 MHz single-core ARM1176JZF-S	900 MHz 32-bit quad-core ARM Cortex-A7	1.2 GHz 64-bit quad-core ARM Cortex-A53
GPU	Broadcom VideoCore IV @ 250 MHz (BCM2837: 3D part of GPU @ 300 MHz, video part of GPU @ 400 MHz) OpenGL ES 2.0 (BCM2835, BCM2836: 24 GFLOPS / BCM2837: 28.8 GFLOPS) MPEG-2 and VC-1 (with license), H.264/MPEG-4 AVC high-profile decoder and encoder	Broadcom VideoCore IV @ 250 MHz (BCM2837: 3D part of GPU @ 300 MHz, video part of GPU @ 400 MHz) OpenGL ES 2.0 (BCM2835, BCM2836: 24 GFLOPS / BCM2837: 28.8 GFLOPS) MPEG-2 and VC-1 (with license), H.264/MPEG-4 AVC high-profile decoder and encoder	Broadcom VideoCore IV @ 250 MHz (BCM2837: 3D part of GPU @ 300 MHz, video part of GPU @ 400 MHz) OpenGL ES 2.0 (BCM2835, BCM2836: 24 GFLOPS / BCM2837: 28.8 GFLOPS) MPEG-2 and VC-1 (with license), H.264/MPEG-4 AVC high-profile decoder and encoder
Compatibilidad con Arduino	✓	✓	✓
GPIO (76 pines disponibles)	✓	✓	✓
Módulo WIFI	✗	✗	✓
Gigabit Ethernet	✓	✓	✓
Puerto SATA	✗	✗	✓

Fuente: Investigador [14]

Tabla 0.5 Características de la Raspberry Pi 3

Raspberry Pi 3	Características
	1.2 GHz 64-bit quad-core ARM Cortex-A53
	RAM DDR3 1GB
	76 pines disponibles compatibles con Arduino
	HDMI LVDS + Touch
	1 micro USB (1 OTG)
	2 USB 2.0 tipo A y 1 conector USB 2.0 interno
	Conectores analógicos de audio y micrófono
	Conexión de cámara CSI
	Lector de tarjetas Micro SD

	Fuente de alimentación (6-15 V CC) y conector de batería externa
	10/100 Mbit/s Ethernet, Bluetooth 4.1
	Módulo WiFi 802.11n

Fuente: Investigador [14]

1.22.2 Comunicación con la placa Raspberry Pi 3

Los datos son enviados por vía serial a la placa Raspberry Pi 3, la misma que se conecta a una pantalla previamente calibrada para la presentación de la información. Es muy importante considerar que ciertos pines no se pueden configurar en diferentes estados, ya que esto pudiese causar un daño a la placa. Por ello se debe conocer la manera adecuada de manejar el tratamiento de los datos con el procesador ARM Cortex-A53 que la placa posee para el manejo de los sistemas operativos Android o Linux, para desarrollar cualquier tipo de proyecto.

La Raspberry Pi3 posee pines dedicados a establecer la comunicación serial con dispositivos como microcontroladores, memorias o sistemas de medición, sin embargo, para evitar inconvenientes con el sistema de enlace para la recepción de datos, es más seguro utilizar un cable de comunicación usb, para comunicar la placa del circuito de medición con la Raspberry Pi 3 a través del puerto USB.

1.22.3 Configuración de Software

Raspbian 4.9 es la distribución que se ha seleccionado para cargar en la Raspberry Pi 3 a través de la imagen que es grabada en la memoria micro SD clase 10. Permite implementar los recursos necesarios para poner en marcha el Sistema Operativo que facilitará la instalación de cada uno de los drivers para el manejo del hardware necesario.

Se ha elegido a Python como el lenguaje de programación que permita escribir el código fuente que servirá para desarrollar las funciones que establecen la comunicación, gestionar la base de datos y manejar las ventanas de la interfaz para el monitoreo de los parámetros eléctricos. El Ide para diseñar el programa en Python que se seleccionó es PyCharm en su versión profesional, debido principalmente a la facilidad que presenta en la manipulación de las librerías, tanto en el desarrollo de la aplicación gráfica, como en el sistema de comunicación y las funciones matemáticas necesarias para algunos procesos.

El código realizado en Python que se encuentra incluido como parte del proyecto principal implementado para la aplicación GUI se presenta en el Anexo B.

1.22.4 Diseño de la Interfaz Gráfica

Con respecto al modelamiento de la interfaz gráfica se ha determinado como la mejor opción implementar un diseño de ésta utilizando la herramienta de *Qt Designer*, ya que este software de diseño de *GUI's* brinda la facilidad de transformar el código .ui a la extensión Python (.py) con mucha facilidad para la implementación de funciones y métodos requeridos para el desarrollo de la interfaz del analizador.

Se pensó en la utilización de MyOpenLab como la herramienta principal para el desarrollo de la propuesta de la interfaz, ya que este software, con manejo de licencia libre desarrollado principalmente para sistemas Linux, posee una similitud para el diseño de proyectos con interfaz gráfica al igual que LabView, sin embargo tras haber verificado cada una de las funciones que proporciona, se determinó que muchas de las características que se requieren no se pueden implementar para un monitoreo en tiempo real que se necesita para el proyecto.

La plataforma de desarrollo gráfico Qt tiene gran portabilidad, y aunque principalmente está estructurada para desarrollar proyectos en lenguaje C bajo licencia GNU LGPL, permite utilizar la extensión de *Qt Designer* como una herramienta para diseñar de manera profesional la interfaz requerida.

En *Qt Designer* se ha estructurado el modelo de la interfaz a través de diferentes pantallas, las mismas que poseen una presentación particular para interactuar directamente con el analizador. Se ha procurado implementar esta interfaz a través de un sistema de control y monitoreo de fácil manejo para una interacción amigable con el usuario.

Las pantallas que manejará la interfaz para el NIR son:

- Pantalla de Bienvenida
- Pantalla para mostrar los resultados de la medición
- Pantalla de gráficas

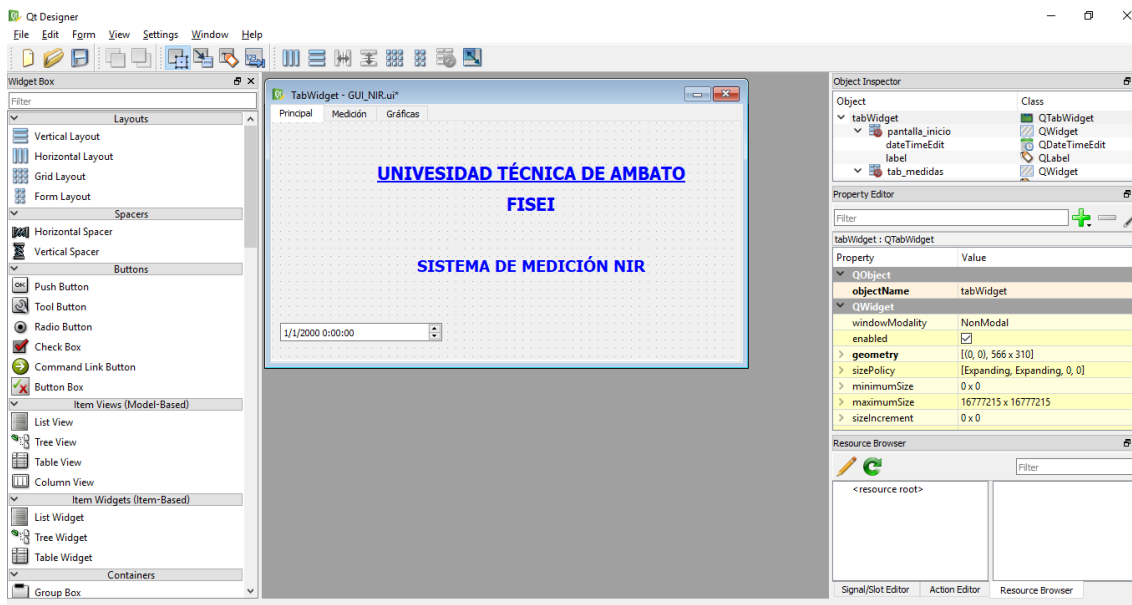


Figura 2. 1 Interfaz en Qt Designer

Fuente: Investigador

Qt Designer permite ejecutar en vista previa el modelo de la interfaz gráfica para el usuario, pudiendo realizar así los cambios necesarios con respecto a los *widgets* que mejor se ajusten a la aplicación. A través de la función incorporada *uic* del paquete *Qt5* para *Python* se realiza la reconstrucción del código *.ui* para poder manejarlo como un *script .py*, que es un formato de código más entendible.

La programación de las funciones para cada una de las pantallas de la aplicación se realiza dentro de la plataforma PyCharm que permite incorporar cada uno de los módulos requeridos para el desarrollo.

1.23 DISEÑO E IMPLEMENTACIÓN DEL SERVIDOR

Para manejar la información dentro de una base de datos y con el fin de mantener un respaldo del registro que se genere con cada análisis realizado para los sistemas eléctricos, se ha implementado un servidor que permite manejar las consultas y gestionar procesos de actualización de manera eficiente.

1.23.1 Instalación del Servidor LAMP

La configuración para el levantamiento de este servidor en el sistema operativo Linux para para la distribución Raspbian de la placa Raspberry PI se detalla en el siguiente bloque.

El servidor que se ha elegido es Apache para manejar las conexiones entrantes y manejo de las bases de datos que se solicitan.

- Para instalarlo se debe ingresar por terminal la siguiente instrucción:

```
sudo apt-get install apache2
```

- Se comprueba que los derechos para el acceso sean los correctos:

```
sudo groupadd www-data
```

```
sudo usermod -g www-data www-data
```

```
sudo chown -R www-data: www-data / var / www
```

- A continuación, se debe instalar el intérprete php para el servidor de la siguiente manera:

```
sudo apt-get install php5 php5-cli
```

- El servidor titular para la base de datos MySQL generalmente se instala de manera automática en los pasos previos, sin embargo, hay que asegurarse de su instalación con la siguiente instrucción:

sudo apt-get install mysql-server mysql-client

MySQL permite insertar los datos necesitados en una base de datos, además ayuda a la creación de páginas web para interactuar con los datos que se van a generar con la aplicación. La gran ventaja que se puede resaltar con respecto a este servidor es que se puede manejar con facilidad el conector mysql dentro de los *scripts* de Python.

Para la integración de la base de datos sql con el intérprete php es necesario instalar phpmyadmin. A través de esta herramienta se podrá controlar fácilmente a través de un acceso cifrado la base de datos y la gestión con las páginas web que se asocien. [34]

- Para realizar la instalación de phpmyadmin se ingresa la siguiente instrucción:

sudo apt-get install php5-mysql phpmyadmin php5-curl nano

- Una vez completada la instalación, se agrega phpmyadmin a la configuración de Apache:

sudo nano /etc/apache2/apache2.conf

- Se debe añadir la cadena de configuración de phpmyadmin apache en la parte inferior del archivo:

include /etc/phpmyadmin/apache.conf

- Finalmente se debe reiniciar el servicio apache, para que cada una de las configuraciones tengan efecto:

sudo service apache2 restart

- Para verificar si todo está configurado correctamente se abre el navegador y se ingresa la siguiente dirección:

http://localhost

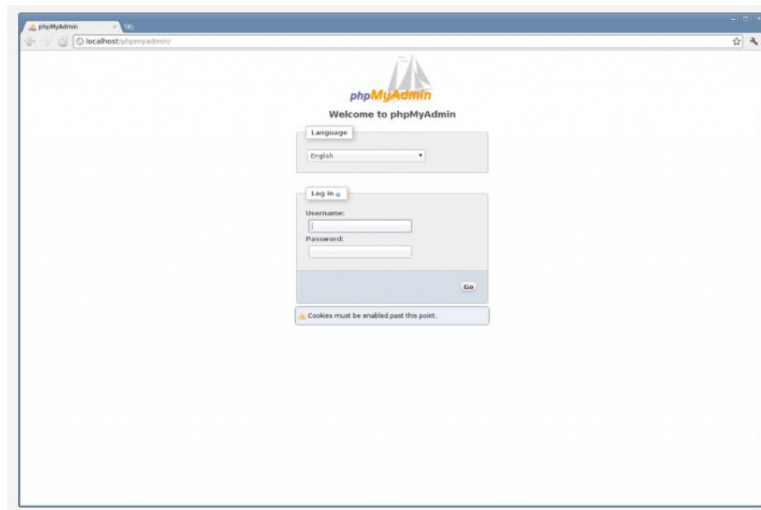


Figura 2. 2 Página de acceso a phpmyadmin [34]

1.2.3.2 Base de Datos

Para el manejo de la información se ha visto factible crear una base general que maneje dos tablas para recolección de información acerca de los principales parámetros.

En la siguiente figura se puede apreciar cada uno de los registros que se manejaron para la tabla de datos del NIR.

FECHA	CANAL_R	CANAL_S	CANAL_T	CANAL_U	CANAL_V	CANAL_W	TEMPERATURA
2018-02-01 19:13:23	858.58	206.14	58.01	29.88	44.34	25.04	27
2018-02-01 19:13:25	859.88	206.14	58.01	29.88	44.34	25.04	27
2018-02-01 19:13:28	858.58	206.14	58.01	29.88	44.34	25.04	27
2018-02-01 19:13:30	858.58	206.14	58.01	29.88	44.34	25.04	27
2018-02-01 19:13:33	858.58	206.14	58.01	29.88	44.34	25.04	27
2018-02-01 19:13:36	859.88	206.14	58.01	29.88	44.34	25.04	27
2018-02-01 19:13:38	858.58	206.14	58.01	29.88	44.34	25.04	27
2018-02-01 19:13:41	858.58	206.14	58.01	29.88	44.34	25.04	27
2018-02-01 19:13:44	859.88	206.14	58.01	29.88	44.34	25.04	27
2018-02-01 19:13:46	857.28	206.14	58.01	29.88	44.34	25.04	27
2018-02-01 19:13:49	858.58	206.14	58.01	29.88	44.34	25.04	27
2018-02-01 19:13:52	858.58	206.14	58.01	29.88	44.34	25.04	27
2018-02-01 19:13:54	858.58	206.14	58.01	29.88	44.34	25.04	27
2018-02-01 19:13:57	858.58	206.14	58.01	29.88	44.34	25.04	27
2018-02-01 19:14:00	858.58	206.14	58.01	29.88	44.34	25.04	27
2018-02-01 19:14:02	858.58	206.14	58.01	29.88	44.34	25.04	27
2018-02-01 19:14:05	859.88	206.14	58.01	29.88	44.34	25.04	27
2018-02-01 19:14:08	858.58	206.14	56.89	29.88	44.34	25.04	27
2018-02-01 19:14:10	858.58	206.14	58.01	29.88	44.34	25.04	27
2018-02-01 19:14:13	857.28	206.14	58.01	29.88	44.34	25.04	27
2018-02-01 19:34:38	807.84	197.05	53.55	27.89	41.38	23.85	28
2018-02-01 19:34:40	805.24	196.04	53.55	27.89	41.38	23.85	28
2018-02-01 19:34:43	806.54	196.04	53.55	27.89	41.38	23.85	28
2018-02-01 19:34:46	805.24	196.04	53.55	27.89	41.38	23.85	28
2018-02-01 19:34:48	806.54	196.04	53.55	27.89	41.38	23.85	28

Figura 2. 3 Base de Datos Utilizada en el Sistema

Fuente: Investigador

1.23.3 Integración con Python

Dentro de la interfaz gráfica de usuario se puede tener acceso a la base en forma dinámica, accediendo a ésta para las operaciones tanto de consulta como de escritura, mediante la configuración del conector mysql para Python. El programa principal ejecuta cada uno de los procesos que permiten adquirir la información acerca del análisis de calidad para interactuar con la base. El procesamiento de la información en este proyecto basa su programación principalmente orientada en Arduino IDE, Python y SQL.

En Python se realizó el tratamiento de los datos alojados en la base de datos llamada DATOS_NIR para lo cual se extrajeron dichos datos en un documento con extensión .txt para luego mediante las librerías pandas, plotly.graph_objs, sklearn.preprocessing, numpy, como se muestra en el Anexo C, crear un gráfico de los componentes principales de la muestra y así obtener una lectura de los componente de proteína, grasa y humedad presentes. La gráfica de los componentes principales se muestra en la figura 4.11.

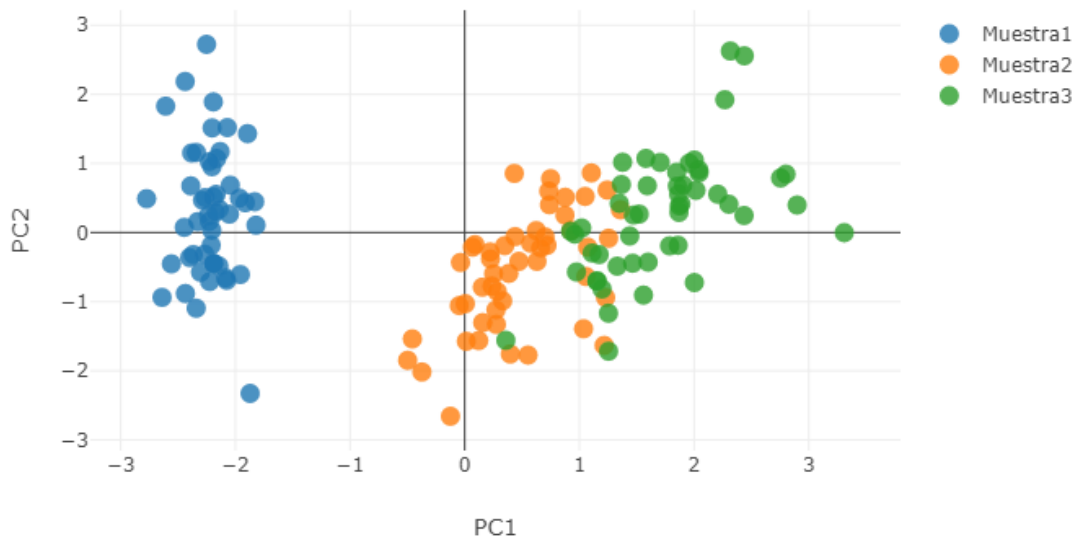


Figura 0.11 Componentes principales de las muestras

Fuente: Investigador

1.24 ANÁLISIS DE RESULTADOS

La implementación se la puede observar en el Anexo D, el sistema se probó con una muestra de harina de maíz, la calibración se la hizo como recomienda la normativa ISO 12099:2017 con la toma de 40 muestras y se realizó el cálculo, obteniéndose los resultados siguientes

En la tabla 4.6 se muestran la cantidad de observaciones con los resultados obtenidos en total, ésta servirá para la obtención de los componentes principales.

Tabla 0.6 Estadísticos descriptivos (Datos cuantitativos)

Variable	Observaciones	Obs. con datos perdidos	Obs. sin datos perdidos	Mínimo	Máximo	Media	Desv. típica
R	40	0	40	805,240	859,880	832,365	26,691
S	40	0	40	196,040	206,140	201,115	5,091
T	40	0	40	53,550	58,010	55,752	2,237
U	40	0	40	27,890	29,880	28,885	1,008
V	40	0	40	41,380	44,340	42,860	1,499
W	40	0	40	23,850	25,040	24,445	0,603

El Análisis de Componentes Principales se muestra en la Tabla 4.7:

Tabla 0.7 Valores propios

	F1	F2	F3	F4
Valor propio	5,994	0,005	0,001	0,000
Variabilidad (%)	99,893	0,086	0,012	0,008
% acumulado	99,893	99,979	99,992	100,000

Después de analizar los valores obtenidos de la muestra de harina de maíz, lo posterior es determinar el modelo predictivo que se ajuste mejor al rango del espectro resultante de la absorbancia para el total de muestras.

Los valores propios son usados para comprimir la cantidad de datos en una dimensión menor, en la tabla 4.7 los valores están comprimidos en la columna F1.

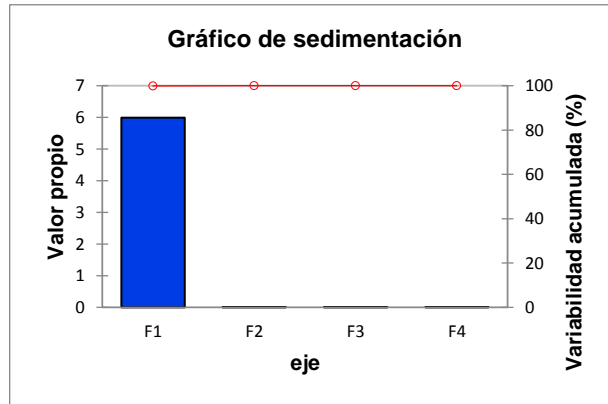


Figura 0.12 Gráfico de sedimentación

En la tabla 4.8 se tienen los vectores propios de cada canal, asociado con los valores propios observados en la tabla 4.7. Se puede ver que en F1 están la mayor cantidad de datos de cada uno de los canales a excepción de los canales R, S y T que se encuentran en F4, F3 y F2 respectivamente.

Tabla 0.8 Vectores propios:

	F1	F2	F3	F4
R	0,408	-0,198	0,298	0,840
S	0,408	-0,218	0,727	-0,507
T	0,408	0,912	0,045	0,000
U	0,408	-0,165	-0,356	-0,111
V	0,408	-0,165	-0,356	-0,111
W	0,408	-0,165	-0,356	-0,111

Tabla 0.9 Cargas factoriales:

	F1	F2	F3	F4
R	1,000	-0,014	0,008	0,019
S	1,000	-0,016	0,020	-0,011
T	0,998	0,066	0,001	0,000
U	1,000	-0,012	-0,010	-0,002
V	1,000	-0,012	-0,010	-0,002
W	1,000	-0,012	-0,010	-0,002

Tabla 0.10 Correlaciones entre las variables y los factores:

	F1	F2	F3	F4
R	1,000	-0,014	0,008	0,019
S	1,000	-0,016	0,020	-0,011
T	0,998	0,066	0,001	0,000
U	1,000	-0,012	-0,010	-0,002
V	1,000	-0,012	-0,010	-0,002
W	1,000	-0,012	-0,010	-0,002

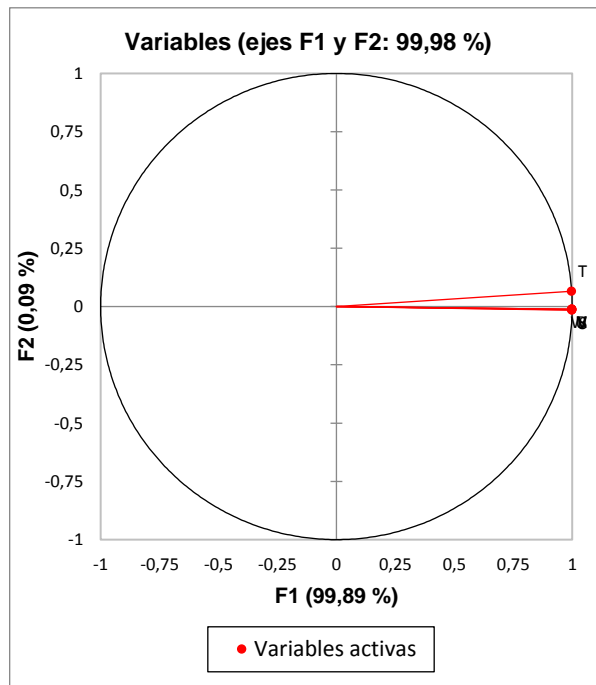


Figura 0.13 Variables (ejes F1 y F2: 99,98 %)

Tabla 0.11 Contribuciones de las variables (%):

	F1	F2	F3	F4
R	16,674	3,906	8,859	70,561
S	16,672	4,774	52,811	25,744
T	16,613	83,181	0,206	0,000
U	16,680	2,713	12,708	1,232
V	16,680	2,713	12,708	1,232
W	16,680	2,713	12,708	1,232

Tabla 0.12 Cosenos cuadrados de las variables:

	F1	F2	F3	F4
R	0,999	0,000	0,000	0,000
S	0,999	0,000	0,000	0,000
T	0,996	0,004	0,000	0,000
U	1,000	0,000	0,000	0,000
V	1,000	0,000	0,000	0,000
W	1,000	0,000	0,000	0,000

Los valores en negrita corresponden para cada variable al factor para el cual el coseno cuadrado es el mayor

Tabla 0.13 Puntuaciones factoriales:

	F1	F2	F3	F4
Obs1	2,456	0,023	-0,001	-0,004
Obs2	2,476	0,013	0,014	0,037
Obs3	2,456	0,023	-0,001	-0,004
Obs4	2,456	0,023	-0,001	-0,004
Obs5	2,456	0,023	-0,001	-0,004
Obs6	2,476	0,013	0,014	0,037
Obs7	2,456	0,023	-0,001	-0,004
Obs8	2,456	0,023	-0,001	-0,004
Obs9	2,476	0,013	0,014	0,037
Obs10	2,436	0,033	-0,015	-0,046
Obs11	2,456	0,023	-0,001	-0,004
Obs12	2,456	0,023	-0,001	-0,004
Obs13	2,456	0,023	-0,001	-0,004
Obs14	2,456	0,023	-0,001	-0,004

Obs15	2,456	0,023	-0,001	-0,004
Obs16	2,456	0,023	-0,001	-0,004
Obs17	2,476	0,013	0,014	0,037
Obs18	2,250	-0,439	-0,024	-0,004
Obs19	2,456	0,023	-0,001	-0,004
Obs20	2,436	0,033	-0,015	-0,046
Obs21	-2,342	-0,054	0,160	-0,039
Obs22	-2,464	0,009	-0,016	-0,020
Obs23	-2,444	-0,001	-0,001	0,022
Obs24	-2,464	0,009	-0,016	-0,020
Obs25	-2,444	-0,001	-0,001	0,022
Obs26	-2,464	0,009	-0,016	-0,020
Obs27	-2,444	-0,001	-0,001	0,022
Obs28	-2,444	-0,001	-0,001	0,022
Obs29	-2,444	-0,001	-0,001	0,022
Obs30	-2,464	0,009	-0,016	-0,020
Obs31	-2,464	0,009	-0,016	-0,020
Obs32	-2,464	0,009	-0,016	-0,020
Obs33	-2,444	-0,001	-0,001	0,022
Obs34	-2,464	0,009	-0,016	-0,020
Obs35	-2,444	-0,001	-0,001	0,022
Obs36	-2,444	-0,001	-0,001	0,022
Obs37	-2,444	-0,001	-0,001	0,022
Obs38	-2,464	0,009	-0,016	-0,020
Obs39	-2,464	0,009	-0,016	-0,020
Obs40	-2,444	-0,001	-0,001	0,022

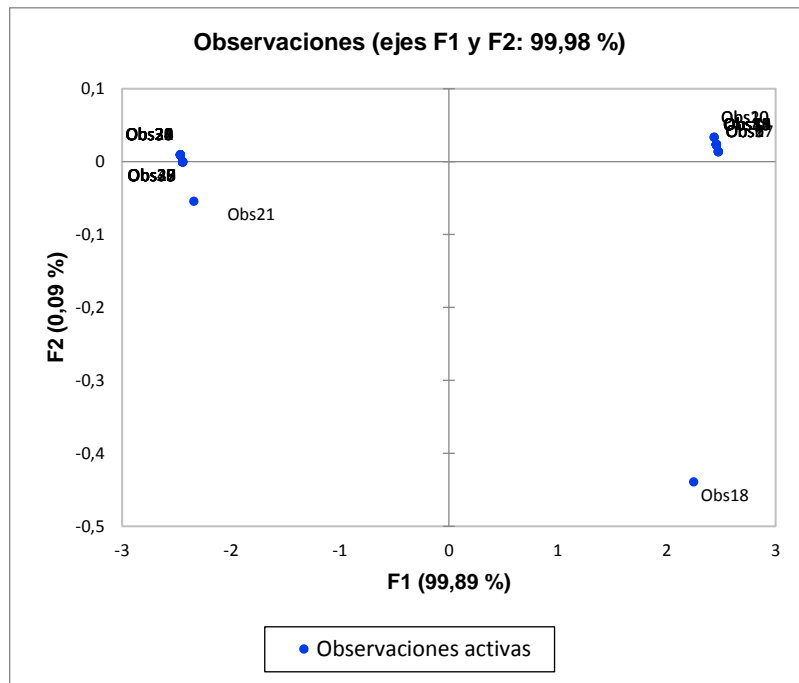


Figura 0.14 Observaciones (ejes F1 y F2: 99,98 %)

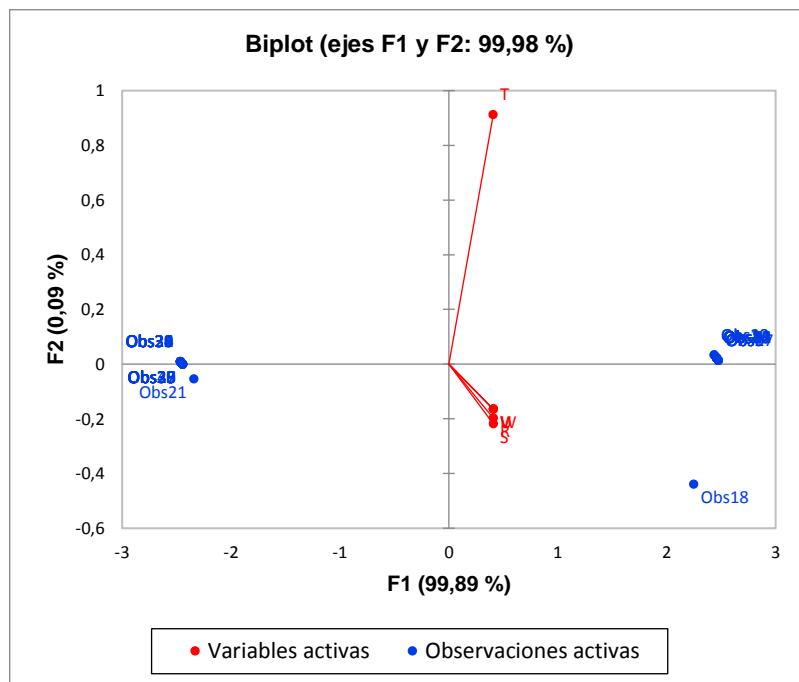


Figura 0.15 Biplot (ejes F1 y F2: 99,98 %)

De acuerdo con la tabla 4.14, cada una de las observaciones realizadas, contribuyen en el porcentaje señalado para la creación del modelo PCA.

Tabla 0.14 Contribuciones de las observaciones (%):

	F1	F2	F3	F4
Obs1	2,516	0,262	0,001	0,089
Obs2	2,558	0,088	0,673	7,049
Obs3	2,516	0,262	0,001	0,089
Obs4	2,516	0,262	0,001	0,089
Obs5	2,516	0,262	0,001	0,089
Obs6	2,558	0,088	0,673	7,049
Obs7	2,516	0,262	0,001	0,089
Obs8	2,516	0,262	0,001	0,089
Obs9	2,558	0,088	0,673	7,049
Obs10	2,475	0,527	0,788	10,578
Obs11	2,516	0,262	0,001	0,089
Obs12	2,516	0,262	0,001	0,089
Obs13	2,516	0,262	0,001	0,089
Obs14	2,516	0,262	0,001	0,089
Obs15	2,516	0,262	0,001	0,089
Obs16	2,516	0,262	0,001	0,089
Obs17	2,558	0,088	0,673	7,049
Obs18	2,111	93,405	1,888	0,097
Obs19	2,516	0,262	0,001	0,089
Obs20	2,475	0,527	0,788	10,578
Obs21	2,287	1,438	86,133	7,641
Obs22	2,533	0,038	0,850	1,974
Obs23	2,491	0,000	0,005	2,398

Obs24	2,533	0,038	0,850	1,974
Obs25	2,491	0,000	0,005	2,398
Obs26	2,533	0,038	0,850	1,974
Obs27	2,491	0,000	0,005	2,398
Obs28	2,491	0,000	0,005	2,398
Obs29	2,491	0,000	0,005	2,398
Obs30	2,533	0,038	0,850	1,974
Obs31	2,533	0,038	0,850	1,974
Obs32	2,533	0,038	0,850	1,974
Obs33	2,491	0,000	0,005	2,398
Obs34	2,533	0,038	0,850	1,974
Obs35	2,491	0,000	0,005	2,398
Obs36	2,491	0,000	0,005	2,398
Obs37	2,491	0,000	0,005	2,398
Obs38	2,533	0,038	0,850	1,974
Obs39	2,533	0,038	0,850	1,974
Obs40	2,491	0,000	0,005	2,398

En la figura 4.15, se desea filtrar las observaciones y las variables con suma de cosenos al cuadrado mayor de 0.5. Para hacerlo, se activa la opción Filtro en XLSTAT y se selecciona $>\cos^2$, introduciendo el valor 0.5.

Tabla 0.15 Cosenos cuadrados de las observaciones:

	F1	F2	F3	F4
Obs1	1,000	0,000	0,000	0,000
Obs2	1,000	0,000	0,000	0,000
Obs3	1,000	0,000	0,000	0,000
Obs4	1,000	0,000	0,000	0,000

Obs5	1,000	0,000	0,000	0,000
Obs6	1,000	0,000	0,000	0,000
Obs7	1,000	0,000	0,000	0,000
Obs8	1,000	0,000	0,000	0,000
Obs9	1,000	0,000	0,000	0,000
Obs10	0,999	0,000	0,000	0,000
Obs11	1,000	0,000	0,000	0,000
Obs12	1,000	0,000	0,000	0,000
Obs13	1,000	0,000	0,000	0,000
Obs14	1,000	0,000	0,000	0,000
Obs15	1,000	0,000	0,000	0,000
Obs16	1,000	0,000	0,000	0,000
Obs17	1,000	0,000	0,000	0,000
Obs18	0,963	0,037	0,000	0,000
Obs19	1,000	0,000	0,000	0,000
Obs20	0,999	0,000	0,000	0,000
Obs21	0,995	0,001	0,005	0,000
Obs22	1,000	0,000	0,000	0,000
Obs23	1,000	0,000	0,000	0,000
Obs24	1,000	0,000	0,000	0,000
Obs25	1,000	0,000	0,000	0,000
Obs26	1,000	0,000	0,000	0,000
Obs27	1,000	0,000	0,000	0,000
Obs28	1,000	0,000	0,000	0,000
Obs29	1,000	0,000	0,000	0,000
Obs30	1,000	0,000	0,000	0,000
Obs31	1,000	0,000	0,000	0,000

Obs32	1,000	0,000	0,000	0,000
Obs33	1,000	0,000	0,000	0,000
Obs34	1,000	0,000	0,000	0,000
Obs35	1,000	0,000	0,000	0,000
Obs36	1,000	0,000	0,000	0,000
Obs37	1,000	0,000	0,000	0,000
Obs38	1,000	0,000	0,000	0,000
Obs39	1,000	0,000	0,000	0,000
Obs40	1,000	0,000	0,000	0,000

Los valores en negrita corresponden para cada observación al factor para el cual el coseno cuadrado es el mayor

Con la obtención de los resultados del método PCA, se pudo desarrollar el método que permite la obtención de las cantidades de proteína, grasa y humedad de la muestra de harina de maíz. Los valores reales son los obtenidos de la tabla de contenidos de nutrientes en alimentos seleccionados de Organización De Las Naciones Unidas Para La Alimentación y Agricultura [35], lo que permitió obtener los siguientes valores que se muestran en la tabla 4.16:

Tabla 0.16 Resultados comparativos

Valores Reales (%)			Valores obtenidos (%)			Error relativo (%)		
Proteína	Grasa	Humedad	Proteína	Grasa	Humedad	Proteína	Grasa	Humedad
9,4	1	12,3	8,4	1,3	11,8	1	-0,3	0,5
9,4	1	12,3	7,5	2,4	12,4	1,9	-1,4	-0,1
9,4	1	12,3	9,3	1,9	11,8	0,1	-0,9	0,5
9,4	1	12,3	8,8	2,1	10,8	0,6	-1,1	1,5
9,4	1	12,3	9,25	1	11,6	0,15	0	0,7
9,4	1	12,3	9,55	0,9	9,1	-0,15	0,1	3,2
9,4	1	12,3	9,85	1,2	12,6	-0,45	-0,2	-0,3
9,4	1	12,3	10,15	0,8	13,1	-0,75	0,2	-0,8
9,4	1	12,3	10,45	0,76	11,8	-1,05	0,24	0,5
9,4	1	12,3	10,75	1,5	12,7	-1,35	-0,5	-0,4

1.25 ANÁLISIS ECONÓMICO DEL PROYECTO

Los costos de los dispositivos necesarios para la consecución del proyecto se muestran en la tabla 4.6.

Tabla 0.17 Análisis del costo

ITEM	CANTIDAD	UNIDAD	DESCRIPCIÓN	PRECIO UNITARIO	PRECIO TOTAL
1	1	c/u	Arduino UNO R3	\$24.95	\$24.95
2	1	c/u	SparkFun Qwiic Shield para Arduino	\$ 5.95	\$ 5.95
3	1	c/u	SparkFun Spectral Sensor Breakout - AS7263 NIR (Qwiic)	\$24.95	\$24.95
4	1	c/u	SparkFun OpenPIR	\$14.95	\$14.95
5	1	c/u	Raspberry Pi LCD - 7" Touchscreen	\$64.95	\$64.95
6	1	c/u	Raspberry Pi 3 Starter Kit	\$89.95	\$89.95
7	3	c/u	Qwiic Cable - 100mm	\$1.50	\$4.50
8	1	c/u	Norma ISO 12099:2017	\$157.00	\$157.00
9	1	c/u	SmartPi Touch	\$24.95	\$24.95
10	100	horas	Diseño	\$3.75	\$375.00
	TOTAL				\$787.15

Fuente: Investigador

El costo unitario de las horas de diseño se obtuvo basado en el sueldo, de 600 dólares mensuales y basado en una carga horaria mensual de 160 horas, de un Ingeniero electrónico para desarrollo y ejecución de proyectos en la empresa Electroprotecciones, encargado del desarrollo de nuevos proyectos en el área de protección contra disturbios eléctricos y calidad de energía. [36]

CAPÍTULO 5

CONCLUSIONES Y RECOMENDACIONES

1.26 CONCLUSIONES

- Se determinó que debido al reducido rango de espectro que es capaz de analizar el circuito AS7263, la calidad de la observación no es la adecuada para determinar con mayor fidelidad los porcentajes de proteína, grasa y humedad presentes en los alimentos.
- La versatilidad de la espectroscopía de infrarrojo cercano en el análisis de alimentos es muy amplia, pero depende de que el sistema esté bien calibrado y sus validaciones puedan ser comprobables.
- El uso de técnicas de análisis de multivariable ha hecho que los métodos aplicados sean estadísticos y esto a su vez ha dado un impulso mayor al desarrollo de plataformas de lenguajes tales como Python o R en la creación de algoritmos que ayuden a los investigadores en el análisis de los datos.

- Un término muy usado durante el desarrollo de esta investigación es el de learning machine que se refiere a algoritmos de inteligencia artificial, que no fueron tratados en este trabajo. Su uso sería de mucha ayuda ya que los sets de datos podrían ser utilizados por la máquina dedicada al análisis de los mismos y así sacar calibraciones automáticas de acuerdo a las observaciones.

1.27 RECOMENDACIONES

- Para obtener mejores resultados se recomienda tomar en cuenta los detalles que propone la norma ISO 12099:2017 en lo que se refiere a validación de muestras y los métodos sugeridos para el análisis de resultados.
- El presente prototipo está diseñado para uso didáctico, por lo que se recomienda que se utilice sólo en ambientes educativos para formar a los usuarios en la utilización de la espectroscopía en mayor escala.
- Para tener una mejor calidad de resultados se recomienda utilizar cámaras que detecten un mayor rango de longitudes de onda ya que, como se vio, las mejores observaciones se dan en el espectro de mayor longitud.
- El análisis de multivariados es muy extendido en diferentes ramas del conocimiento humano, por lo que se recomienda que se añadan estos conocimientos en la malla de la Ingeniería Electrónica para, de esta manera, involucrar a los estudiantes en las nuevas tecnologías de Big Data.

BIBLIOGRAFÍA

- [1] J. Bello Gutiérrez, Ciencia bromatológica: principios generales de los alimentos, España: Ediciones Diaz de Santos S.A., 2000.
- [2] P. Anna, «La espectroscopía NIR en la determinación de propiedades físicas y composición química de intermedios de producción y productos acabados,» Universidad Autónoma de Barcelona, Barcelona, 2010.
- [3] INEN, «www.normalizacion.gob.ec,» 11 Noviembre 2016. [En línea]. Available: <http://apps.normalizacion.gob.ec/tarifario/tarifario.pdf>. [Último acceso: 28 Marzo 2018].
- [4] V. L. Coral Torres, «Pontificia Universidad Católica del Ecuador. Repositorio PUCE,» 2014. [En línea]. Available: <http://repositorio.puce.edu.ec/bitstream/handle/22000/8924/Determinaci%C3%B3n%20proximal%20de%20los%20principales%20componentes%20nutricionales%20de%20siete%20alimentos.pdf?sequence=1>. [Último acceso: 10 Abril 2018].
- [5] Departamento de Producción Animal de la Universidad de Córdoba, «Producción animal y gestión de empresas,» Universidad de Córdoba, Febrero 2014. [En línea]. Available: <https://www.uco.es/zootecniaygestion/menu.php?tema=146>. [Último acceso: 28 Marzo 2018].
- [6] Grupo Selecta, «Método Kjeldah JP Selecta S.A. Notas de Aplicación,» JP Selecta S.A. , 10 Octubre 2012. [En línea]. Available: <http://www.grupo-selecta.com/notasdeaplicaciones/sin-categoria/metodo-kjeldahl/>. [Último acceso: 10 Abril 2018].
- [7] J. F. Ocampo, «Estandarización de las curvas de calibración por la metodología NIR y la química húmeda,» 2015. [En línea]. Available:

http://repository.lasallista.edu.co/dspace/bitstream/10567/1571/1/Estandarizacion_curvas_calibracion_NIR.pdf. [Último acceso: 02 11 2017].

- [8] Sigma-Aldrich Química, S.L., «Vetec: Reactivos Grado ACS,» 2013. [En línea]. Available: <https://www.sigmaaldrich.com/mexico/promo/sigma-vetec-mexico.html>. [Último acceso: 29 Marzo 2018].
- [9] J. Castañeda, L. García, V. Jaramillo y J. Pineda, «Diseño de un dispositivo mecatrónico de espectroscopia para el monitoreo de sistemas biológicos,» IV Congreso Internacional 314 De Ingeniería Mecatrónica Y Automatización - Ciima 2015, Manizales, 2015.
- [10] H. You, Y. Kim, J. Hyung Lee y S. Choi, «Classification of food powders using handheld NIR spectrometer,» *IEEE Conference Publications*, pp. 732-734, 2017.
- [11] W. Song, H. Wang, P. Maguire y O. Nibouche, «Differentiation of Organic and Non-organic Apples,» *IEEE Conferetion*, vol. 1, n° 24, pp. 234-239, 2016.
- [12] Excma. Diputación Provincial de Pontevedra, «Aplicación de la espectroscopía NIR al análisis de algunos parámetros nutricionales en habas,» Laboratorio Mouriscade, Pontevedra, 2016.
- [13] E. Serenato Madalozzo, E. Sauer y N. Nagata, «Determination of fat, protein and moisture in ricotta cheese,» *J Food Sci Technol*, vol. 3, n° 52, p. 1649–1655, 2015.
- [14] F. Estienne, L. Pasti, V. Centner, B. Walczak, F. Despagne, D. Jouan Rimbaud, O. de Noord y D. L. Massart, «A comparison of multivariate calibration techniques applied to experimental NIR data sets Part II. Predictive ability under extrapolation conditions,» *Chemometrics and Intelligent Laboratory Systems*, n° 58, p. 195–211, 2002.
- [15] C. Haiyan y H. Yong, «“Theory and application of near infrared reflectance spectroscopy in determination of food quality”,» *Trends in Food Science & Technology*, n° 18, pp. 72-83, 2007.

- [16] T. Miller y N. Krogan, «COMPASS: A complex of proteins associated with a trithorax-related SET domain protein,» *The National Academy of Sciences*, vol. 98, nº 23, p. 101, 2001.
- [17] M. D. L. (. L. -. D. H. K. M. D. S. Tokkie Groenewald, «Engormix,» 26 Marzo 2015. [En línea]. Available: http://www.engormix.com/s_articles_view.asp?art=577&AREA=AV. [Último acceso: 10 Noviembre 2017].
- [18] D. Martínez B., O. D. Gil Novoa y A. J. Barón González, «ESPECTROMETRIA DE FLUORESCENCIA DE RAYOS X,» *Revista Colombiana de Física*, vol. 38, nº 4, pp. 790-793, 2006.
- [19] A. Lara Barragán, G. Cerpa , M. E. Rodríguez y H. Núñez, «Movimiento armónico,» de *Física Para Bachillerato Cinemática*, México, Pearson Education, 2006, p. 192.
- [20] A. Steinbach, «NIR Spectroscopy - A guide to near-infrared spectroscopic analysis,» Metrohm AG, Herisau, 2013.
- [21] O. Lombardi y J. C. Martínez González, «Sci. stud,» 2012. [En línea]. Available: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1678-31662012000400002. [Último acceso: 12 Abril 2018].
- [22] Testo Argentina, «Calibración instrumentos de medición,» Testo Argentina S.A., 2018. [En línea]. Available: <https://www.testo.com/es-AR/servicios/calibracion>. [Último acceso: 1 Abril 2018].
- [23] C. M. Ruiz, «Universidad de Sevilla,» Junio 2017. [En línea]. Available: <https://idus.us.es/xmlui/bitstream/handle/11441/63208/M%C3%A1rquez%20Ruiz%20Cristina%20TFG.PDF?sequence=1>. [Último acceso: 19 Enero 2018].
- [24] M. G. Redrobán Matute, «Implementación Del Algoritmo De Componentes Principales Recursivo (Rpca) Para El Análisis Del Comportamiento Del Ciclo De

Potencia De Vapor En Dos Plantas Interconectadas Para Generación Eléctrica Y Estudio De Su Aplicabilidad En La Identificación De Fall,» EPN, Quito, 2011.

- [25] NN. [En línea]. Available: <http://www.steren.com.mx/raspberry-pi-3-modelo-b.html>.
- [26] NN, «comohacer.eu,» [En línea]. Available: <https://comohacer.eu/comparativa-y-analisis-raspberry-pi-vs-competencia/>.
- [27] J. Pastor, «Xataka,» 20 enero 2017. [En línea]. Available: <https://www.xataka.com/accesorios/la-tinkerboard-de-asus-intenta-competir-con-la-raspberry-pi-pero-lo-tiene-crudo>.
- [28] I. Huertas, «DataTons,» easyadmin, 8 Abril 2016. [En línea]. Available: <https://blog.datatons.com/2016/04/08/que-es-lenguaje-programacion-r/>. [Último acceso: 1 Abril 2018].
- [29] D. Martínez Pradales, «Python, un lenguaje simple para comprender la complejidad del mundo,» nobbot, 18 Julio 2017. [En línea]. Available: <https://www.nobbot.com/general/python-lenguaje-programacion/>. [Último acceso: 2 Abril 2018].
- [30] Santillana, «7 razones para programar en Python,» BeJob, 18 Septiembre 2016. [En línea]. Available: <https://www.bejob.com/7-razones-para-programar-en-python/>. [Último acceso: 2 Abril 2018].
- [31] ISO, «ISO 12009:2017,» ISO, Geneva, 2017.
- [32] ams AG, «AS7263 6-Channel NIR Spectral_ID Device with Electronic Shutter and Smart Interface,» ams AG, Premstaetten, 2016.
- [33] Sparkfun, «Sparkfun Electronics,» 11 2016. [En línea]. Available: <https://learn.sparkfun.com/tutorials/qwiic-shield-for-arduino--photon-hookup-guide>. [Último acceso: 04 02 2018].

- [34] UDOO, «SERVIDOR WEB UDOO,» [En línea]. Available: <https://www.udoo.org/tutorial/udoo-web-server/>.
- [35] Organización de las Naciones Unidas para la Alimentación y la Agricultura, «Anexo3: Contenido de nutrientes en alimentos seleccionados,» 2010. [En línea]. Available: <http://www.fao.org/docrep/006/w0073s/w0073s1x.htm>. [Último acceso: 11 Abril 2018].
- [36] jooble, «jooble.com,» 17 Marzo 2018. [En línea]. Available: <https://ec.jooble.org/desc/1913613500918654333?ckey=ingeniero+electronico&rgn=0&pos=3&elckey=-7280682526588462200&sid=-8853747328874770899&age=856&relb=100&brelb=100&bscr=832,2099&scr=832,2099&iid=-8270111462114526363>. [Último acceso: 13 Abril 2018].
- [37] M. d. E. y. E. Renovable, «Tips para el uso eficiente de la energía eléctrica,» [En línea]. Available: <http://www.energia.gob.ec/tips-para-el-uso-eficiente-de-la-energia-electrica>.
- [38] J. Rivadeneira y C. Flores, «Diseño e implementación de un sistema microprocesado para adquisición de datos en forma remota de un medidor digital de consumo de energía eléctrica tipo industrial, mediante telefonía celular,» Repositorio Escuela Politécnica Nacional, Quito, 2006.
- [39] D. Sergio de la Fé, «EcuRed,» 27 Junio 2014. [En línea]. Available: https://www.ecured.cu/Caracter%C3%ADsticas_Generales_de_las_Red_de_Distribuci%C3%B3n.
- [40] J. Seymour y T. Horsley, «Los siete tipos de problemas en el suministro eléctrico,» APC Legendary Reliability, 2005.
- [41] NN, «ElectroniLab,» [En línea]. Available: <https://electronilab.co/tienda/sensor-de-corriente-alterna-100a-transformador-sct-013-000/>.

- [42] A. Tirado, «Aprende Física,» 29 marzo 2012. [En línea]. Available: <http://aprendefisika.blogspot.com/p/precision-de-instrumentos-y-errores.html>.
- [43] J. L. Hernández , «Tuveras.com,» 01 julio 2017. [En línea]. Available: <http://www.tuveras.com/fdp/fdp.htm#compensaciontrifasica>.
- [44] S. Wolf y R. F.M. Smith, Guía para Mediciones Electrónicas y Prácticas de Laboratorio, México: Prentice Hall Hispanoamérica S.A., 1992, pp. 278,279.
- [45] NN, «ECURED,» 13 octubre 2017. [En línea]. Available: https://www.ecured.cu/Amplitud_de_la_corriente_alterna.
- [46] O. J. Vélez Osorio y D. Delgado Peláez , «Evaluación de la calidad de la energía en la Universidad Tecnológica de Pereira,» Universidad Tecnológica de Pereira, Facultad de Ingenierías. Pereira, 2008. [En línea]. Available: <http://repositorio.utp.edu.co/dspace/handle/11059/1047>.
- [47] Devices, Analog, «ADE7758 Rev. E,» 2004-2011. [En línea]. Available: <http://www.analog.com/media/en/technical-documentation/data-sheets/ADE7758.pdf>.

ANEXOS

ANEXO A

Código de la librería AS726X.h para Arduino UNO

```
/*  
  
Name:          AS726X.h  
  
Created: 7/11/2017 12:06:22 PM  
  
Author:  andrew.England  
  
Editor:  http://www.visualmicro.com  
  
*/  
  
#ifndef _AS726X_h  
#define _AS726X_h  
#include "Arduino.h"  
#include "Wire.h"  
class AS726X {  
public:  
    AS726X();  
    void begin(TwoWire &wirePort = Wire, byte gain = 3, byte measurementMode  
= 3);  
    void takeMeasurements();  
    void takeMeasurementsWithBulb();  
    void printMeasurements();  
    void printUncalibratedMeasurements();  
    byte getTemperature();
```

```
float getTemperatureF();

void setMeasurementMode(byte mode);

boolean dataAvailable();

void enableIndicator();

void disableIndicator();

void setIndicatorCurrent(byte current);

void enableBulb();

void disableBulb();

void setBulbCurrent(byte current);

void softReset();

void setGain(byte gain);

void setIntegrationTime(byte integrationValue);

void enableInterrupt();

void disableInterrupt();

//Get the various color readings

int getViolet();

int getBlue();

int getGreen();

int getYellow();

int getOrange();

int getRed();

//Get the various NIR readings

int getR();
```

```
int getS();  
int getT();  
int getU();  
int getV();  
int getW();
```

```
//Returns the various calibration data
```

```
float getCalibratedViolet();  
float getCalibratedBlue();  
float getCalibratedGreen();  
float getCalibratedYellow();  
float getCalibratedOrange();  
float getCalibratedRed();
```

```
float getCalibratedR();  
float getCalibratedS();  
float getCalibratedT();  
float getCalibratedU();  
float getCalibratedV();  
float getCalibratedW();
```

```
private:
```

```
TwoWire *_i2cPort;  
int getChannel(byte channelRegister);
```

```

float getCalibratedValue(byte calAddress);

float convertBytesToFloat(uint32_t myLong);

boolean clearDataAvailable();

byte virtualReadRegister(byte virtualAddr);

void virtualWriteRegister(byte virtualAddr, byte dataToWrite);

void writeRegister(byte addr, byte val);

byte readRegister(byte addr);

#define AS726X_ADDR 0x49 //7-bit unshifted default I2C Address

#define SENSORTYPE_AS7262 0x3E

#define SENSORTYPE_AS7263 0x3F

//Register addresses

#define AS726x_DEVICE_TYPE 0x00

#define AS726x_HW_VERSION 0x01

#define AS726x_CONTROL_SETUP 0x04

#define AS726x_INT_T 0x05

#define AS726x_DEVICE_TEMP 0x06

#define AS726x_LED_CONTROL 0x07

#define AS72XX_SLAVE_STATUS_REG 0x00

#define AS72XX_SLAVE_WRITE_REG 0x01

#define AS72XX_SLAVE_READ_REG 0x02

```

//The same register locations are shared between the AS7262 and AS7263,
they're just called something different

//AS7262 Registers

```
#define AS7262_V 0x08  
#define AS7262_B 0x0A  
#define AS7262_G 0x0C  
#define AS7262_Y 0x0E  
#define AS7262_O 0x10  
#define AS7262_R 0x12  
#define AS7262_V_CAL 0x14  
#define AS7262_B_CAL 0x18  
#define AS7262_G_CAL 0x1C  
#define AS7262_Y_CAL 0x20  
#define AS7262_O_CAL 0x24  
#define AS7262_R_CAL 0x28
```

//AS7263 Registers

```
#define AS7263_R 0x08  
#define AS7263_S 0x0A  
#define AS7263_T 0x0C  
#define AS7263_U 0x0E  
#define AS7263_V 0x10  
#define AS7263_W 0x12  
#define AS7263_R_CAL 0x14
```

```
#define AS7263_S_CAL 0x18
#define AS7263_T_CAL 0x1C
#define AS7263_U_CAL 0x20
#define AS7263_V_CAL 0x24
#define AS7263_W_CAL 0x28

#define AS72XX_SLAVE_TX_VALID 0x02
#define AS72XX_SLAVE_RX_VALID 0x01

#define SENSORTYPE_AS7262 0x3E
#define SENSORTYPE_AS7263 0x3F

#define POLLING_DELAY 5 //Amount of ms to wait between checking for virtual
register changes

    byte _sensorVersion = 0;
};

#endif
```

ANEXO A

Código de la GUI

```
# -*- coding: utf-8 -*-

# Form implementation generated from reading ui file 'untitled.ui'
#
# Created by: PyQt5 UI code generator 5.9
#
# WARNING! All changes made in this file will be lost!

from PyQt5 import QtCore, QtGui, QtWidgets
from PyQt5.QtCore import QObject

import sys

try:
    _fromUtf8 = QtCore.QString.fromUtf8
except AttributeError:
    def _fromUtf8(s):
        return s

try:
    _encoding = QtGui.QApplication.UnicodeUTF8
    def _translate(context, text, disambig):
        return QtGui.QApplication.translate(context, text, disambig, _encoding)
except AttributeError:
    def _translate(context, text, disambig):
        return QtGui.QApplication.translate(context, text, disambig)

class Ui_tabWidget(QObject):
    def __init__(self):
        super(Ui_tabWidget,self).__init__()

    def setupUi(self, tabWidget):
        tabWidget.setObjectName("tabWidget")
        tabWidget.resize(566, 310)
        self.pantalla_inicio = QtWidgets.QWidget()
        self.pantalla_inicio.setObjectName("pantalla_inicio")
```



```

self.label = QtWidgets.QLabel(self.pantalla_inicio)
self.label.setGeometry(QtCore.QRect(120, 30, 381, 171))
self.label.setStyleSheet("color: rgb(0, 0, 255);")
self.label.setObjectName("label")
self.dateTimeEdit = QtWidgets.QDateTimeEdit(self.pantalla_inicio)
self.dateTimeEdit.setGeometry(QtCore.QRect(10, 240, 194, 22))
self.dateTimeEdit.setObjectName("dateTimeEdit")
tabWidget.addTab(self.pantalla_inicio, "")
self.tab_medidas = QtWidgets.QWidget()
self.tab_medidas.setObjectName("tab_medidas")
self.pushButton_start = QtWidgets.QPushButton(self.tab_medidas)
self.pushButton_start.setGeometry(QtCore.QRect(50, 30, 171, 31))
self.pushButton_start.setStyleSheet("background-color: rgb(255, 170, 0);")
self.pushButton_start.setObjectName("pushButton_start")
self.textEdit_humedad = QtWidgets.QTextEdit(self.tab_medidas)
self.textEdit_humedad.setGeometry(QtCore.QRect(230, 80, 161, 31))
self.textEdit_humedad.setObjectName("textEdit_humedad")
self.textEdit_proteina = QtWidgets.QTextEdit(self.tab_medidas)
self.textEdit_proteina.setGeometry(QtCore.QRect(230, 120, 161, 31))
self.textEdit_proteina.setObjectName("textEdit_proteina")
self.textEdit_grasa = QtWidgets.QTextEdit(self.tab_medidas)
self.textEdit_grasa.setGeometry(QtCore.QRect(230, 160, 161, 31))
self.textEdit_grasa.setObjectName("textEdit_grasa")
self.label_2 = QtWidgets.QLabel(self.tab_medidas)
self.label_2.setGeometry(QtCore.QRect(60, 80, 161, 31))
self.label_2.setObjectName("label_2")
self.label_3 = QtWidgets.QLabel(self.tab_medidas)
self.label_3.setGeometry(QtCore.QRect(60, 120, 161, 31))
self.label_3.setObjectName("label_3")
self.label_4 = QtWidgets.QLabel(self.tab_medidas)
self.label_4.setGeometry(QtCore.QRect(60, 160, 161, 31))
self.label_4.setObjectName("label_4")
tabWidget.addTab(self.tab_medidas, "")
self.tab_grafica = QtWidgets.QWidget()
self.tab_grafica.setObjectName("tab_grafica")
self.label_5 = QtWidgets.QLabel(self.tab_grafica)
self.label_5.setGeometry(QtCore.QRect(190, 130, 161, 31))
self.label_5.setObjectName("label_5")
self.label_6 = QtWidgets.QLabel(self.tab_grafica)
self.label_6.setGeometry(QtCore.QRect(190, 50, 161, 31))
self.label_6.setObjectName("label_6")
self.label_7 = QtWidgets.QLabel(self.tab_grafica)
self.label_7.setGeometry(QtCore.QRect(190, 90, 161, 31))
self.label_7.setObjectName("label_7")
self.radioButton_humedad = QtWidgets.QRadioButton(self.tab_grafica)
self.radioButton_humedad.setGeometry(QtCore.QRect(160, 60, 81, 21))

```

```

self.radioButton_humedad.setText("")
self.radioButton_humedad.setObjectName("radioButton_humedad")
self.radioButton_proteina = QtWidgets.QRadioButton(self.tab_grafica)
self.radioButton_proteina.setGeometry(QtCore.QRect(160, 96, 81, 21))
self.radioButton_proteina.setText("")
self.radioButton_proteina.setObjectName("radioButton_proteina")
self.radioButton_grasa = QtWidgets.QRadioButton(self.tab_grafica)
self.radioButton_grasa.setGeometry(QtCore.QRect(160, 132, 81, 21))
self.radioButton_grasa.setText("")
self.radioButton_grasa.setObjectName("radioButton_grasa")
self.label_8 = QtWidgets.QLabel(self.tab_grafica)
self.label_8.setGeometry(QtCore.QRect(370, 60, 141, 81))
self.label_8.setObjectName("label_8")
self.pushButton_mostrar_grafica = QtWidgets.QPushButton(self.tab_grafica)
self.pushButton_mostrar_grafica.setGeometry(QtCore.QRect(150, 180, 171, 23))

self.pushButton_mostrar_grafica.setObjectName("pushButton_mostrar_grafica")
tabWidget.addTab(self.tab_grafica, "")

self.retranslateUi(tabWidget)
tabWidget.setCurrentIndex(2)
QtCore.QMetaObject.connectSlotsByName(tabWidget)

def retranslateUi(self, tabWidget):
    _translate = QtCore.QCoreApplication.translate
    tabWidget.setWindowTitle(_translate("tabWidget", "TabWidget"))
    self.label.setText(_translate("tabWidget", "<html><head><body><p
align=\"center\"><span style=\" font-size:16pt; font-weight:600; text-decoration:
underline;\">UNIVERSIDAD TÉCNICA DE AMBATO</span></p><p
align=\"center\"><span style=\" font-size:16pt; font-
weight:600;\">FISEI</span></p><p align=\"center\"><br/></p><p
align=\"center\"><span style=\" font-size:14pt; font-weight:600;\">SISTEMA DE
MEDICIÓN NIR</span></p></body></html>"))
    tabWidget.setTabText(tabWidget.indexOf(self.pantalla_inicio),
_translate("tabWidget", "Principal"))
    self.pushButton_start.setText(_translate("tabWidget", "INICIO"))
    self.label_2.setText(_translate("tabWidget", "<html><head><body><p><span
style=\" font-size:12pt;\">% Humedad:</span></p></body></html>"))
    self.label_3.setText(_translate("tabWidget", "<html><head><body><p><span
style=\" font-size:12pt;\">% Proteina:</span></p></body></html>"))
    self.label_4.setText(_translate("tabWidget", "<html><head><body><p><span
style=\" font-size:12pt;\">% Grasa:</span></p></body></html>"))
    tabWidget.setTabText(tabWidget.indexOf(self.tab_medidas),
_translate("tabWidget", "Medición"))
    self.label_5.setText(_translate("tabWidget", "<html><head><body><p><span
style=\" font-size:12pt;\">% Grasa:</span></p></body></html>"))

```

```

        self.label_6.setText(_translate("tabWidget", "<html><head/><body><p><span
style=\" font-size:12pt;\">% Humedad:</span></p></body></html>"))
        self.label_7.setText(_translate("tabWidget", "<html><head/><body><p><span
style=\" font-size:12pt;\">% Proteina:</span></p></body></html>"))
        self.label_8.setText(_translate("tabWidget", "<!DOCTYPE HTML PUBLIC \"-
//W3C//DTD HTML 4.0//EN\" \"http://www.w3.org/TR/REC-
html40/strict.dtd\">\n"
"<html><head><meta name=\"qrichtext\" content=\"1\" /><style
type=\"text/css\">\n"
"p, li { white-space: pre-wrap; }\n"
"</style></head><body style=\" font-family:'MS Shell Dlg 2'; font-size:8pt; font-
weight:400; font-style:normal;\">\n"
"<p align=\"center\" style=\" margin-top:12px; margin-bottom:12px; margin-
left:0px; margin-right:0px; -qt-block-indent:0; text-indent:0px;\"><span style=\"
color:#aa0000;\">Seleccione una característica <br />de la cual desee obtener<br
/> su gráfica y luego <br />presione el botón <br />Mostrar Gráfica</span></p>\n"
"<p align=\"center\" style=\" -qt-paragraph-type:empty; margin-top:12px;
margin-bottom:12px; margin-left:0px; margin-right:0px; -qt-block-indent:0; text-
indent:0px; color:#aa0000;\"><br /></p>\n"
"<p align=\"center\" style=\" -qt-paragraph-type:empty; margin-top:12px;
margin-bottom:12px; margin-left:0px; margin-right:0px; -qt-block-indent:0; text-
indent:0px; color:#aa0000;\"><br /></p></body></html>"))
        self.pushButton_mostrar_grafica.setText(_translate("tabWidget", "Mostrar
Gráfica"))
        tabWidget.setTabText(tabWidget.indexOf(self.tab_grafica),
_translate("tabWidget", "Gráficas"))

if __name__ == "__main__":
    app=QtWidgets.QApplication(sys.argv)
    tabWidget=QtWidgets.QTabWidget()
    ui=Ui_tabWidget()
    ui.setupUi(tabWidget)
    tabWidget.show()
    sys.exit(app.exec_())

```

ANEXO C

Código del análisis PCA de las muestras

```
import pandas as pd
import plotly as py
from plotly.graph_objs import *
from sklearn.preprocessing import StandardScaler
import numpy as np
from sklearn.decomposition import PCA as sklearnPCA

#init_notebook_mode(connected=True)

df = pd.read_csv(
    filepath_or_buffer=r'C:\Users\luira\PycharmProjects\untitled\Datos_NIR.txt',
    header=None,
    sep=',')

df.columns=['sepal_len', 'sepal_wid', 'petal_len', 'petal_wid', 'class']
df.dropna(how='all', inplace=True) # drops the empty line at file-end

df.tail()

X = df.ix[:,0:4].values
y = df.ix[:,4].values

traces = []

legend = {0:False, 1:False, 2:False, 3:True}

colors = {'Muestra1': 'rgb(31, 119, 180)',
          'Muestra2': 'rgb(255, 127, 14)',
          'Muestra3': 'rgb(44, 160, 44)'}

for col in range(4):
    for key in colors:
        traces.append(Histogram(x=X[y==key, col],
                                opacity=0.75,
                                xaxis='x%s' %(col+1),
                                marker=Marker(color=colors[key]),
                                name=key,
                                showlegend=legend[col]))

data = Data(traces)

layout = Layout(barmode='overlay',
                xaxis=XAxis(domain=[0, 0.25], title='sepal length (cm)'),
                xaxis2=XAxis(domain=[0.3, 0.5], title='sepal width (cm)'),
                xaxis3=XAxis(domain=[0.55, 0.75], title='petal length (cm)'),
                xaxis4=XAxis(domain=[0.8, 1], title='petal width (cm)'),
                yaxis=YAxis(title='count'),
                title='Distribution of the different Iris flower features')
```

```

fig = Figure(data=data, layout=layout)
#py.offline.plot(fig)

X_std = StandardScaler().fit_transform(X)

"""
mean_vec = np.mean(X_std, axis=0)
cov_mat = (X_std - mean_vec).T.dot((X_std - mean_vec)) / (X_std.shape[0]-1)
print('Covariance matrix \n%s' %cov_mat)"""

print('NumPy covariance matrix: \n%s' %np.cov(X_std.T))

cov_mat = np.cov(X_std.T)

eig_vals, eig_vecs = np.linalg.eig(cov_mat)

print('Eigenvectors \n%s' %eig_vecs)
print('\nEigenvalues \n%s' %eig_vals)

cor_mat1 = np.corrcoef(X_std.T)

eig_vals1, eig_vecs1 = np.linalg.eig(cor_mat1)

print('Eigenvectors \n%s' %eig_vecs1)
print('\nEigenvalues \n%s' %eig_vals1)

cor_mat2 = np.corrcoef(X.T)

eig_vals, eig_vecs = np.linalg.eig(cor_mat2)

print('Eigenvectors \n%s' %eig_vecs)
print('\nEigenvalues \n%s' %eig_vals)

u,s,v = np.linalg.svd(X_std.T)
#u
for ev in eig_vecs:
    np.testing.assert_array_almost_equal(1.0, np.linalg.norm(ev))
print('Everything ok!')

# Make a list of (eigenvalue, eigenvector) tuples
eig_pairs = [(np.abs(eig_vals[i]), eig_vecs[:,i]) for i in range(len(eig_vals))]

# Sort the (eigenvalue, eigenvector) tuples from high to low
eig_pairs.sort()
eig_pairs.reverse()

# Visually confirm that the list is correctly sorted by decreasing eigenvalues
print('Eigenvalues in descending order:')
for i in eig_pairs:
    print(i[0])

tot = sum(eig_vals)
var_exp = [(i / tot)*100 for i in sorted(eig_vals, reverse=True)]

```

```

cum_var_exp = np.cumsum(var_exp)

trace1 = Bar(
    x=['PC %s' %i for i in range(1,5)],
    y=var_exp,
    showlegend=False)

trace2 = Scatter(
    x=['PC %s' %i for i in range(1,5)],
    y=cum_var_exp,
    name='cumulative explained variance')

data = Data([trace1, trace2])

layout=Layout(
    yaxis=YAxis(title='Explained variance in percent'),
    title='Explained variance by different principal components')

fig = Figure(data=data, layout=layout)
#py.offline.plot(fig)
matrix_w = np.hstack((eig_pairs[0][1].reshape(4,1),
    eig_pairs[1][1].reshape(4,1)))

print('Matrix W:\n', matrix_w)

Y = X_std.dot(matrix_w)
traces = []

for name in ('Muestra1', 'Muestra2', 'Muestra3'):

    trace = Scatter(
        x=Y[y==name,0],
        y=Y[y==name,1],
        mode='markers',
        name=name,
        marker=Marker(
            size=12,
            line=Line(
                color='rgba(217, 217, 217, 0.14)',
                width=0.5),
            opacity=0.8))
    traces.append(trace)

data = Data(traces)
layout = Layout(showlegend=True,
    scene=Scene(xaxis=XAxis(title='PC1'),
        yaxis=YAxis(title='PC2'),))

fig = Figure(data=data, layout=layout)
#py.offline.plot(fig)

sklearn_pca = sklearnPCA(n_components=2)
Y_sklearn = sklearn_pca.fit_transform(X_std)
traces = []

```

```

for name in ('Muestra1', 'Muestra2', 'Muestra3'):

    trace = Scatter(
        x=Y_sklearn[y==name,0],
        y=Y_sklearn[y==name,1],
        mode='markers',
        name=name,
        marker=Marker(
            size=12,
            line=Line(
                color='rgba(217, 217, 217, 0.14)',
                width=0.5),
            opacity=0.8))
    traces.append(trace)

data = Data(traces)
layout = Layout(xaxis=XAxis(title='PC1', showline=False),
                yaxis=YAxis(title='PC2', showline=False))
fig = Figure(data=data, layout=layout)
py.offline.plot(fig)

```

ANEXO D

Implementación del prototipo

