



UNIVERSIDAD SURCOLOMBIANA
GESTIÓN DE BIBLIOTECAS



CARTA DE AUTORIZACIÓN

CÓDIGO

AP-BIB-FO-06

VERSIÓN

1

VIGENCIA

2014

PÁGINA

1 de 2

Neiva, 31 de mayo de 2023

Señores

CENTRO DE INFORMACIÓN Y DOCUMENTACIÓN

UNIVERSIDAD SURCOLOMBIANA

Ciudad

El (Los) suscrito(s):

Mariana Rojas Romero, con C.C. No. 1.020.827.348,

Edward Reina Hortúa, con C.C. No. 1.000.850.650,

Autor(es) de la tesis y/o trabajo de grado titulado **APLICACIÓN DE ESPECTROSCOPIA INFRARROJA POR TRANSFORMADA DE FOURIER (FTIR) PARA GENERAR UN PATRÓN DE REFERENCIA CON EL FIN DE COMPARAR Y DETECTAR POSIBLES ADULTERACIONES EN MUESTRAS DE CAFÉ TOSTADO Y MOLIDO EN SUPERMERCADOS LOCALES DEL DEPARTAMENTO DEL HUILA Y COLOMBIA**

presentado y aprobado en el año 2023 como requisito para optar al título de

INGENIERO ELECTRÓNICO;

Autorizo (amos) al CENTRO DE INFORMACIÓN Y DOCUMENTACIÓN de la Universidad Surcolombiana para que, con fines académicos, muestre al país y el exterior la producción intelectual de la Universidad Surcolombiana, a través de la visibilidad de su contenido de la siguiente manera:

- Los usuarios puedan consultar el contenido de este trabajo de grado en los sitios web que administra la Universidad, en bases de datos, repositorio digital, catálogos y en otros sitios web, redes y sistemas de información nacionales e internacionales “open access” y en las redes de información con las cuales tenga convenio la Institución.
- Permita la consulta, la reproducción y préstamo a los usuarios interesados en el contenido de este trabajo, para todos los usos que tengan finalidad académica, ya sea en formato Cd-Rom o digital desde internet, intranet, etc., y en general para cualquier formato conocido o por conocer, dentro de los términos establecidos en la Ley 23 de 1982, Ley 44 de 1993, Decisión Andina 351 de 1993, Decreto 460 de 1995 y demás normas generales sobre la materia.
- Continúo conservando los correspondientes derechos sin modificación o restricción alguna; puesto que, de acuerdo con la legislación colombiana aplicable, el presente es un acuerdo jurídico que en ningún caso conlleva la enajenación del derecho de autor y sus conexos.

De conformidad con lo establecido en el artículo 30 de la Ley 23 de 1982 y el artículo 11 de la Decisión Andina 351 de 1993, “Los derechos morales sobre el trabajo son propiedad de los autores” , los cuales son irrenunciables, imprescriptibles, inembargables e inalienables.

Vigilada Mineducación

La versión vigente y controlada de este documento, solo podrá ser consultada a través del sitio web Institucional www.usco.edu.co, link Sistema Gestión de Calidad. La copia o impresión diferente a la publicada, será considerada como documento no controlado y su uso indebido no es de responsabilidad de la Universidad Surcolombiana.



**UNIVERSIDAD SURCOLOMBIANA
GESTIÓN DE BIBLIOTECAS**



CARTA DE AUTORIZACIÓN

CÓDIGO

AP-BIB-FO-06

VERSIÓN

1

VIGENCIA

2014

PÁGINA

2 de 2

EL AUTOR/ESTUDIANTE: Mariana Rojas Romero

EL AUTOR/ESTUDIANTE: Edward Reina Hortúa

Firma: _____

Firma: _____

Vigilada Mineducación

La versión vigente y controlada de este documento, solo podrá ser consultada a través del sitio web Institucional www.usco.edu.co, link Sistema Gestión de Calidad. La copia o impresión diferente a la publicada, será considerada como documento no controlado y su uso indebido no es de responsabilidad de la Universidad Surcolombiana.



TÍTULO COMPLETO DEL TRABAJO: APLICACIÓN DE ESPECTROSCOPÍA INFRARROJA POR TRANSFORMADA DE FOURIER (FTIR) PARA GENERAR UN PATRÓN DE REFERENCIA CON EL FIN DE COMPARAR Y DETECTAR POSIBLES ADULTERACIONES EN MUESTRAS DE CAFÉ TOSTADO Y MOLIDO EN SUPERMERCADOS LOCALES DEL DEPARTAMENTO DEL HUILA Y COLOMBIA

AUTOR O AUTORES:

Primero y Segundo Apellido	Primero y Segundo Nombre
ROJAS ROMERO	MARIANA
REINA HORTÚA	EDWARD

DIRECTOR Y CODIRECTOR TESIS:

Primero y Segundo Apellido	Primero y Segundo Nombre
MOLINA MOSQUERA	JOHAN JULIAN
SARRIA GONZALEZ	JUAN CAMILO

ASESOR (ES):

Primero y Segundo Apellido	Primero y Segundo Nombre
No aplica	No aplica

PARA OPTAR AL TÍTULO DE: INGENIERO ELECTRÓNICO

FACULTAD: INGENIERÍA

PROGRAMA O POSGRADO: INGENIERÍA ELECTRÓNICA

CIUDAD: NEIVA

AÑO DE PRESENTACIÓN: 2023

NÚMERO DE PÁGINAS: 69

TIPO DE ILUSTRACIONES (Marcar con una X):

Diagramas X Fotografías___ Grabaciones en discos___ Ilustraciones en general X Grabados___
Láminas___ Litografías___ Mapas___ Música impresa___ Planos___ Retratos___ Sin ilustraciones___ Tablas
o Cuadros X

Vigilada Mineducación

La versión vigente y controlada de este documento, solo podrá ser consultada a través del sitio web Institucional www.usco.edu.co, link Sistema Gestión de Calidad. La copia o impresión diferente a la publicada, será considerada como documento no controlado y su uso indebido no es de responsabilidad de la Universidad Surcolombiana.

SOFTWARE requerido y/o especializado para la lectura del documento:

MATERIAL ANEXO:

PREMIO O DISTINCIÓN (En caso de ser LAUREADAS o Meritoria):

PALABRAS CLAVES EN ESPAÑOL E INGLÉS:

<u>Español</u>	<u>Inglés</u>	<u>Español</u>	<u>Inglés</u>
1. <u>Adulteraciones</u>	<u>Adulterations</u>	6. <u>Clasificación</u>	<u>Classification</u>
2. <u>Análisis estadístico</u>	<u>Statistical análisis</u>	7. <u>Comercial</u>	<u>Commercial</u>
3. <u>Análisis sensorial</u>	<u>Sensory analysis</u>	8. <u>FTIR</u>	<u>FTIR</u>
4. <u>Café</u>	<u>Coffee</u>	9. <u>Grupos funcionales</u>	<u>Functional groups</u>
5. <u>Calidad</u>	<u>Quality</u>	10. <u>Inteligencia artificial</u>	<u>Artificial intelligence</u>

RESUMEN DEL CONTENIDO: (Máximo 250 palabras)

El objetivo de este proyecto es crear un algoritmo capaz de clasificar componentes del café tostado molido comercial consumido en el departamento del Huila realizando un análisis en el espectrómetro FTIR en donde se obtiene una gráfica por cada muestra de café y con el programa de ingeniería MATLAB se usan distintas herramientas para lograr dicha clasificación. Como objetivo adicional, se analizan las marcas de café comercial por un equipo de catadores semi-entrenados del Centro Surcolombiano de investigación en Café CESURCAFÉ para obtener un análisis más amplio sobre las características de este café. Este trabajo se realiza con el fin de buscar posibles adulteraciones en las muestras de café y suministrar información de la efectividad del método de espectroscopía infrarroja por Transformada de Fourier (FTIR) para encontrar dichas adulteraciones o características.

Para cumplir este propósito, se realiza una encuesta diligenciada por los consumidores locales. Se realizan varios análisis cuantitativos con el espectrómetro FTIR por marca de café que se obtuvo mediante la encuesta, para tener información espectral, es decir, las gráficas de absorbancia de las muestras. Se realiza un algoritmo de clasificación mediante técnicas de aprendizaje supervisado, con el fin de entrenar los modelos, se extraen las medidas estadísticas que representan cada muestra. Para la ejecución del modelo se usan los datos estadísticos, el 80% de resultados para entrenamiento y 20% de resultados para pruebas y tener un modelo entrenado capaz de realizar la clasificación, mostrando las etiquetas en la gráfica obtenida por el FTIR de los componentes del café por marca.

También se realiza una sesión para el análisis sensorial, donde se califica el café según sus características. Esto para tener una visión diferente a la técnica por computador que pueda identificar posibles



DESCRIPCIÓN DE LA TESIS Y/O TRABAJOS DE GRADO

CÓDIGO

AP-BIB-FO-07

VERSIÓN

1

VIGENCIA

2014

PÁGINA

3 de 5

adulteraciones o componentes del café comercial que pueden ser claves para reconocer el café de calidad.

Este proyecto representa una contribución en el conocimiento de los consumidores de dicho café, en donde el análisis investigativo es de gran importancia para demostrar si es recomendable para su consumo habitual. Además de su contribución en el campo tecnológico aplicando metodologías de inteligencia artificial para el desarrollo intelectual del departamento del Huila.

ABSTRACT: (Máximo 250 palabras)

The objective of this project is to create an algorithm capable of classifying components of commercial roasted ground coffee consumed in the department of Huila by performing an analysis in the FTIR spectrometer where a graph is obtained for each coffee sample and with the MATLAB engineering program different tools are used to achieve such classification. As an additional objective, commercial coffee brands are analyzed by a team of semi-trained tasters from the “Centro Surcolombiano de investigación en Café CESURCAFÉ” to obtain a broader analysis of the characteristics of this coffee. This work is carried out with the purpose of searching for possible adulterations in the coffee samples and to provide information on the effectiveness of the Fourier Transform Infrared Spectroscopy (FTIR) method to find such adulterations or



characteristics.

To fulfill this purpose, a survey is carried out and filled out by local consumers. Several quantitative analyses are carried out with the FTIR spectrometer per brand of coffee obtained from the survey, in order to obtain spectral information, that is to say, the absorbance graphs of the samples. A classification algorithm is performed using supervised learning techniques, in order to train the models, the statistical measures that represent each sample are extracted. For the execution of the model the statistical data are used, 80% of results for training and 20% of results for testing and to have a trained model capable of performing the classification, showing the labels on the graph obtained by the FTIR of the coffee components by brand.

A session for sensory analysis is also carried out, where the coffee is graded according to its characteristics. This is to have a different vision to the computer technique that can identify possible adulterations or components of commercial coffee that can be key to recognize quality coffee.

This project represents a contribution to the knowledge of the consumers of such coffee, where the investigative analysis is of great importance to demonstrate if it is recommendable for its habitual consumption. In addition to its contribution in the technological field by applying artificial intelligence methodologies for the intellectual development of the department of Huila.



APROBACION DE LA TESIS

Nombre Jurado: Nelson Gutiérrez Guzmán

Firma:

Nombre Jurado: José de Jesús Salgado Patrón

Firma:

APLICACIÓN DE ESPECTROSCOPÍA INFRARROJA POR TRANSFORMADA
DE FOURIER (FTIR) PARA GENERAR UN PATRÓN DE REFERENCIA CON EL
FIN DE COMPARAR Y DETECTAR POSIBLES ADULTERACIONES EN
MUESTRAS DE CAFÉ TOSTADO Y MOLIDO EN SUPERMERCADOS LOCALES
DEL DEPARTAMENTO DEL HUILA Y COLOMBIA

EDWARD REINA HORTÚA
Código: 20172164250
MARIANA ROJAS ROMERO
Código: 20172162111

UNIVERSIDAD SURCOLOMBIANA
FACULTAD DE INGENIERÍA
PROGRAMA DE INGENIERÍA ELECTRÓNICA
NEIVA - HUILA
2023

APLICACIÓN DE ESPECTROSCOPIA INFRARROJA POR TRANSFORMADA DE
FOURIER (FTIR) PARA GENERAR UN PATRÓN DE REFERENCIA CON EL FIN
DE COMPARAR Y DETECTAR POSIBLES ADULTERACIONES EN MUESTRAS
DE CAFÉ TOSTADO Y MOLIDO EN SUPERMERCADOS LOCALES DEL
DEPARTAMENTO DEL HUILA Y COLOMBIA

EDWARD REINA HORTÚA
MARIANA ROJAS ROMERO

TRABAJO DE GRADO PRESENTADO PARA OPTAR POR EL TÍTULO DE
INGENIERO/A ELECTRÓNICO/A

Director:
Ingeniero Electrónico Johan Julián Molina Mosquera MSc.

Codirector:
Ingeniero Electrónico Juan Camilo Sarria González MSc.

UNIVERSIDAD SURCOLOMBIANA
FACULTAD DE INGENIERÍA
PROGRAMA DE INGENIERÍA ELECTRÓNICA
NEIVA - HUILA
2023

Nota de Aceptación

José de Jesús Salgado Patrón

Jurado

Nelson Gutiérrez Guzmán

Jurado

Neiva, 8 de mayo de 2023

*Dedico este trabajo primeramente a Dios y en especial a mi papá, Rodrigo Rojas Zúñiga, aunque no estés presente en vida, sé que estarás siempre conmigo, por tu apoyo y amor incondicional este y todos mis logros son por ti y para ti.
A mi mamá, Driana Romero Sandoval,
A mi abuelita Margarita Sandoval.*

Mariana.

Dedico el presente trabajo a mis padres, Vilma Yaneth Hortua y Ancizar Reina, ustedes son pieza clave en la persona que soy, los logros que obtengo y obtendré.

Edward.

AGRADECIMIENTOS

Agradezco infinitamente a Dios por permitirme cumplir uno de mis más grandes sueños, porque me dio la fuerza para persistir y terminar con uno de los mejores procesos que he vivido.

A mis papás Driana Romero y Rodrigo Rojas, por ayudarme y apoyarme, por darme las bases para construir mi vida y creer en mí. A mi abuelita Margarita por siempre animarme y escucharme. A mis tías Mayerly y Laura Rojas, por recibirme en su casa y ayudarme para que no me faltara nada, por su paciencia, cariño, colaboración y ánimo. Y a toda mi familia, quienes estuvieron acompañándome en todo este proceso.

Agradezco a Sebastián Gutiérrez, por su incondicionalidad, amor, paciencia, apoyo, por levantarme y acompañarme.

A todos los profesores del programa de ingeniería electrónica, en especial a Julián Molina, por confiar en mí y por su aporte en todo este camino, y a Camilo Sarria, por ser pieza clave y parte de este trabajo, por su compromiso y apoyo como nuestro codirector del proyecto de grado.

Mariana.

Agradezco en primer lugar a mis padres Ancizar Reina y Vilma Yaneth Hortua, mi padre es el más grande ejemplo que tengo de humildad y sencillez, mi madre es la mujer más fuerte, amable y alegre, gracias a ellos conozco el valor del sacrificio, desde siempre han buscado lo mejor para mí y para que pueda cumplir mis metas y a mi familia, en la cual he sentido la unión y por su presencia en mi camino.

A mis amigos, en especial a Juan Esteban Narváez y Mariana Rojas, los cuales han estado presentes a lo largo de estos años y han hecho de la universidad un camino más agradable de recorrer.

A mis profesores del programa de Ingeniería Electrónica, no solo por los conocimientos sino por ser el ejemplo del profesional que quiero ser, su entrega, empeño y dedicación, al Centro Surcolombiano de Investigación en Café CESURCAFE por su colaboración, a Camilo Sarria, quien desde el primer momento en el que iniciamos el proyecto de grado ha estado presente, por guiarnos y siempre tener la mejor disposición y actitud.

Edward.

TABLA DE CONTENIDO

	Pág.
0. INTRODUCCIÓN	15
1. OBJETIVOS	17
1.1 OBJETIVO GENERAL	17
1.2 OBJETIVOS ESPECÍFICOS	17
2. FUNDAMENTOS BÁSICOS	18
2.1 ADULTERACIONES DE LOS ALIMENTOS	18
2.2 CAFÉ CALIDAD (ESPECIALES VS COMERCIALES)	19
2.3 DIFERENCIA ENTRE CAFÉ DE ESPECIALIDAD Y CAFÉ COMERCIAL	20
2.4 EVALUACIÓN SENSORIAL	21
2.5 ESPECTROSCOPIA	21
2.5.1 Principios de la espectroscopía infrarroja	21
2.5.2 Espectroscopía infrarroja por transformada de Fourier	24
2.6 GRUPOS FUNCIONALES	25
2.7 INTELIGENCIA ARTIFICIAL	26
2.7.1 Machine Learning	27
2.7.2 Métodos de clasificación – MATLAB	27
2.8 MEDIDAS ESTADÍSTICAS	28
2.8.1 Medidas de centralización	29
2.8.2 Medidas de dispersión	29

2.8.3	Medidas de posición o cuantiles	30
2.8.4	Medidas estadísticas de forma	30
3.	MATERIALES Y MÉTODOS	32
3.1	MUESTRAS DE CAFÉ	33
3.2	ANÁLISIS DEL ESPECTRO INFRARROJO CON FTIR	33
3.3	ANÁLISIS SENSORIAL	34
3.4	DESARROLLO DEL ALGORITMO DE CLASIFICACIÓN	35
3.5	ANÁLISIS ESTADÍSTICO	37
4.	RESULTADOS Y DISCUSIÓN	38
4.1	RESULTADOS DE ENCUESTA VIRTUAL	38
4.2	RESULTADOS DE ANÁLISIS SENSORIAL	39
4.3	RESULTADOS ANÁLISIS FTIR	40
4.4	RESULTADOS ALGORITMO DESARROLLADO	41
4.4.1	Arquitectura clasificadores de redes neuronales - MATLAB.	48
4.5	RESULTADOS ESTADÍSTICOS	49
5.	CONCLUSIONES	53
6.	RECOMENDACIONES	56
	BIBLIOGRAFÍA	57
	ANEXOS	62

LISTA DE TABLAS

	Pág.
Tabla 1. Entrenamiento modelos de clasificación nivel de cafeína.	42
Tabla 2. Entrenamiento modelos de clasificación tipo de café.	45

LISTA DE GRÁFICAS

	Pág.
Grafica 1. Resultados encuesta virtual	38
Gráfica 2. Resultados Análisis Sensorial	39
Gráfica 3. Niveles de absorbancia por longitud de onda. Análisis FTIR	40

LISTA DE FIGURAS

	Pág.
Figura 1. Espectro electromagnético	22
Figura 2. Clasificación espectro infrarrojo	23
Figura 3. Interacción de la radiación electromagnética con la materia	24
Figura 4. Grados de asimetría	31
Figura 5. Curtosis en datos	31
Figura 6. Diseño experimental para el desarrollo del proyecto	32
Figura 7. Espectrómetro FTIR Cary 630	34
Figura 8. Escala evaluación análisis sensorial	35
Figura 9. Diagrama de flujo del algoritmo de clasificación	37
Figura 10. Matriz de confusión.	41
Figura 11. Resumen entrenamiento.	43
Figura 12. Matriz de confusión modelo clasificador nivel cafeína.	44
Figura 13. Curva ROC modelo clasificador nivel de cafeína.	44
Figura 14. Resumen entrenamiento.	46
Figura 15. Matriz de confusión modelo clasificador tipo de café.	46
Figura 16. Curva ROC modelo clasificador nivel de cafeína.	47
Figura 17. Estructura clasificador red neuronal	48
Figura 18. Compuestos del FTIR característicos de las muestras de café.	49

LISTA DE ANEXOS

	Pág.
Anexo A. Algoritmo de clasificación.	62
Anexo B. Código de la función encargada de calcular las medidas estadísticas.	64

GLOSARIO

IA: Inteligencia Artificial, todo sistema o máquina que imitan la inteligencia humana para realizar tareas y que pueden mejorar iterativamente mediante la información que recopilan.

IDE: Un entorno de desarrollo integrado, es una aplicación de software que ayuda a los programadores a desarrollar código de software de manera eficiente. Aumenta la productividad de los desarrolladores al combinar capacidades como editar, crear, probar y empaquetar software en una aplicación fácil de usar.

FTIR: Espectroscopia Infrarroja por Transformada de Fourier, es una técnica empleada para obtener la huella dactilar molecular de una muestra biológica que absorbe la onda infrarroja de acuerdo con los enlaces químicos y estructurales de las moléculas.

RESUMEN

El objetivo de este proyecto es crear un algoritmo capaz de clasificar componentes del café tostado molido comercial consumido en el departamento del Huila realizando un análisis en el espectrómetro FTIR en donde se obtiene una gráfica por cada muestra de café y con el programa de ingeniería MATLAB se usan distintas herramientas para lograr dicha clasificación. Como objetivo adicional, se analizan las marcas de café comercial por un equipo de catadores semi-entrenados del Centro Surcolombiano de investigación en Café CESURCAFÉ para obtener un análisis más amplio sobre las características de este café. Este trabajo se realiza con el fin de buscar posibles adulteraciones en las muestras de café y suministrar información de la efectividad del método de espectroscopía infrarroja por Transformada de Fourier (FTIR) para encontrar dichas adulteraciones o características.

Para cumplir este propósito, se realiza una encuesta diligenciada por los consumidores locales. Se realizan varios análisis cuantitativos con el espectrómetro FTIR por marca de café que se obtuvo mediante la encuesta, para tener información espectral, es decir, las gráficas de absorbancia de las muestras. Se realiza un algoritmo de clasificación mediante técnicas de aprendizaje supervisado, con el fin de entrenar los modelos, se extraen las medidas estadísticas que representan cada muestra. Para la ejecución del modelo se usan los datos estadísticos, el 80% de resultados para entrenamiento y 20% de resultados para pruebas y tener un modelo entrenado capaz de realizar la clasificación, mostrando las etiquetas en la gráfica obtenida por el FTIR de los componentes del café por marca.

También se realiza una sesión para el análisis sensorial, donde se califica el café según sus características. Esto para tener una visión diferente a la técnica por computador que pueda identificar posibles adulteraciones o componentes del café comercial que pueden ser claves para reconocer el café de calidad.

Este proyecto representa una contribución en el conocimiento de los consumidores de dicho café, en donde el análisis investigativo es de gran importancia para demostrar si es recomendable para su consumo habitual. Además de su contribución en el campo tecnológico aplicando metodologías de inteligencia artificial para el desarrollo intelectual del departamento del Huila.

PALABRAS CLAVE: Adulteraciones, análisis estadístico, análisis sensorial, café, calidad, clasificación, comercial, FTIR, grupos funcionales, inteligencia artificial, muestras.

ABSTRACT

The objective of this project is to create an algorithm capable of classifying components of commercial roasted ground coffee consumed in the department of Huila by performing an analysis in the FTIR spectrometer where a graph is obtained for each coffee sample and with the MATLAB engineering program different tools are used to achieve such classification. As an additional objective, commercial coffee brands are analyzed by a team of semi-trained tasters from the “**Centro Surcolombiano de investigación en Café CESURCAFÉ**” to obtain a broader analysis of the characteristics of this coffee. This work is carried out with the purpose of searching for possible adulterations in the coffee samples and to provide information on the effectiveness of the Fourier Transform Infrared Spectroscopy (FTIR) method to find such adulterations or characteristics.

To fulfill this purpose, a survey is carried out and filled out by local consumers. Several quantitative analyses are carried out with the FTIR spectrometer per brand of coffee obtained from the survey, in order to obtain spectral information, that is to say, the absorbance graphs of the samples. A classification algorithm is performed using supervised learning techniques, in order to train the models, the statistical measures that represent each sample are extracted. For the execution of the model the statistical data are used, 80% of results for training and 20% of results for testing and to have a trained model capable of performing the classification, showing the labels on the graph obtained by the FTIR of the coffee components by brand.

A session for sensory analysis is also carried out, where the coffee is graded according to its characteristics. This is to have a different vision to the computer technique that can identify possible adulterations or components of commercial coffee that can be key to recognize quality coffee.

This project represents a contribution to the knowledge of the consumers of such coffee, where the investigative analysis is of great importance to demonstrate if it is recommendable for its habitual consumption. In addition to its contribution in the technological field by applying artificial intelligence methodologies for the intellectual development of the department of Huila.

KEY WORDS: Adulterations, statistical analysis, sensory analysis, coffee, quality, classification, commercial, FTIR, functional groups, artificial intelligence, samples.

0. INTRODUCCIÓN

El café es uno de los productos primarios más valiosos y la segunda bebida más consumida a nivel mundial, y, además, juega un papel fundamental para la economía de muchos países. Hoy en día, el café en el comercio internacional tiene un amplio protagonismo, ya que en el mundo se consumen un aproximado de 3 millones de tazas al día. Colombia es el tercer mayor consumidor de café en América Latina, con pronóstico de 2 millones de sacos de 60 kilogramos en el año 2021.¹ Además de ser el mayor productor de café arábigo suave lavado con un total de 12,6 millones de sacos de 60 kilogramos de café verde exportado en el año 2021. La calidad en la bebida del café se atribuye a la confluencia de factores como la variedad de granos (café arábica y café robusta), las mezclas con granos defectuosos (negro, verde y agrio, BGS) y/o materia extraña a los granos de café (piedras, maíz, entre otros), además de los diferentes grados de tueste (ligero, medio y oscuro), que en conjunto pueden alterar las propiedades organolépticas de la bebida y, en consecuencia, en su calidad.

Existen formas mediante las cuales se logra identificar posibles adulteraciones que hacen que el café presente baja calidad, sabores desagradables y hasta componentes perjudiciales para la salud. Los catadores son personas entrenadas para clasificar el café en base a sus sentidos organolépticos, con el inconveniente de que poseen un coste relativamente elevado de preparación y mano de obra, y además incluye la posibilidad de clasificación errónea debido a la dificultad del análisis a gran escala y, principalmente, a la subjetividad generada por los mismos. Por lo tanto, es necesario buscar técnicas que permitan un análisis rápido y fiable, y que demuestren también posibles componentes que desnaturalizan el café de consumo.

La espectroscopia infrarroja por transformada de Fourier (FTIR) es una técnica ampliamente difundida en la rutina de los laboratorios para el análisis directo de muestras de alimentos, que además de aplicarse con éxito para medir la calidad del café, podría ser una herramienta útil para evaluar la calidad de las marcas comerciales de café molido tostado. El presente estudio tiene como objetivo

¹ STATISTA, Consumo de café en Colombia en la temporada de cultivo 2022, por tipo. [sitio web]. Colombia, 2022, Disponible en: <https://es.statista.com/estadisticas/1284331/colombia-consumo-de-cafe-por-tipo/#:~:text=Colombia%3A%20consumo%20de%20caf%C3%A9%202021%2D2021%2C%20por%20tipo&text=Se%20pronostic%C3%B3%20que%20el%20consumo,del%20pa%C3%ADs%20sudamericano%20ese%20a%C3%B1o>

encontrar si el café que se consume por parte de los consumidores posee una buena o mala calidad o si posee también adulteraciones realizando un análisis con el espectrómetro y con la clasificación de componentes presentes en las muestras de café con un algoritmo de clasificación en MATLAB.

1. OBJETIVOS

1.1 OBJETIVO GENERAL

Emplear el método de análisis FTIR en muestras de café tostado molido comercializadas en el departamento del Huila para medir su calidad y posibles adulteraciones.

1.2 OBJETIVOS ESPECÍFICOS

- Selección y preparación de las muestras de café comercializadas en el Huila.
- Obtener información espectral a partir del análisis FTIR de las muestras de café por medio de una metodología que permita generar un patrón de referencia del café tostado.
- Realizar un análisis sensorial a las muestras de café por medio de un panel de catadores entrenados con el fin de obtener descriptores y atributos sensoriales de las muestras analizadas.
- Verificar y validar los resultados obtenidos en los componentes espectrales de las muestras analizadas mediante una técnica de clasificación estadística e Inteligencia Artificial.

2. FUNDAMENTOS BÁSICOS

2.1 ADULTERACIONES DE LOS ALIMENTOS

Los alimentos, sean productos naturales o procesados, deben contribuir a las funciones normales del organismo, a la salud. Los alimentos son matrices complejas y heterogéneas formadas por diversos componentes; la evaluación de calidad de estos, características sensoriales, valor nutricional y estabilidad son el principal interés de quien trabaja en la industria alimentaria.²

Existe una gran variedad de peligros asociados a los alimentos, y tanto los ingredientes como los aditivos presentan una compleja problemática porque, generalmente, no tienen una característica funcional única o que pueda evaluarse visualmente, lo que no permite su discriminación de otros ingredientes similares o de posibles adulterantes.³ Ejemplos de adulteraciones alimentarias de gran impacto mundial incluyen, en Inglaterra (2005) la presencia de colorantes no autorizados en chile en polvo que fueron utilizados en salsas Worcestershire, en China (2007) el aumento del contenido proteico en leche con el agregado de gluten de trigo y melamina lo cual derivó en muertes infantiles y miles de hospitalizaciones, en India (2012) leche adulterada con detergente, grasa y urea. Otras adulteraciones de alto impacto en la salud publica fueron la presencia de óxido de plomo (Pb_3O_4) en pimentones y pimienta de cayena y la contaminación de carne de cerdo con dioxinas (Irlanda, 2012) que constituyo la mayor crisis que afectó la economía y la salud nacional. Recientemente, en Europa, se impidió la importación de miel asiática por la presencia ilegal de antibióticos y también de mieles fabricadas a partir de edulcorantes artificiales, y en 2013 en Taiwán, se detectaron contaminaciones de aceite de oliva con aceites baratos de diferentes orígenes y clorofilina, así como de bebidas alcohólicas, leche y arroz.⁴ Entre las materias primas sujetas a un mayor número de adulteraciones se encuentran especias, aceites comestibles, miel, leche y productos lácteos, frutas y jugos de fruta, café, harina, productos cárnicos.

En la Ley 09 de 1979, en su artículo 304 del Ministerio de Salud y Protección social establece que los alimentos que no se consideren aptos para el consumo humano

² OPS, Clasificación de los alimentos y sus implicaciones en la salud, [sitio web]. Disponible en: <https://www3.paho.org/ecu/1135-clasificacion-alimentos-sus-implicaciones-salud.html>

³ COZZOLINO, Daniel, 2015. Foodomics and infrared spectroscopy: from compounds to functionality. Current Opinion in Food Science. DOI: <https://doi.org/10.1016/j.cofs.2015.05.003>

⁴ LOHUMI, Santosh, et al, 2015. A review of vibrational spectroscopic techniques for the detection of food authenticity and adulteration. Trends in Food Science & Technology. DOI: <https://doi.org/10.1016/j.tifs.2015.08.003>.

los alimentos o bebidas alteradas, adulterados, falsificados, contaminados o los que por otras características anormales puedan afectar la salud del consumidor.

Además, en su artículo 426 de dicha ley, aclara que, en cualquier tipo de alimento o bebida, la presencia de antibióticos u otras sustancias no permitidas, será causal de decomiso del producto.

2.2 CAFÉ CALIDAD (ESPECIALES VS COMERCIALES)

Muchos perciben el café de especialidad como costoso, y la mayoría de los bebedores habituales de café solo están familiarizados con el café comercial. Es posible que algunos ni siquiera se den cuenta de que existe una diferencia entre el café comercial y el de especialidad. Piensan que todo el café es amargo y ácido.

La mayor diferencia entre los dos radica en la forma en que se cosechan y procesan para el consumo. El café de especialidad implica más cuidado y atención, mientras que el café comercial suele producirse en masa.

El café de especialidad, también conocido como café gourmet, es el café de mayor grado en términos de calidad y sabor. Esto implica controles estrictos a lo largo de la cadena de suministro para garantizar la máxima calidad en cada paso. El café de especialidad ha ganado visibilidad gracias a la tercera ola de café.

La tercera ola del café comenzó a principios de la década de 2000, mostrando un cambio en la forma en que se presenta el café. Ahora, muchos ven el café de especialidad como artesanal, al mismo nivel que los vinos y quesos caros. Esto creó una demanda de café de alta calidad tanto en grano como en taza. Ahora no se trata solo de cómo sabe el café. También se trata de cómo se cultiva, cosecha y procesa. El café de especialidad proviene de los países productores de café entre los Trópicos de Cáncer y Capricornio, principalmente de América del Sur y Central, África y Asia.⁵

La calidad del café se puede cuantificar subjetivamente mediante un **“Puntaje de café”**. En un nivel más técnico, la Asociación de Cafés Especiales (SCA) tiene una escala de calificación detallada para juzgar si el café es de especialidad o no. Muchos amantes del vino pueden comparar esta escala de café con la escala de

⁵ VALERIO, Yker, Café comercial versus café de especialidad: ¿cuál es la diferencia?. [en línea]. Era of We, Diciembre 11, 2021. Disponible en: <https://www.eraofwe.com/coffee-lab/es/articles/cafe-comercial-versus-cafe-de>.

100 puntos de Wine Spectator. Un café con un puntaje mayor o igual a 80 sobre 100 puntos se considera un café de especialidad, y el nivel superior obtiene más de 90 puntos (Café Presidencial). Para que un café sea considerado café de especialidad debe cumplir con al menos estos tres criterios mínimos:

- El café debe cosecharse a mano.
- Obtiene más de 80 puntos en la escala de calificación de la SCA.
- Debe tener menos de 5 defectos por cada 12 onzas.⁶

El café comercial es cualquier café producido en masa vendido por grandes marcas, tostado y vendido en grandes cantidades se llama café comercial. La mayoría de las personas en Colombia conocen el café a través de las marcas de café comercial. Por lo general, se presenta en forma de café instantáneo y tostado molido.⁷

2.3 DIFERENCIA ENTRE CAFÉ DE ESPECIALIDAD Y CAFÉ COMERCIAL

Existen varias diferencias entre el café comercial y el de especialidad:

Tipo de granos: Hay dos tipos principales de variedades de café: Arábica y Robusta. Arábica es el tipo dominante, que comprende el 60% del mercado. Los granos de Arábica tienen más valor ya que son menos amargos y tienen un paladar más amplio de sabores. Los frijoles robusta son más resistentes mientras crecen, pero también son más amargos.

El café de especialidad es exclusivamente Arábica, mientras que la mayoría del café comercial es completamente Robusta o una mezcla de Arábica y Robusta.

Costo de producción: Uno de los puntos de diferencia más obvios entre el café comercial y el de especialidad es el precio. El café de especialidad es más caro debido a que requiere más mano de obra para su recolección y procesamiento. Además, como es Arábica, es más difícil de cultivar. Por lo tanto, los rendimientos de la cosecha son más bajos, lo que aumenta aún más el precio.

El café de especialidad también cuesta más debido al cuidado que se pone en cada paso del procesamiento, desde el tostado hasta el envasado. Todo se hace en lotes

⁶ Ibíd.

⁷ Ibíd.

pequeños, y con más intervención humana, lo que lo encarece. El beneficio de esto es la transparencia en todos los niveles y la alta calidad del producto final.

Origen de los granos: Los países productores donde se cultiva la mayor parte del café del mundo, se encuentran entre los Trópicos de Cáncer y Capricornio. La mayoría de los cafés especiales se obtienen de esta región. El origen de su café afecta su sabor y calidad en general. Factores como el clima, la altitud, el tipo de suelo y otras condiciones de crecimiento afectan el sabor final.

El origen es importante para muchos consumidores, especialmente para los más involucrados en el café. El café de especialidad de origen único es un mercado enorme donde el origen es el factor fundamental. El café comercial, por otro lado, a menudo se obtiene de múltiples mercados para reducir el costo. Sin embargo, esto da como resultado un sabor mixto y un perfil de sabor poco claro.⁸

2.4 EVALUACIÓN SENSORIAL

La evaluación o análisis sensorial es una disciplina científica usada para discriminar, analizar, interpretar y medir aquellas características que se perciben por los sentidos del ser humano, es decir, es la evaluación de la apariencia, olor, aroma, textura y sabor de un alimento o materia prima.

Este tipo de análisis comprende un conjunto de técnicas para la medida precisa de las respuestas humanas a los alimentos y minimiza los potenciales efectos de desviación que la identidad de la marca y otras informaciones pueden ejercer sobre el juicio del consumidor. Es decir, intenta aislar las propiedades sensoriales u organolépticas de los alimentos o productos en sí mismos y aporta información muy útil para su desarrollo o mejora.⁹

2.5 ESPECTROSCOPÍA

2.5.1 Principios de la espectroscopia infrarroja

La luz tiene la particularidad de comportarse como partícula, que se le conoce como fotón, y como onda, la cual viaja en el vacío una velocidad de $2.99792458 \times 10^8 \frac{Km}{s}$ y la velocidad disminuye cuando se propaga en otros medios. La luz como onda

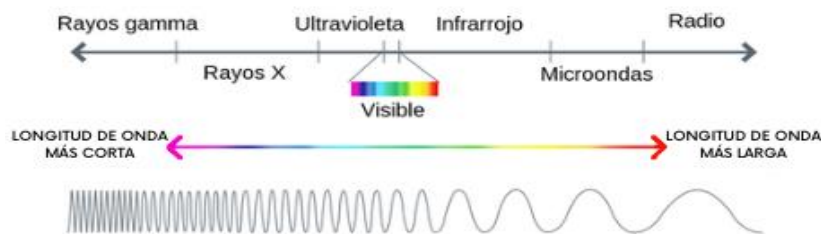
⁸ Ibíd.

⁹ GARCÍA AHUED, Maricela, 2014. Análisis sensorial de alimentos. Padi Boletín Científico De Ciencias Básicas E Ingenierías Del ICBI. DOI: <https://doi.org/10.29057/icbi.v2i3.533>

presenta la característica de tener una parte eléctrica y otra magnética, es decir que la luz se propaga generando una perturbación oscilante de la naturaleza electromagnética. Una de las características más importantes de una onda electromagnética es que puede transportar energía de un lugar a otro. En 1900, el físico alemán, Max Planck afirmó que la radiación era emitida en forma de cuantos, paquetes de energía de frecuencia determinada, a los que más tarde Albert Einstein los llamaría fotones. La energía transportada por una onda electromagnética es función al cambio de su longitud de onda, es decir, a menor longitud de onda más cantidad de energía posee y viceversa.¹⁰

De acuerdo con la energía que transportan y/o a la magnitud de la longitud de onda de las ondas electromagnéticas, estas se han organizado en el denominado espectro electromagnético. En donde se puede observar, que existen ondas que se encuentran desde los nanómetros hasta los kilómetros, por lo tanto, también hay distintos tipos de niveles energéticos, los cuales han sido clasificados como: rayos gamma, rayos X, ultravioleta, visible, infrarrojo o microondas.

Figura 1. Espectro electromagnético



Fuente: elaboración propia, con base en Labster Theory, Espectro electromagnético. [sitio web]. Disponible en: <https://theory.labster.com/electromagnetic-spectrum-es/>

La representación de las propiedades energéticas de las ondas en función a su tamaño: a menor longitud de onda, mayor energía, a mayor longitud de onda, menor energía.

¹⁰ MONDRAGÓN CORTEZ, Pedro. Espectroscopia de infrarrojo para todos. Centro de Investigación y Asistencia en Tecnología y Diseño del Estado de Jalisco, A. C., 2017. ISBN: 978-607-97548-4-6

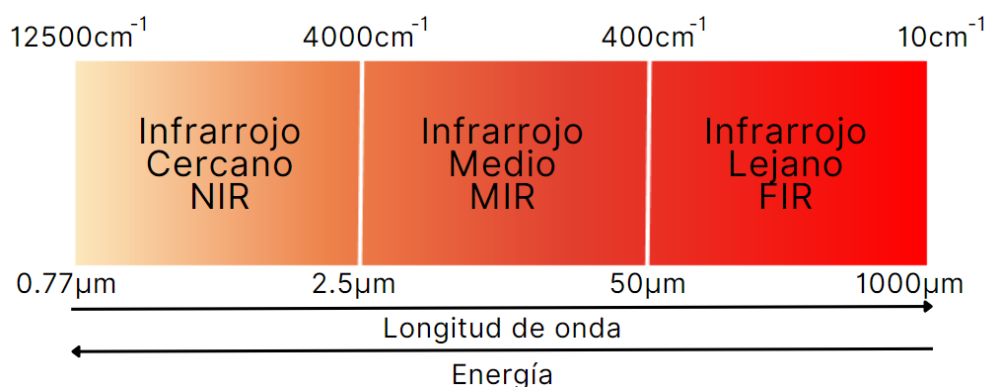
La radiación infrarroja se encuentra en el intervalo de 13000cm^{-1} y 10cm^{-1} , entre 0.77 y $100\mu\text{m}$. En función de las distintas aplicaciones de la espectroscopia de infrarrojo por conveniencia la radiación infrarroja se clasifica en tres intervalos:

Infrarrojo cercano: Se encuentra en el intervalo de $12500\text{cm}^{-1} - 4000\text{cm}^{-1}$ ($0.77\mu\text{m} - 2.5\mu\text{m}$)

Infrarrojo medio: Se encuentra en el intervalo de $4000\text{cm}^{-1} - 400\text{cm}^{-1}$ ($2.5\mu\text{m} - 50\mu\text{m}$)

Infrarrojo lejano: Se encuentra en el intervalo de $400\text{cm}^{-1} - 10\text{cm}^{-1}$ ($50\mu\text{m} - 1000\mu\text{m}$).¹¹

Figura 2. Clasificación Espectro Infrarrojo



Fuente: elaboración propia, con base en TÉLLEZ MESA, Clara. Aplicaciones de la espectroscopía infrarroja en el análisis de alimentos. Facultad de farmacia universidad de Sevilla, 2019. 38 p.

La espectroscopia de infrarrojo es actualmente una de las técnicas analíticas más utilizadas en todo el mundo, por analistas, científicos y estudiantes, todo ello en un amplio campo de aplicaciones, por ejemplo, en la medicina, alimentos, ambiental, ciencia de los materiales, biotecnología, ciencias forenses, etc.

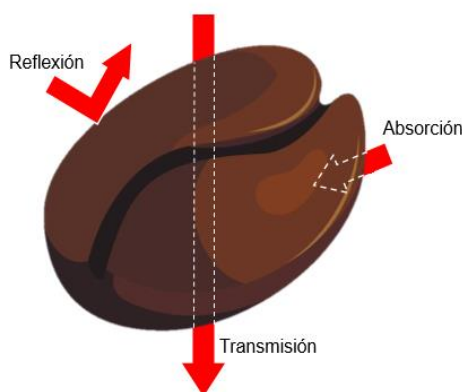
La espectroscopia estudia la interacción entre la radiación electromagnética y la materia. La radiación incidente sobre la materia es diferente a la saliente por efecto de la interacción. El resultado de la interacción proporciona información útil sobre la sustancia involucrada con relación a su estructura molecular.

¹¹ Ibí.

La espectrometría infrarroja se basa en el hecho de que los enlaces químicos de las sustancias tienen frecuencias de vibración específicas, que corresponden a los niveles de energía de la molécula. Estas frecuencias dependen de la forma de la superficie de energía potencial de la molécula, la geometría molecular, las masas atómicas y, posiblemente, el acoplamiento vibracional.¹²

La radiación electromagnética puede interactuar con la materia de diversas formas: reflexión, absorción y transmisión. La luz puede reflejarse, absorberse, transmitirse, o en una combinación de ellas en la materia.

Figura 3. Interacción de la radiación electromagnética con la materia



Fuente: elaboración propia

Una interacción electromagnética puede llevarse a cabo en la materia en cualquiera de sus estados físicos. La radiación electromagnética una vez interactúa con la materia, se pueden presentar ciertos fenómenos a nivel atómico y molecular.¹³

2.5.2 Espectroscopía infrarroja por transformada de Fourier

En primer lugar, a través de esta técnica analítica de espectroscopia se permite identificar la familia química que posee una sustancia en específico. Su principal función estriba en que “FTIR mide la absorbancia de la luz infrarroja de una muestra

¹² Ibíd.

¹³ Ibíd.

y genera un espectro basado en **los grupos funcionales** del material".¹⁴ (Eurofins, 2022).

Por tanto, la FTIR consiste en un método empleado para determinar las estructuras específicas de aquellas moléculas las cuales cumplan la característica base como lo es la absorción de la radiación infrarroja de acuerdo con la vibración molecular de la misma, ya que dicha característica es en realidad única para cada molécula en particular; es decir, se está en presencia de una propiedad física única, un único registro de identificación para la molécula.

Es por esto por lo que, la FTIR es considerada como una técnica funcional y potencialmente versátil, ya que a través de esta es posible el estudio y la determinación de cualquier tipo de muestra, indiferentemente del estado en que la misma se encuentre -líquidos, gel, gases, fibras, superficies, entre otros; la misma "es una técnica espectroscópica poderosa y versátil, que ofrece selectividad y especificidad. Por ello, se ha convertido en una herramienta muy útil para el análisis de innumerables sustancias en muestras complejas, sin la necesidad de tratamientos complejos previos a la muestra".¹⁵ (Ayala, C et al, 2009, p.239).

Este tipo de técnica es utilizada en cualquier estado, por lo que llega a ser útil en industrias como la farmacéutica, médica, química clínica; en procesos industriales para el control de procesos; manufactura, la industria química, ambiental, agroquímica e incluso en la industria alimentaria. Esto, debido principalmente al carácter versátil de la técnica, ya que, una vez establecida la calibración del equipo, el análisis es rápido, promoviendo el que no se utilicen los reactivos y un bajo mantenimiento del equipo.

Con respecto al sector de materiales, la importancia de su aplicación estriba principalmente en que es posible la caracterización de diversas mezclas complejas de materiales, esto con el fin de garantizar la identificación de diversos grupos funcionales orgánicos.

¹⁴ Tecno Soluciones, 2022. Espectroscopia infrarroja por transformada de Fourier y sus aplicaciones en la industria. [sitio web]. Disponible en: <https://tecnosolucionescr.net/blog/635-espectroscopia-infrarroja-por-transformada-de-fourier-y-sus-aplicaciones-en-la-industria>

¹⁵ AYALA, Carlos. Determination of atenolol in pharmaceutical dosages by Fourier transform infrared spectrometry (FTIR). Revista Técnica de la Facultad de Ingeniería Universidad del Zulia. ISSN: 0254-0770

2.6 GRUPOS FUNCIONALES

Las propiedades de los compuestos de carbono dependen del arreglo de sus cadenas y tipos de átomos a los que están unidos, esto es, a su estructura.

Un grupo funcional es un átomo o un arreglo de átomos que siempre reaccionan de una forma determinada; además, es la parte de la molécula responsable de su comportamiento químico ya que le confiere propiedades características. Muchos compuestos orgánicos contienen más de un grupo funcional. Hay gran cantidad de grupos funcionales.

Algunos ejemplos de grupos funcionales pueden ser: ácidos carboxílicos, aldehídos, cetonas, aminas, etc. Estos se pueden encontrar estructurados en los alimentos como lípidos, carbohidratos, cafeína, polisacáridos, azúcares, alcaloides, etc.

En el análisis con el espectro infrarrojo se pueden identificar los grupos funcionales de distintos alimentos según el valor de longitud de onda y la amplitud de la absorbancia.

2.7 INTELIGENCIA ARTIFICIAL

La inteligencia artificial, es la capacidad que tienen los ordenadores o máquinas para realizar tareas y/o tomar decisiones propias de seres humanos, con la ventaja de que dichas maquinas no necesitan descansar, pueden analizar grandes cantidades de información en un corto tiempo y la cantidad de errores con respecto a un ser humano, es significativamente menor.¹⁶

La inteligencia no es un formato único, sino una amplia variedad de diversas capacidades las cuales permiten, a partir de datos o situaciones previas, procesar determinada información y tomar decisiones, de igual forma, la inteligencia artificial puede utilizar una gran cantidad de técnicas con el fin de realizar determinada labor; el hecho de que un programa computacional, pueda aprender a tomar decisiones es algo muy importante y se debe estar consciente de esto, ya que las aplicaciones basadas en inteligencia artificial están siendo cada vez más utilizadas con el fin de

¹⁶ PETTERI ROUHIAINEN, Lasse, 2018. Inteligencia artificial. 101 cosas que debes saber hoy sobre nuestro futuro. España: Editorial Planeta, S.A., 2018. ISBN: 978-84-17568-08-5. [En línea]. Disponible en: https://static0planetadelibroscom.cdnstatics.com/libros_contenido_extra/40/39308_Inteligencia_artificial.pdf

favorecerse de las mejoras que trae y de la eficiencia que brinda en gran cantidad de procesos.¹⁷

2.7.1 Machine Learning

Es una disciplina del campo de la inteligencia artificial, el cual, mediante técnicas y algoritmos, se entrenan modelos con el fin de que aprendan de los datos observados, analizando datos masivos, identificando patrones, variables con gran relevancia, inferir, aplicar reglas y elaborar predicciones, dichas técnicas permiten a los ordenadores cumplir con tareas específicas de manera autónoma, es decir sin ser programados.¹⁸

Machine Learning utiliza dos tipos de técnicas, aprendizaje supervisado, mediante el cual se entrenan modelos con datos conocidos, con el fin de predecir resultados futuros, la segunda técnica es aprendizaje no supervisado, en el cual los modelos identifican patrones o estructuras intrínsecas de los datos, algunos de los tipos de algoritmos de Machine Learning, son algoritmos de regresión, bayesianos, de agrupación, de árbol de decisión, redes neuronales, reducción de dimensión, y aprendizaje profundo.¹⁹

2.7.2 Métodos de clasificación – MATLAB

“MATLAB es una plataforma de programación y cálculo numérico utilizada por millones de ingenieros y científicos para analizar datos, desarrollar algoritmos y crear modelos”.²⁰ Es un entorno de desarrollo integrado o IDE que cuenta con un lenguaje de programación propio el cual es el lenguaje M, en MATLAB se pueden generar funciones, representar datos, manipular matrices, crear interfaces de usuario o GUI e implementar algoritmos, todas estas herramientas permiten

¹⁷ A. BODEN, Margaret, 2016. Inteligencia artificial. Madrid:Turner Publicaciones S.L., 2017, ISBN: 978-84-16714-22-3. [En línea]. Disponible en:

https://books.google.com.co/books?hl=es&lr=&id=LCnYDwAAQBAJ&oi=fnd&pg=PT3&dq=Inteligencia+artificial+PDF&ots=drWpAVcMq6&sig=7CFI1P_mlxgp88w7KXoAUzjYq5A#v=onepage&q&f=false

¹⁸ Iberdrola, S.A. Qué es el 'machine learning'. [sitio web]. Disponible en:

[https://www.iberdrola.com/innovacion/machine-learning-aprendizaje-automatico#:~:text=El%20Machine%20Learning%20es%20una,elaborar%20predicciones%20\(an%C3%A1lisis%20predictivo\).](https://www.iberdrola.com/innovacion/machine-learning-aprendizaje-automatico#:~:text=El%20Machine%20Learning%20es%20una,elaborar%20predicciones%20(an%C3%A1lisis%20predictivo).)

¹⁹ MathWorks, Inc. Machine learning. [sitio web]. Disponible en:

<https://la.mathworks.com/discovery/machine-learning.html>

²⁰ MathWorks, Inc. Matlab. [sitio web]. Disponible en:

<https://la.mathworks.com/products/matlab.html>

discriminar información en base al análisis de datos, además de tener compatibilidad con programas desarrollados en otros lenguajes y comunicarse con diferentes dispositivos hardware.²¹

Mediante el uso de Machine Learning, se pueden realizar métodos de clasificación de datos, los cuales pueden ser métodos de clasificación supervisados o no supervisados, particularmente, los métodos de clasificación supervisados se entrenan mediante un grupo de datos los cuales se categorizan y se etiquetan de manera manual de acuerdo con una pauta específica, de esta manera se conforma una clase. La finalidad de los clasificadores ya entrenados es decidir en qué categoría debe ir un dato nuevo, teniendo en cuenta un conjunto de características propias del dato en cuestión, como pueden ser medidas estadísticas.²²

La aplicación “Classification Learner” de MATLAB, permite entrenar modelos de clasificación de datos, dicha aplicación se basa en el proceso de Machine Learning supervisado, llevando a cabo un entrenamiento automatizado con el fin de encontrar el mejor modelo de clasificación, dichos modelos incluyen árboles de decisión, análisis discriminantes, máquinas de soporte vectorial (SVM), regresión logística, vecinos más próximos (KNN), algoritmos probabilísticos (Naive Bayes), aproximación de kernel, ensemble y clasificación de redes neuronales.²³

2.8 MEDIDAS ESTADÍSTICAS

La estadística, se ha convertido en una de las herramientas más importantes en la investigación, ya que permite recolectar, organizar, representar, interpretar y analizar la información, permite comunicar de una manera detallada la información basada en datos cuantitativos, obteniendo así conclusiones objetivas y mejores recomendaciones, lo cual permite ser más exacto en los procedimientos y la toma de decisiones.²⁴

²¹ Wikipedia, 26 de enero 2023. Matlab. [sitio web]. Disponible en: <https://es.wikipedia.org/wiki/MATLAB>

²² VANEGAS, René. Aplicaciones de inteligencia artificial para la clasificación automatizada de propósitos comunicativos en informes de ingeniería. [En línea]. Revista Signos, 2021. Disponible en: <https://www.semanticscholar.org/paper/Aplicaciones-de-inteligencia-artificial-para-la-de-Venegas/46a3502e80d904808a4427ae5154973fb8f3e30c>

²³ MathWorks, Inc. Classification Learner. [sitio web]. Disponible en:

²⁴ Manuel Fortún, 2012. Importancia de la estadística. [Blog]. Disponible en: http://materiaestadistica.blogspot.com/2012/01/importancia-de-la-estadistica_15.html

Las medidas o parámetros estadísticos son valores numéricos calculados a partir de un conjunto de datos, por lo tanto, permiten resumir la información contenida en dichos datos, como sus diferentes patrones de comportamiento, además de permitir la comparación entre diferentes muestras y los datos que la componen, así como llegar a predecir sus resultados.²⁵

2.8.1 Medidas de centralización

Expresan un valor central representativo del conjunto de datos, las más comunes son:

Media: Es el promedio de los datos de la muestra, es decir es el cociente entre la suma de los datos y el número total de datos.

$$\bar{X} = \frac{\sum_{i=1}^n x_i}{n}$$

Mediana: Es el valor central de los datos ordenados, de modo que la mitad de los datos sean menores que la mediana y la mitad mayor.

Para números impares:

$$M_e = X \frac{n+1}{2}$$

Para números pares: ²⁶

$$M_e = \frac{1}{2} \left(\frac{Xn}{2} + \frac{Xn}{2} + 1 \right)$$

2.8.2 Medidas de dispersión

Son medidas descriptivas, que complementan las medidas de centralización, por lo tanto, indican la desviación de los datos respecto a las medidas centrales.

Desviación estándar o típica: Expresa información sobre la dispersión de los datos respecto al promedio:

$$\sigma = \sqrt{\frac{\sum_i^N (X_i - \underline{X})^2}{N}}$$

²⁵ JUÁREZ, Marc Gisbert. Medidas estadísticas. [sitio web]. Disponible en: <https://www.probabilidadyestadistica.net/medidas-estadisticas/>

²⁶ ALVAREZ BALBUENA, Maria. Medidas estadísticas de centralización. [en línea]. Disponible en: https://rpubs.com/maria_alvarez_b/markdown1

Desviación media absoluta: Es la media de las desviaciones absolutas respecto a la media aritmética.²⁷

$$D_m = \frac{1}{N} \sum_{i=1}^N |X_i - \underline{X}|$$

2.8.3 Medidas de posición o cuantiles

Son valores que dividen la distribución en partes iguales, o en intervalos los cuales tienen el mismo número de valores.

Percentiles: Son valores que dividen en cien partes iguales el conjunto de datos, $P_1, P_2 \dots P_{100}$, dichos valores corresponden del 1% al 100% de los datos.

Cuartiles: Son tres valores que dividen los datos en cuatro partes iguales Q_1, Q_2 y Q_3 , los cuales, respecto a los datos, representan el 25%, 50% y 75% de los datos.²⁸

2.8.4 Medidas estadísticas de forma

Permiten describir una distribución según su forma, es decir, según los datos de la muestra, que forma tiene la curva que los representa.

Asimetría: Es el grado de simetría o asimetría de la distribución, la cual puede ser simétrica, asimétrica, asimétrica negativa y asimétrica positiva, para calcular el coeficiente de asimetría se puede utilizar el coeficiente de Fisher “ g_1 ”:

$$g_1 = \frac{\sum (x_i - \underline{X})^3 n_i}{N s^2}$$

Por lo tanto:

Si $g_1 = 0$ la distribución es simétrica.

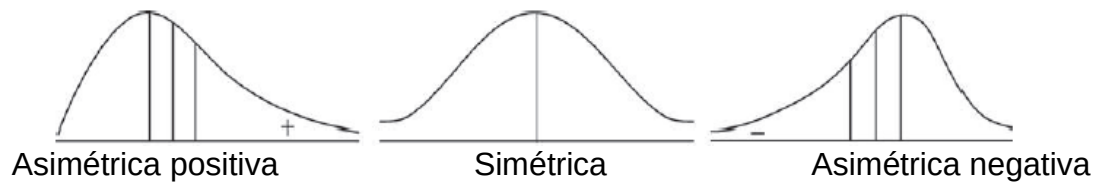
Si $g_1 > 0$ la distribución es asimétrica positiva.

Si $g_1 < 0$ la distribución es asimétrica negativa.

²⁷ VILLASANTE, Paula. Las medidas de dispersión en estadística. [sitio web]. Disponible en: <https://lamenteesmaravillosa.com/las-medidas-de-dispersion-en-estadistica/>

²⁸ HERNÁNDEZ, Juan Luis. Medidas de Posición: Cuantiles. [sitio web]. Disponible en: <https://www.tuveras.com/estadistica/estadistica02.htm>

Figura 4: Grados de asimetría



Fuente: MARTÍNEZ BENCARDINO, Ciro. Estadística y muestreo. Bogotá: Ecoe Ediciones, 2012. ISBN 978-958-648-702-3

Curtosis: Es el grado de concentración de los datos de la muestra, alrededor de la zona central de distribución o normal, si la curva es normal se le considera mesocúrtica, si es más plana que la normal, se considera achatada o platicúrtica y si es más aguda se considera apuntada leptocúrtica, se define mediante la siguiente formula:

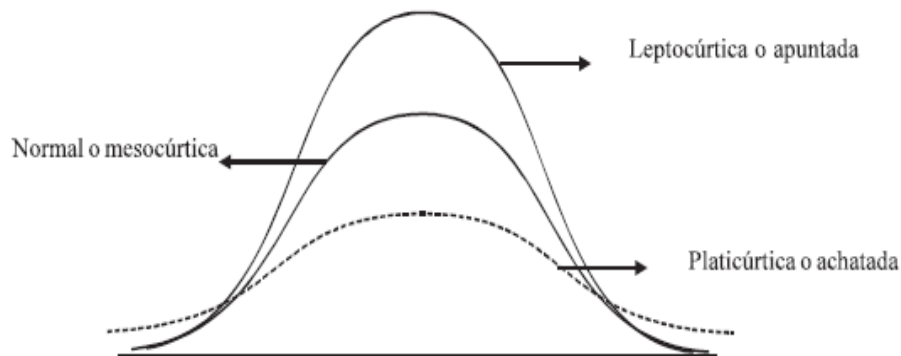
$$g_2 = \frac{\sum (x_i - \underline{X})^4 n_i}{N s^4}$$

Si $g_2 = 3$ la distribución es normal o mesocúrtica.

Si $g_2 > 3$ la curva es apuntada o leptocúrtica.

Si $g_2 < 3$ la curva es achatada o platicúrtica.²⁹

Figura 5. Curtosis en datos



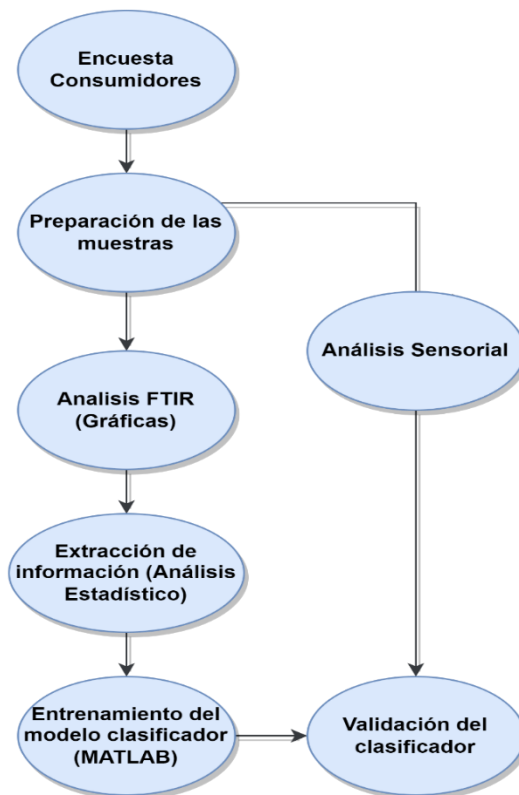
Fuente: MARTÍNEZ BENCARDINO, Ciro. Estadística y muestreo. Bogotá: Ecoe Ediciones, 2012. ISBN 978-958-648-702-3

²⁹ NAVARRO FRÓMETA, Amado Enrique. Medidas de forma. [sitio web]. Disponible en: <https://navarrof.orgfree.com/Docencia/MatematicasIII/M3UT3/m3ut3mf.htm>

3. MATERIALES Y MÉTODOS

Los experimentos, preparación de las muestras, mediciones analíticas y sensoriales se realizaron en el Centro Surcolombiano de investigación en Café CESURCAFÉ, de la Facultad de Ingeniería de la Universidad Surcolombiana en Neiva, Huila. El diseño experimental que se siguió fue el siguiente (Fig. 6):

Figura 6. Diseño Experimental para el desarrollo del proyecto



Fuente: elaboración propia

3.1 MUESTRAS DE CAFÉ

Para realizar la elección de muestras a analizar se realizó una encuesta virtual por medio de la plataforma Google Forms, esta encuesta fue diligenciada por 180 consumidores de café con edades variables. De los resultados obtenidos, se compraron en diferentes supermercados del departamento del Huila las 10 muestras que obtuvieron una mayor puntuación por parte de los usuarios. Además,

se codificaron sus nombres comerciales para salvaguardar conflictos de intereses (academia-industria) y generar así un anonimato en la investigación.

Las muestras se codificaron así:

Primero la palabra Muestra, seguida de la letra que corresponde con la marca comercial, y finalmente los números que indican la repetición de la muestra. Para un total de 11 identificadores, agregando al final el indicativo de la sub-muestra con números “01”, “02”, etc....

- Muestra_A_01
- Muestra_B_01
- Muestra_C_01
- Muestra_D_01
- Muestra_E_01
- Muestra_F_01
- Muestra_G_01
- Muestra_H_01
- Muestra_I_01
- Muestra_J_01
- Muestra_K_01
- Muestra_L_01

Para la muestra testigo se solicitó y analizó una muestra de café especial que fue suplida por CESURCAFÉ, la cual se catalogó como de alta calidad, debido a sus características físicas y sensoriales.

3.2 ANÁLISIS DEL ESPECTRO INFRARROJO CON FTIR

Para el análisis de las muestras se usó el espectrómetro FTIR Cary 630 (Agilent, USA) con un detector DLATGS y un accesorio de muestreo ATR. El software utilizado para realizar el escaneo de las muestras con el FTIR es MicroLab el cual viene instalado con el equipo.

Figura 7. Espectrómetro FTIR Cary 630



Fuente: Satapathi lab. Fourier-Transform Infrared Spectroscopy (FTIR). [sitio web]. Disponible en: <https://satapathilab.iitr.ac.in/facilities/>

Se limpió el equipo y se pesó aproximadamente un gramo del café molido tostado de las muestras seleccionadas, se cerraba el accesorio de acople de muestras y se proseguía con la lectura, se seleccionó el rango de $4000 - 650\text{cm}^{-1}$, con una resolución de 8cm^{-1} y 64 escaneos; al final se muestra la gráfica con los niveles de absorbancia de los compuestos en las diferentes longitudes de onda. Se realizaron 5 análisis por cada muestra.

3.3 ANÁLISIS SENSORIAL

Se realizó el análisis sensorial con el método de infusión, asimilando la preparación convencional del consumo de café comercial. Para la preparación de las muestras se realizó una relación de 7 gramos por 100 mililitros de agua filtrada con carbón activado a una temperatura de 90 grados Celsius de acuerdo con lo recomendado por Sánchez & Chambers (2015), para las muestras de café molido tostado.

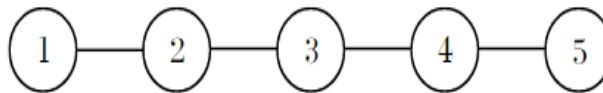
Para las muestras de café instantáneo se usó la relación de 2 gramos por 150 mililitros de agua filtrada con carbón activado a una temperatura de 90 grados Celsius de acuerdo con lo recomendado por Oliveira, et al., (2009).

Se realizó una sesión de catación conformada por 5 jueces semi-entrenados. Se consideraron como jueces semi-entrenados a aquellos que tenían conocimiento teórico, práctico y habilidad en el análisis sensorial de café, y como consumidores habituales a las personas que no realizan evaluaciones sensoriales periódicas, pero que consumen habitualmente café.

Se nombraron las muestras con números aleatorios de tres cifras y se les retiró cualquier empaque o distintivo con el que se pudiera conocer qué marca de café era para no sesgar la evaluación ya que podrían tenerse conclusiones apresuradas o comentarios de la experiencia con el café seleccionado.

A cada juez se le proporcionó 30 mililitros de la bebida de café preparada la cual iba marcada con el número aleatorio indicado para cada marca y el formato elaborado en el cual tenían que evaluar los atributos del café que son: color, aroma, sabor, sabor residual y cuerpo, en la siguiente escala:

Figura 8. Escala evaluación análisis sensorial



Fuente: elaboración propia

1. Me disgusta mucho
2. Me disgusta poco
3. Ni me gusta ni me disgusta
4. Me gusta poco
5. Me gusta mucho

Lo que debían hacer era rellenar el círculo dependiendo la evaluación que consideraran adecuada para cada muestra de café.

3.4 DESARROLLO DEL ALGORITMO DE CLASIFICACIÓN

El algoritmo de clasificación se realizó mediante el IDE MATLAB, primero, se obtuvieron los datos de las muestras arrojadas por el FTIR en archivos “.csv”, los cuales se adecuaron eliminando información irrelevante y se nombraron de acuerdo a la muestra y el número de muestra de cada marca seleccionada, para que fuera más fácil manipular los archivos “.csv” en el algoritmo, cada archivo contiene dos columnas, la primera columna corresponde a los datos de longitud de onda denominada “XLabelWavenumber” y la segunda columna corresponde a los datos de absorbancia de la muestra denominada “YLabelAbsorbance”.

Se crea un almacén de datos y se guarda en una variable con el fin de recopilar todos los archivos cuya extensión sea “.csv”, esto con el fin de agrupar los datos y

más adelante poderlos manipular, mediante un ciclo iterativo, se obtienen los datos de absorbancia de las muestras para cada uno de los compuestos seleccionados, dichos datos son los que se utilizan para el entrenamiento de los métodos de clasificación.

Se llama la función que calcula las diferentes medidas estadísticas seleccionadas, los argumentos de entrada de la función son los datos de absorbancia de cada muestra, dichas medidas estadísticas se utilizaron como predictores para entrenar los modelos de clasificación.

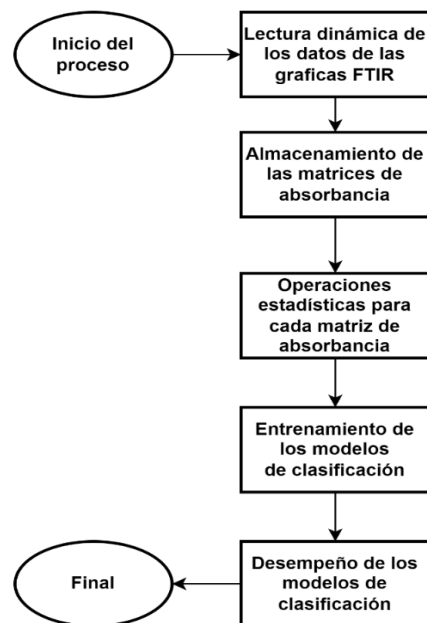
Se crea un archivo “.csv” el cual corresponde a la clasificación manual que se le asignó a las muestras, dicho archivo se guarda en una variable de tipo categórica.

Se realizaron dos clasificaciones, la primera corresponde el nivel de cafeína de las muestras (alto o bajo) y la segunda a el tipo de café (instantáneo o tostado molido), se obtienen dichas clasificaciones en variables independientes, luego se declaran dos variables correspondientes a el tipo de clasificación, en cada una se guarda la tabla de medidas estadísticas y se le agrega en su última columna la clasificación manual correspondiente.

Se realiza la partición de los datos, seleccionando el 80% de los datos para entrenamiento y el 20% para prueba o testeo, dichos datos se guardan en variables disponibles en el “Workspace” de MATLAB, se procede a entrenar los modelos mediante la aplicación “Classification Learner”, para lo cual se ingresan los datos de entrenamiento y se ejecutan todos los tipos de clasificadores, esto con el fin de determinar qué tipo de clasificador tuvo un porcentaje más alto de exactitud, por último, mediante los datos de testeo se analiza el desempeño de los diferentes modelos de clasificación.

A continuación se presenta el diagrama de flujo del algoritmo de clasificación:

Figura 9. Diagrama de flujo del algoritmo de clasificación



Fuente: elaboración propia

3.5 ANÁLISIS ESTADÍSTICO

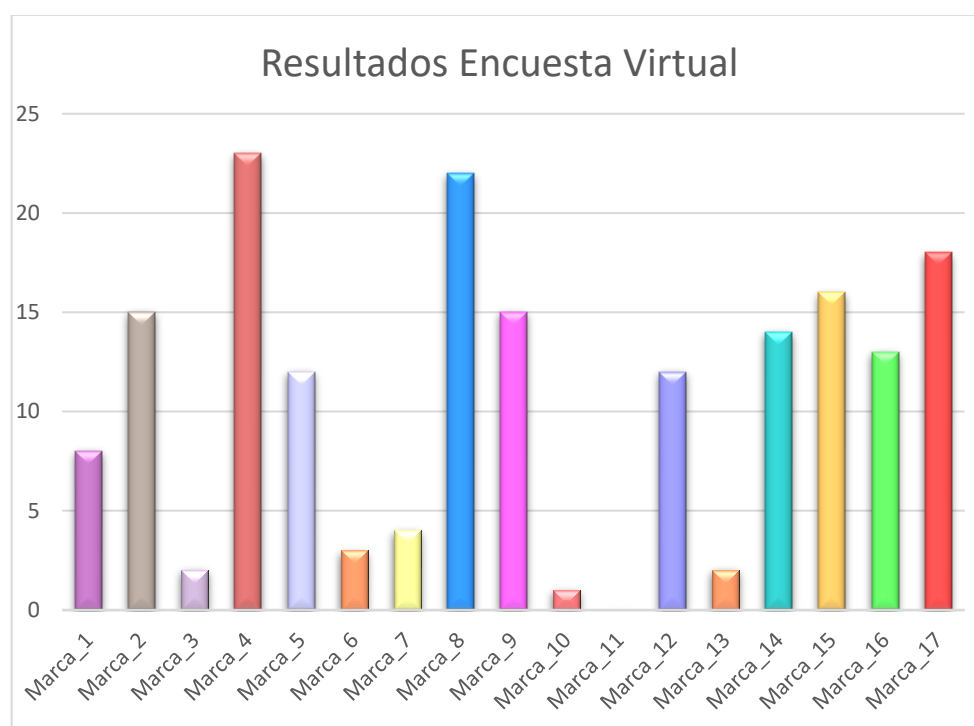
El análisis estadístico se realizó en MATLAB mediante el desarrollo de una función que se nombró "tablaDatosEstadísticos", la cual se encarga de aplicar las diferentes medidas estadísticas seleccionadas a los datos de absorbancia de todas las muestras, en cada uno de los compuestos seleccionados, las medidas estadísticas utilizadas fueron mínimo, máximo, media, mediana, desviación estándar, desviación media absoluta, percentil 25 y 75, rango inter cuantil, y curtosis, cada medida estadística se aplicó a los datos mediante funciones propias de Matlab, a las cuales como parámetro de entrada se ingresaron los datos de las gráficas y como salida se obtiene dicha medida estadística los valores de las medidas estadísticas se guardan en una variable de tipo tabular, dichos valores se utilizaron para entrenar los diferentes modelos de clasificación.

4. RESULTADOS Y DISCUSIÓN

4.1 RESULTADOS DE ENCUESTA VIRTUAL

Los resultados de la encuesta permitieron elegir las marcas de café comercial. Fue necesario el estudio de las marcas que están a la venta en los mercados de la región para así poder seleccionar según la disponibilidad y el fácil acceso de los consumidores. En la siguiente gráfica se muestran los resultados de la encuesta, donde se observan todas las marcas y las que tuvieron resultados por encima de 10, fueron las que se eligieron para realizar el análisis.

Grafica 1. Resultados encuesta virtual



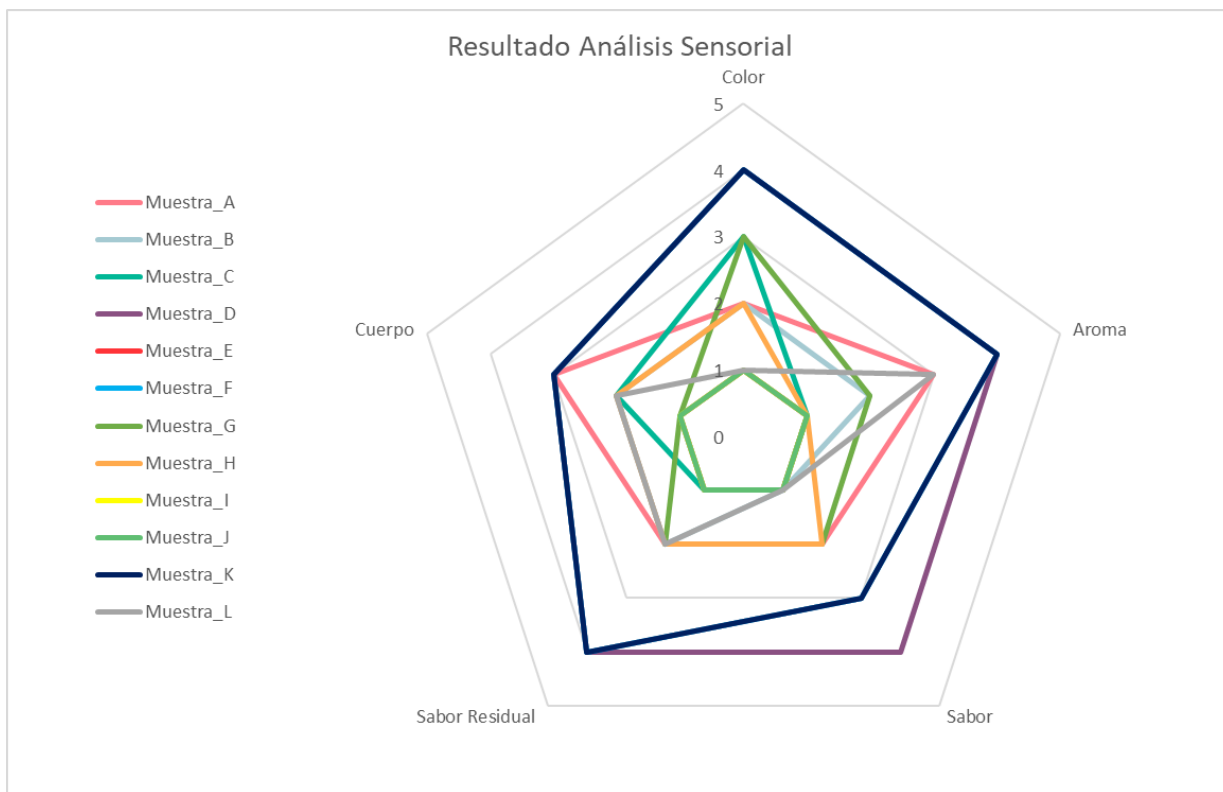
Fuente: elaboración propia

Se seleccionaron 10 marcas para realizar el proyecto. Además, se elige una muestra testigo la cual es un café auspiciado por Centro Surcolombiano de investigación en Café CESURCAFÉ, el cual cuenta con una excelente calidad y con un cuidadoso proceso de producción.

4.2 RESULTADOS DE ANÁLISIS SENSORIAL

El siguiente gráfico de araña representa el resultado de la evaluación sensorial por el panel de catadores para cada una de las muestras analizadas con el FTIR:

Gráfica 2. Resultados Análisis Sensorial



Fuente: elaboración propia

Como se puede observar, las marcas que tuvieron mejores resultados de acuerdo con el sabor fueron la Muestra D y la muestra K, las cuales también tuvieron los mejores resultados en cuanto al color, aroma y sabor. En cuerpo, los mejores resultados fueron de las muestras K y A. Las muestras con los resultados más desfavorables fueron las muestras I, J, y E, en todas las características, tuvieron mayor cantidad de votos en 1 de la escala que corresponde a “me disgusta mucho”.

Las características evaluadas fueron elegidas por ser lo que suele tener un café que son fácilmente captadas por los sentidos y que pueden ser fácilmente percibidas por cualquier consumidor.

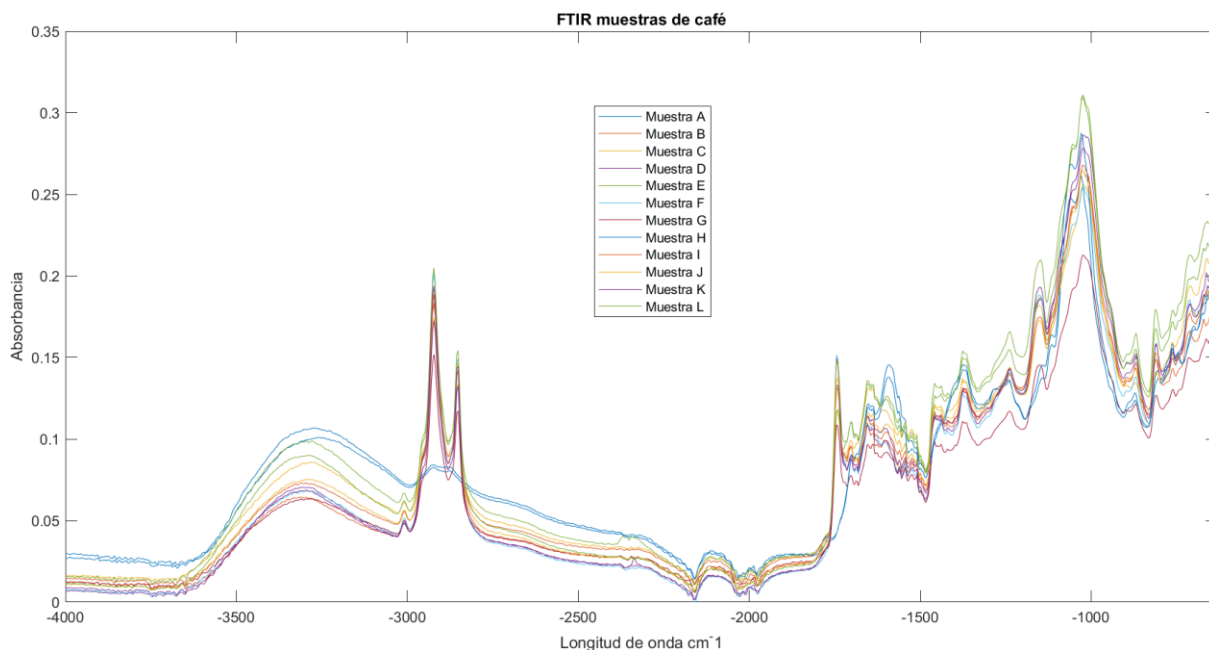
4.3 RESULTADOS ANÁLISIS FTIR

Para obtener la información espectral se configura el espectrómetro FTIR con las siguientes características:

- Rango: $4000 - 650 \text{ cm}^{-1}$, indicando que se trabajará en infrarrojo medio.
- Background scans (Escaneo de fondo): 8.
- Sample scan (Escaneo de la muestra): 64.
- Resolution (Resolución): 8 cm^{-1} .
- En eje Y se selecciona *absorbancia*.

Se realiza el muestreo por marca de café y se obtiene la siguiente grafica de valores promedio en absorbancia por longitud de onda:

Gráfica 3. Niveles de absorbancia por longitud de onda. Análisis FTIR



Fuente: elaboración propia

Se observan distintos niveles de absorbancia y picos característicos.

Esta información espectral es la base para el análisis de los componentes y posibles adulteraciones de las muestras de café.

4.4 RESULTADOS ALGORITMO DESARROLLADO

Con el fin de analizar el desempeño de los diferentes modelos de clasificación entrenados, se deben definir métricas que permitan analizar el comportamiento de los modelos de clasificación, como lo son la matriz de confusión y la curva de ROC.

La matriz de confusión permite visualizar la cantidad de aciertos y errores que presenta el modelo de clasificación.

Figura 10. Matriz de confusión

		Predicción	
		Positivos	Negativos
Observación	Positivos	Verdaderos Positivos (VP)	Falsos Negativos (FN)
	Negativos	Falsos Positivos (FP)	Verdaderos Negativos (VN)

Fuente: Zelad, Carlos. Evaluación de modelos de clasificación. [en línea]. Disponible en: <https://rpubs.com/chzelada/275494>

- **VP:** Cantidad de positivos clasificados correctamente como positivos.
- **FN:** Cantidad de positivos clasificados incorrectamente como negativos.
- **FP:** Cantidad de negativos clasificados incorrectamente como positivos.
- **VN:** Cantidad de negativos clasificados correctamente como negativos.³⁰

A partir de la matriz de confusión es posible obtener la exactitud (Accuracy) así como diferentes métricas que permiten entender el rendimiento del modelo.

La curva de ROC, muestra la tasa de verdaderos positivos frente a la tasa de falsos positivos para diferentes umbrales de puntajes de clasificación, dicha curva se suele acompañar por el valor AUC (área bajo la curva), el cual es un cálculo integral de la superficie que está debajo de la curva del modelo, el valor AUC se encuentra entre 0 a 1, entre más se acerque a 1, significa que el clasificador tiene un mejor

³⁰ ZELAD, Carlos. Evaluación de modelos de clasificación. [en línea]. Disponible en: <https://rpubs.com/chzelada/275494>

rendimiento, por lo tanto el valor AUC indica la calidad general del clasificador.³¹ A continuación, se muestran los resultados del entrenamiento y testeo de los modelos de clasificación:

Primero se observan los resultados de entrenamiento de la clasificación del nivel de cafeína (alto o bajo).

Tabla 1. Entrenamiento modelos de clasificación nivel de cafeína (alto o bajo)

#	Nombre del modelo	Exactitud (%)
1	Fine Tree	100
2	Medium Tree	100
3	Coarse Tree	100
4	Linear Discriminant	100
5	Quadratic Discriminant	Fallido
6	Logistic Regression	100
7	Gaussian Naive Bayes	100
8	Kernel Naive Bayes	100
9	Linear SVM	100
10	Quadratic SVM	100
11	Cubic SVM	100
12	Fine Gaussian SVM	93.8
13	Medium Gaussian SVM	100
14	Coarse Gaussian SVM	100
15	Fine KNN	100
16	Medium KNN	100
17	Coarse KNN	83.3
18	Cosine KNN	100
19	Cubic KNN	100
20	Weighted KNN	100
21	Boosted Trees	83.3
22	Bagged Trees	100
23	Subspace Discriminant	100
24	Subspace KNN	100
25	RUSBoosted Trees	83.3
26	Narrow Neuronal Network	100
27	Medium Neuronal Network	100
28	Wide Neuronal Network	100

³¹ MathWorks, Inc. Evaluación del rendimiento de los modelos de clasificación de machine learning. [sitio web]. Disponible en: <https://la.mathworks.com/discovery/roc-curve.html>.

Tabla 1. (Continuación)

29	Bilayered Neuronal Network	100
30	Trilayered Neuronal Network	100
31	SVM Kernel	100
32	Logistic Regression Kernel	100

Fuente: elaboración propia

Se puede observar que la mayoría de los modelos tuvo una exactitud del 100% o mayor a 80%.

Para el análisis de desempeño, se seleccionó el modelo 29 Red neuronal bicapa (Bilayered Neural Network).

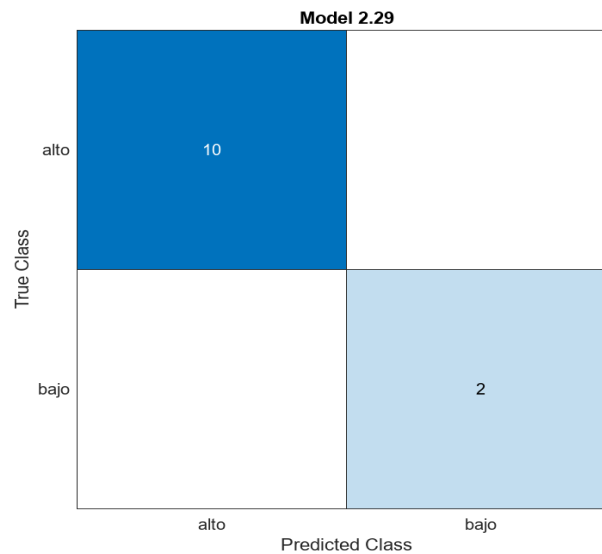
Figura 11. Resumen entrenamiento

Model 2.29: Neural Network	
Status: Trained	
Training Results	
Accuracy (Validation)	100.0%
Total cost (Validation)	Not applicable
Prediction speed	~460 obs/sec
Training time	15.42 sec
Test Results	
Accuracy (Test)	100.0%
Total cost (Test)	Not applicable

Fuente: elaboración propia

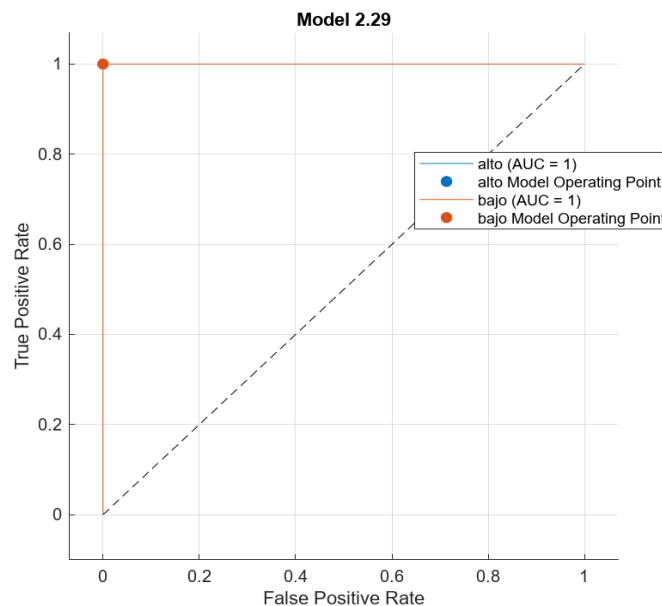
La aplicación “Classification Learner” de MATLAB, permite obtener la matriz de confusión de cada modelo, así como su curva ROC, por lo tanto, procede a probar el desempeño del modelo mediante los datos de prueba.

Figura 12. Matriz de confusión modelo clasificador nivel cafeína



Fuente: elaboración propia

Figura 13. Curva ROC modelo clasificador nivel de cafeína.



Fuente: elaboración propia

Ahora se observan los resultados del entrenamiento de la clasificación por tipo de café (instantáneo o tostado molido).

Tabla 2. Entrenamiento modelos de clasificación tipo de café (Instantáneo o tostado-molido).

N°	Nombre del modelo	Exactitud (%)
1	Fine Tree	100
2	Medium Tree	100
3	Coarse Tree	100
4	Linear Discriminant	100
5	Cuadratic Discriminant	Fallido
6	Logistic Regression	100
7	Gaussian Naive Bayes	100
8	Kernel Naive Bayes	100
9	Linear SVM	100
10	Quadratic SVM	100
11	Cubic SVM	100
12	Fine Gaussian SVM	91.7
13	Medium Gaussian SVM	100
14	Coarse Gaussian SVM	100
15	Fine KNN	100
16	Medium KNN	100
17	Coarse KNN	83.3
18	Cosine KNN	100
19	Cubic KNN	100
20	Weighted KNN	100
21	Boosted Trees	83.3
22	Bagged Trees	100
23	Subspace Discriminant	100
24	Subspace KNN	100
25	RUSBoosted Trees	85.4
26	Narrow Neuronal Network	100
27	Medium Neuronal Network	100
28	Wide Neuronal Network	100
29	Bilayered Neuronal Network	100
30	Trilayered Neuronal Network	100
31	SVM Kernel	100
32	Logistic Regression Kernel	100

Fuente: elaboración propia

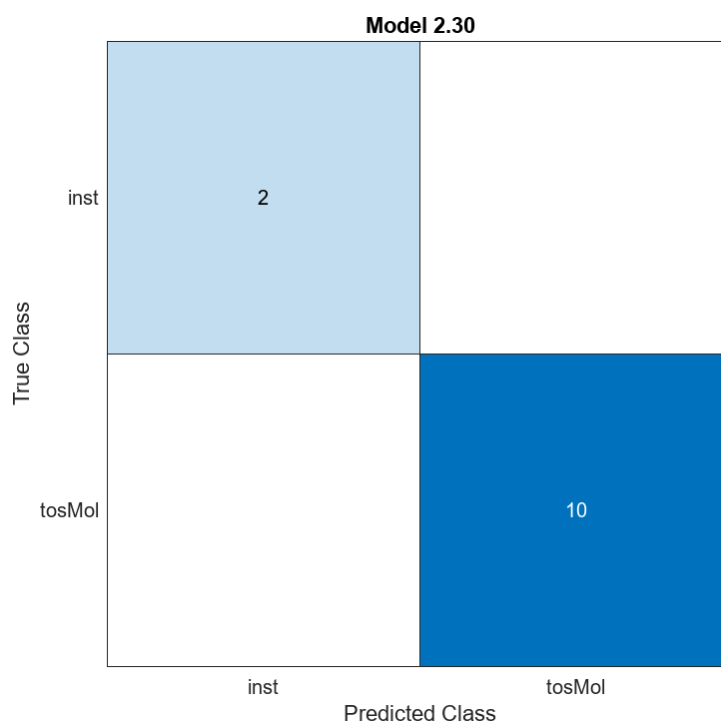
Con el fin de analizar el desempeño para este caso se selecciona el modelo 30 Red neuronal tricapa (Trilayered Neuronal Network).

Figura 14. Resumen entrenamiento

Model 2.30: Neural Network	
Status: Trained	
Training Results	
Accuracy (Validation)	100.0%
Total cost (Validation)	Not applicable
Prediction speed	~210 obs/sec
Training time	15.497 sec
Test Results	
Accuracy (Test)	100.0%
Total cost (Test)	Not applicable

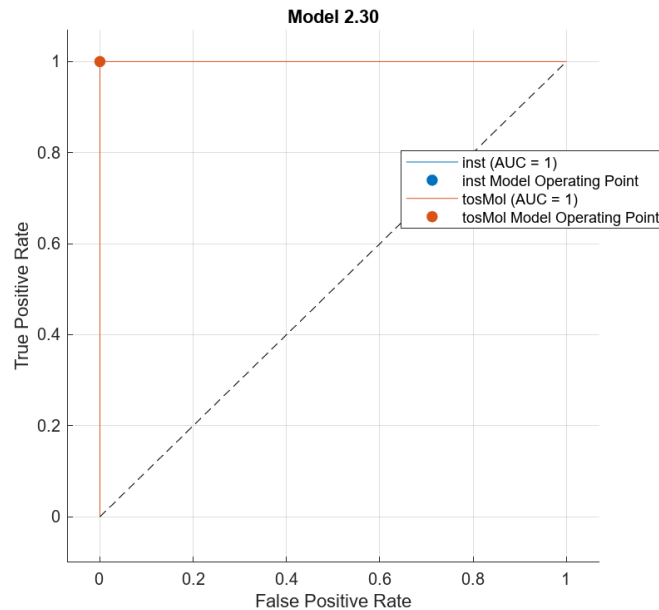
Fuente: elaboración propia

Figura 15. Matriz de confusión modelo clasificador tipo de café



Fuente: elaboración propia

Figura 16. Curva ROC modelo clasificador nivel de cafeína



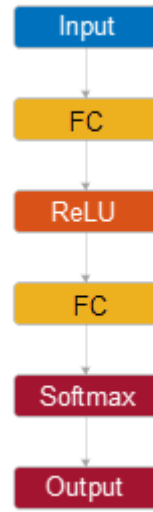
Fuente: elaboración propia

Cada uno de los modelos disponibles en la aplicación Classification Learner de MATLAB, es una red neuronal de avance diseñada para clasificar datos. Estas redes están compuestas por varias capas completamente conectadas, en las que la primera capa recibe la entrada de los datos predictores y las capas subsiguientes reciben la entrada de la capa anterior. En cada capa completamente conectada, se multiplican los datos de entrada por una matriz de pesos y se les agrega un vector de sesgo antes de aplicar una función de activación. La capa final de la red, que también está completamente conectada, utiliza una función softmax para producir la salida de la red, que consiste en puntuaciones de clasificación (probabilidades posteriores) y etiquetas de clasificación predichas.³²

El clasificador de red neuronal predeterminado tiene la siguiente estructura de capas:

³² MathWorks, Inc. App Classification Learner. Choose Classifier Options. [sitio web]. Disponible en: https://www-mathworks-com.translate.goog/help/stats/choose-a-classifier.html?_x_tr_sl=en&_x_tr_tl=es&_x_tr_hl=es&_x_tr_pto=sc

Figura 17. Estructura clasificador red neuronal



Fuente: MathWorks, Inc. Estructura red neuronal. [sitio web]. Disponible en: [https://www-mathworks-com.translate.goog/help/stats/fitnet.html? x tr sl=en& x tr tl=es& x tr hl=es& x tr_pto=sc#mw_a9fa4524-aecf-4d9b-95ef-58914f7780f3](https://www.mathworks-com.translate.goog/help/stats/fitnet.html? x tr sl=en& x tr tl=es& x tr hl=es& x tr_pto=sc#mw_a9fa4524-aecf-4d9b-95ef-58914f7780f3)

Entrada (Input): Esta capa corresponde a los datos predictores.

Primera capa totalmente conectada (FC): Esta capa tiene 10 salidas de forma predeterminada, se puede ampliar la capa o agregar capas más completamente conectadas a la red, además se puede encontrar los pesos y sesgos de esta capa.

Función de activación (ReLU): fitnet aplica la función de activación a la primera capa completamente conectada.

Capa final totalmente conectada (FC): Esta capa tiene K salidas, donde K es el número de clases en la variable de respuesta. Se puede encontrar los pesos y sesgos de esta capa.

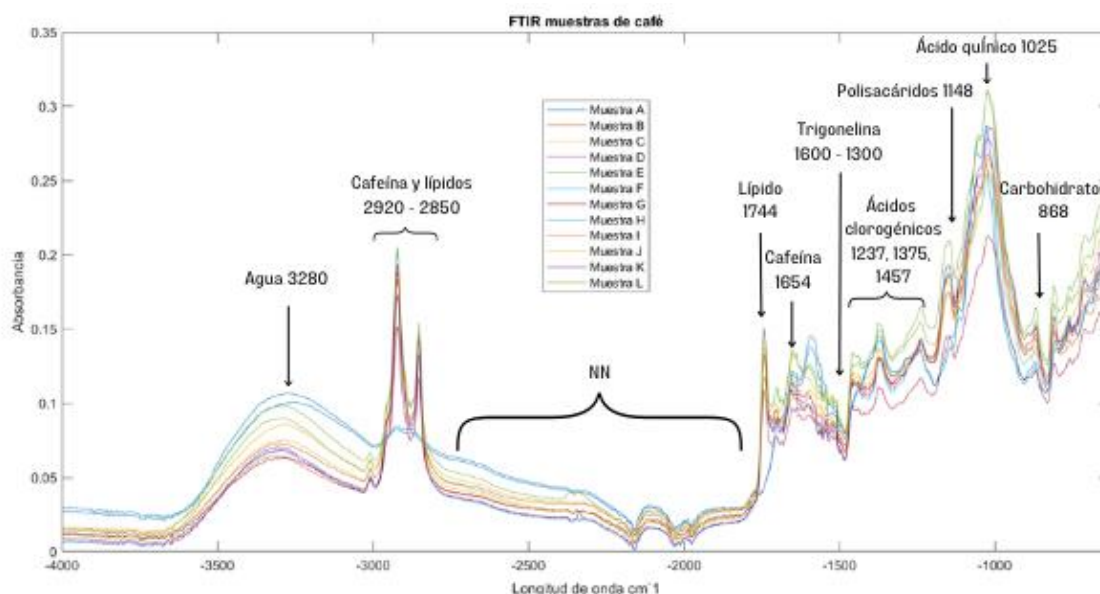
Función Softmax (para clasificación binaria y multiclase): fitnet aplica esta función de activación a la capa final completamente conectada. La función toma cada entrada x_i y devuelve lo siguiente, donde K es el número de clases en la variable de respuesta. Los resultados corresponden a las puntuaciones de clasificación previstas (o probabilidades posteriores).

Salida (Output): esta capa corresponde a las etiquetas de clase predichas.³³

4.5 RESULTADOS ESTADÍSTICOS

Con el fin de generar los suficientes predictores para entrenar los distintos clasificadores, se calculó el mínimo, máximo, media, mediana, desviación estándar, desviación media absoluta, percentil 25 y 75, rango intercuartílico, asimetría y curtosis. Cada medida estadística se calculó por cada compuesto seleccionado de los datos de absorbancia del FTIR a todas las muestras de las diferentes marcas de café.

Figura 18. Compuestos del FTIR característicos de las muestras de café



Fuente: elaboración propia, con base en BARRIOS RODRIGUEZ, Yeison Fernando; SALAS CALDERON, Karen Tatiana y GIRÓN HERNÁNDEZ, Joel, 2018. Comparison of sensory attributes and chemical markers of the infrared spectrum between defective and non-defective Colombian coffee simples. En: Coffee Science, e151659, 2020. Doi: <https://doi.org/10.25186/cs.v15i.1659>

³³ MathWorks, Inc. App Classification Learner. fitcnet [sitio web]. Disponible en: https://www-mathworks-com.translate.goog/help/stats/fitcnet.html?_x_tr_sl=en&_x_tr_tl=es&_x_tr_hl=es&_x_tr_pto=sc#mw_a9fa4524-aecf-4d9b-95ef-58914f7780f3

En la banda de 3280 cm^{-1} se asigna al agua³⁴, se observa que las muestras A y H presentan un mayor nivel de absorbancia con respecto a las demás, por lo que puede que dichas muestras tengan un mayor nivel de agua, las muestras en cuestión corresponden a las dos marcas de café instantáneo que se seleccionaron.

En el rango de 2920 cm^{-1} y 2850 cm^{-1} se asigna la molécula de cafeína^{35 36}, en 1650 cm^{-1} es la banda determinante para la cafeína en café tostado³⁷. Se evidencia que las muestras de café instantáneo A y H, presentan una baja absorbancia en las longitudes correspondientes a la cafeína. El rango 2920 cm^{-1} y 2850 cm^{-1} se asigna a las bandas combinadas a las que contribuyen tanto la cafeína como los lípidos³⁸.

En las muestras de café analizadas, se presentaron lípidos en la longitud de onda 1744 cm^{-1} , algunos autores relacionan la banda de 1740 cm^{-1} con la vibración de carbonilo del grupo éster en triglicéridos o en ésteres alifáticos indicando que podría estar asociada a lípidos^{39 40 41}, en la Figura 18 (página 49) se evidencia en la marca K correspondiente a una marca artesanal del Huila, presentó el pico más alto de absorbancia en la banda de lípidos a diferencia de las marcas de café instantáneos que presentan un bajo nivel de absorbancia en dicha banda, las demás muestras presentaron un nivel similar en cuanto a lípidos.

³⁴ BARRIOS RODRIGUEZ, Yeison Fernando; SALAS CALDERON, Karen Tatiana y GIRÓN HERNÁNDEZ, Joel, 2018. Comparison of sensory attributes and chemical markers of the infrared spectrum between defective and non-defective

Colombian coffee samples. En: Coffee Science, e151659, 2020. Doi: <https://doi.org/10.25186/cs.v15i.1659>

³⁵ CRAIG, Ana Paula; FRANCA, Adriana S y OLIVEIRA, Leonardo S, 2012a. Evaluation of the potential of FTIR and chemometrics for separation between defective and non-defective coffees. Food Chemistry. DOI: <https://doi.org/10.1016/j.foodchem.2011.11.121>

³⁶ PARADKAR, Manish M y IRUDAYARAJ, Joseph, 2004. Rapid determination of caffeine content in soft drinks using FTIR-ATR spectroscopy. Food Chemistry. DOI: [doi.org/10.1016/S0308-8146\(02\)00116-4](https://doi.org/10.1016/S0308-8146(02)00116-4)

³⁷ GARRIGUES, J. M, et al, 2000. Fourier transform infrared determination of caffeine in roasted coffee samples. Fresenius J Anal Chem. DOI: doi.org/10.1007/s002160050063

³⁸ CREMER, Dirk Rainer y KALETUNÇ, Gönül, 2003. Fourier transform infrared microspectroscopic study of the chemical microstructure of corn and oat flour-based extrudates. Carbohydrate Polymers. DOI: [doi.org/10.1016/S0144-8617\(02\)00266-7](https://doi.org/10.1016/S0144-8617(02)00266-7)

³⁹ KEMSLEY, E. K; RUAULT, S y WILSON, R.H, 1995. Discrimination between Coffea arabica and Coffea canephora variant robusta beans using infrared spectroscopy. Food Chemistry. DOI: [doi.org/10.1016/0308-8146\(95\)00030-M](https://doi.org/10.1016/0308-8146(95)00030-M)

⁴⁰ LYMAN, Donald. J, et al, 2003. FTIR-ATR analysis of brewed coffee: Effect of roasting conditions. Journal of Agricultural and Food Chemistry. DOI: [10.1021/jf0209793](https://doi.org/10.1021/jf0209793)

⁴¹ CRAIG, Ana Paula; FRANCA, Adriana S y OLIVEIRA, Leonardo S, 2012b. Discrimination between defective and non-defective roasted coffees by diffuse reflectance infrared Fourier transform spectroscopy. LWT - Food Science and Technology. DOI: doi.org/10.1016/j.lwt.2012.02.016

La banda de los 1591 cm^{-1} corresponde a trigonelina, se asocia a los picos en el rango de $1600 - 1300\text{ cm}^{-1}$ ⁴², las muestras *K* y *H* presentan una mayor absorbancia en los primeros picos en comparación con las muestras de café tostado molido, en los picos siguientes se disminuye el nivel de absorbancia para dicha sustancia en todas las muestras, pero se mantiene la relación de nivel.

El rango de $1450 - 1000\text{ cm}^{-1}$ se asigna a los ácidos clorogénicos⁴³, las muestras analizadas presentaron picos de dicha sustancia en las bandas 1457 cm^{-1} , 1375 cm^{-1} y 1237 cm^{-1} , en el primer pico las muestras de café tostado presentaron diferentes valores, sin embargo, las muestras de café instantáneo *A* y *H* presentaron un aumento gradual, siendo la segunda y tercera banda en la cual todas las muestras presentan picos de absorbancia.

La banda de 1025 cm^{-1} corresponde al ácido quínico, que pertenece a la familia de los ácidos clorogénicos⁴⁴, en la muestra *G* se observa el nivel más bajo de absorbancia para dicho compuesto, las demás muestras presentan un nivel de absorbancia similar.

La banda 1148 cm^{-1} corresponde a los polisacáridos, el café comercial molido; en la fase de tostado se alcanzan temperaturas superiores a los 180°C que producen una reacción exotérmica que involucra compuestos como polisacáridos, entre otros⁴⁵. Estos resultados muestran un desplazamiento en la longitud de onda⁴⁶, asocia la absorción en la banda de 1153 cm^{-1} a polisacáridos, la muestra *E* presenta el mayor nivel de absorbancia lo cual puede corresponder a una marca con nivel de tostión alto, las demás muestras presentan un nivel de absorbancia similar lo cual indica que las muestras pueden presentar un nivel de equivalente.

Las muestras presentan nivel de absorbancia en la banda de 868 cm^{-1} que corresponde a carbohidratos, según el autor⁴⁷, Las muestras *L* y *E* presentaron los

⁴² Ibíd.

⁴³ Ibíd.

⁴⁴ REIS, Nadia; FRANCA, Adriana S y OLIVEIRA, Leonardo S, 2013a. Discrimination between roasted coffee, roasted corn and coffee husks by Diffuse Reflectance Infrared Fourier Transform Spectroscopy. Food Science and Technology. DOI: 10.1016/j.lwt.2012.07.016

⁴⁵ HERNÁNDEZ, J. A: HEYD, B. y TRYSTRAM, G., 2008. Prediction of brightness and surface area kinetics during coffee roasting. Journal of Food Engineering. DOI: 10.1016 / j.jfoodeng.2008.04.026

⁴⁶ CRAIG, Ana Paula; FRANCA, Adriana S y OLIVEIRA, Leonardo S, 2018. Mid infrared spectroscopy and chemometrics as tools for the classification of roasted coffees by cup quality. Food Chemistry. DOI: doi.org/10.1016/j.foodchem.2017.11.066

⁴⁷ LYMAN, Donald. J, et al, 2003. FTIR-ATR analysis of brewed coffee: Effect of roasting conditions. Journal of Agricultural and Food Chemistry. DOI: 10.1021/jf0209793

mayores picos de absorbancia y la muestra G presentó el menor nivel, las demás muestras presentaron un nivel similar. En la banda entre los $1400 - 900\text{ cm}^{-1}$ se pueden presentar carbohidratos, no obstante, puede corresponder a otros compuestos⁴⁸.

⁴⁸ WANG, Jun, et al, 2009. Fourier transform infrared spectroscopy for kona coffee authentication. Journal of Food Science. DOI: 10.1111 / j.1750-3841.2009.01173.x

5. CONCLUSIONES

El método de análisis de Espectroscopia Infrarroja por Transformada de Fourier FTIR es una técnica analítica que permite identificar la familia química que posee una sustancia, en este caso fueron las muestras de café. Mide absorbancia de la muestra y genera un espectro basado en los grupos funcionales, como se puede ver en la Figura 18 (página 49). Se analizaron las muestras comerciales elegidas según la encuesta de estudio de consumidores del departamento del Huila, como son muestras donde no se conoce su proceso de cultivo, selección, producción, tostión y qué tanto tiempo lleva en el empaque, no es posible analizar adulteraciones específicas y su calidad, ya que los grupos funcionales detectados en el análisis son característicos del café pero puede dar información de qué posibles compuestos puede tener y así relacionar si es posible que pueda contener ciertos componentes que según los niveles de absorbancia, se puede mencionar la cantidad de dichos compuestos en las muestras. Estos compuestos pueden dar posible información de tostión, características del sabor como acidez, aroma, entre otros.

Se realizó una metodología generando el patrón de referencia de una muestra de café natural no comercial, donde se conocían todas sus características de cultivo, selección, producción y tostión procesada en las instalaciones del Centro Surcolombiano de Investigación en Café CESURCAFÉ una muestra testigo 100% café. Esto permitió poder comparar la muestra mencionada con las muestras de café comercial realizando el mismo proceso de análisis. En donde se logra evidenciar que también posee niveles de absorbancia y picos característicos en ciertos valores de longitud de onda, como mencionado anteriormente, son picos característicos del café, lo que se llama, huella del café en FTIR.

El análisis sensorial indicó que la calidad de las muestras de café tiene una calificación baja en la mayoría de las marcas evaluadas, los catadores semi-entrenados pudieron evidenciar críticas considerables sobre las características sensoriales del café, las muestras que tuvieron la calificación más baja en todas sus características fueron las muestras *I*, *J*, y *E*. Donde se obtuvieron descriptores en cada característica como: En color: Muy oscuro, tueste alto; en aroma: quemado, ahumado; en sabor: amargo, quemado, desagradable; en sabor residual: amargo, quemado; en cuerpo: pesado, simple, denso. Las marcas evaluadas que tuvieron los mejores resultados fueron la *K* y la *D*, en las cuales se obtuvieron los siguientes suscriptores: en color: tueste medio, agradable; en aroma: leve caramelo, notas frutales, panela; en sabor: dulce, limón, limoncillo; en sabor residual: herbal, delicado, limón, perdurable; y en cuerpo: simple, sedoso, acuoso, ligero. Se logra

percibir por parte de los evaluadores, características sensoriales que pueden estar relacionadas con los grupos funcionales en donde se puede observar que como las muestras evaluadas con la más alta y la más baja escala, en el análisis FTIR las muestras nombradas tienen una respuesta similar, pero no se puede dar una conclusión certera ya que no se conocen las condiciones de producción del café, como su tostión, tiempo de empaquetado, etc.

Se logró desarrollar un algoritmo de clasificación utilizando métodos de Machine Learning, el algoritmo se desarrolló mediante el IDE MATLAB y su aplicación Classification Learner. Implementando aprendizaje supervisado, se entrenaron distintos modelos de clasificación con el fin de encontrar el modelo óptimo.

Se entrenaron los diferentes modelos de clasificación para realizar dos clasificaciones diferentes, la primera clasificación consiste en determinar si la muestra presenta niveles altos o bajos de cafeína, la segunda clasificación determina si la muestra corresponde a café instantáneo o a café tostado molido, los datos se dividieron en 80% para entrenamiento y 20% para testeo o validación, todos los modelos se entrenaron con las configuraciones por defecto, la mayoría de los modelos presentaron una exactitud del 100% o mayor a 80%, los modelos con menor exactitud fueron los mismos para ambas clasificaciones, esto puede indicar que dichos modelos no son los más óptimos para los tipos de datos utilizados en el entrenamiento sin embargo se puede buscar aumentar la exactitud mediante la configuración disponible para cada modelo, se presentó un modelo fallido en el entrenamiento, lo cual se debe a la forma en la cual dicho modelo realiza la clasificación que es mediante discriminación cuadrática, dicha clasificación no se pudo realizar ya que según la documentación de MATLAB los datos de entrenamiento presentan matrices singulares, por lo tanto se debe realizar un ajuste a las configuraciones del modelo en cuestión.

Se determinó que los modelos óptimos para realizar la clasificación son los correspondientes a redes neuronales, ya que son modelos que suelen tener una buena precisión predictiva y se pueden utilizar para clasificación multiclase, sin embargo, se debe tener en cuenta que dichos modelos son difíciles de interpretar y su flexibilidad disminuye al aumentar su tamaño y su cantidad de capas totalmente conectadas.

Con el fin de generar los predictores necesarios para el entrenamiento de los diferentes modelos de clasificación, se aplicaron 11 medidas estadísticas a los datos de absorbancia de las muestras de café obtenidos mediante el análisis FTIR,

primero se realizaron las medidas estadísticas a la gráfica en general por lo que se obtuvieron 11 predictores para el entrenamiento de cada modelo, al realizar el entrenamiento, se obtuvo una exactitud baja en varios modelos, por lo tanto, se decidió aplicar dichas medidas estadísticas a los 11 compuestos definidos en los datos de absorbancia para determinados rangos de longitud de onda, obteniendo así 121 medidas estadísticas por muestra utilizadas como predictores, lo cual llevó a que la precisión de los modelos tuviera un aumento significativo, éste aumento se puede relacionar a que se evalúa cada compuesto representativo de los datos FTIR en el café, ya que un nivel de absorbancia en determinado compuesto, puede estar relacionado a un compuesto en una longitud de onda diferente, de éste modo analizar estadísticamente a cada compuesto y no a los datos de absorbancia en general, ayuda a representar de una mejor manera la muestra de café, mejorando así el desempeño de los modelos de clasificación.

6. RECOMENDACIONES

Para la implementación de la inteligencia artificial, MATLAB ofrece muchas herramientas que facilitan dicha implementación, sin embargo, se deben tener nociones básicas de los métodos que se desean utilizar para así poder utilizar sus funciones al máximo, al presente proyecto se le pueden aplicar posibles mejoras, como lo son la identificación de más características que permitan identificar o clasificar la calidad y posibles adulteraciones en las muestras de café.

Luego de seleccionar el modelo de clasificación que se desea emplear, se pueden realizar varios ajustes al modelo en cuestión mediante las diferentes opciones de configuración ofrecidas por la aplicación Classification Learner de MATLAB, por ejemplo, adaptando el modelo a los datos utilizados, lo cual puede ayudar a mejorar el rendimiento del modelo seleccionado.

El análisis FTIR es una metodología bastante usada y confiable en cuanto a la evaluación de la calidad de los alimentos, sin embargo, es importante conocer las condiciones de dichos alimentos como los posibles adulterantes y en este caso, las características de producción del café ya que no se logra evidenciar adulteraciones y cantidades específicas. Para el análisis de alimentos ya procesados, se podría combinar con otros métodos para poder mencionar más y mejores características.

BIBLIOGRAFÍA

ALVAREZ BALBUENA, Maria. Medidas estadísticas de centralización. [en línea]. Disponible en: https://rpubs.com/maria_alvarez_b/markdownm1

AYALA, Carlos. Determination of atenolol in pharmaceutical dosages by Fourier transform infrared spectrometry (FTIR). Revista Técnica de la Facultad de Ingeniería Universidad del Zulia. ISSN: 0254-0770

BODEN, Margaret, 2016. Inteligencia artificial. Madrid:Turner Publicaciones S.L., 2017, ISBN: 978-84-16714-22-3. [En línea]. Disponible en: https://books.google.com.co/books?hl=es&lr=&id=LCnYDwAAQBAJ&oi=fnd&pg=PT3&dq=Inteligencia+artificial+PDF&ots=drWpAVcMq6&sig=7CFI1P_mlxgp88w7KXoAUziYq5A#v=onepage&q&f=false

BARRIOS RODRIGUEZ, Yeison Fernando; SALAS CALDERON, Karen Tatiana y GIRÓN HERNÁNDEZ, Joel, 2018. Comparison of sensory attributes and chemical markers of Colombian coffee simples. En: Coffee Science, e151659, 2020. Doi: <https://doi.org/10.25186/cs.v15i.1659>

COZZOLINO, Daniel, 2015. Foodomics and infrared spectroscopy: from compounds to functionality. Current Opinion in Food Science. DOI: <https://doi.org/10.1016/j.cofs.2015.05.003>

CRAIG, Ana Paula; FRANCA, Adriana S y OLIVEIRA, Leonardo S, 2012a. Evaluation of the potential of FTIR and chemometrics for separation between defective and non-defective coffees. Food Chemistry. DOI: <https://doi.org/10.1016/j.foodchem.2011.11.121>

CRAIG, Ana Paula; FRANCA, Adriana S y OLIVEIRA, Leonardo S, 2012b. Discrimination between defective and non-defective roasted coffees by diffuse reflectance infrared Fourier transform spectroscopy. LWT - Food Science and Technology. DOI: doi.org/10.1016/j.lwt.2012.02.016

CRAIG, Ana Paula; FRANCA, Adriana S y OLIVEIRA, Leonardo S, 2018. Mid infrared spectroscopy and chemometrics as tools for the classification of roasted coffees by cup quality. Food Chemistry. DOI: doi.org/10.1016/j.foodchem.2017.11.066

CREMER, Dirk Rainer y KALETUNÇ, Gönül, 2003. Fourier transform infrared microspectroscopic study of the chemical microstructure of corn and oat flour-based extrudates. Carbohydrate Polymers. DOI: doi.org/10.1016/S0144-8617(02)00266-7

GARCÍA AHUED, Maricela, 2014. Análisis sensorial de alimentos. Padi Boletín Científico De Ciencias Básicas E Ingenierías Del ICBI. DOI: <https://doi.org/10.29057/icbi.v2i3.533>

GARRIGUES, J. M, et al, 2000. Fourier transform infrared determination of caffeine in roasted coffee samples. Fresenius J Anal Chem. DOI: doi.org/10.1007/s002160050063

HERNÁNDEZ, J. A: HEYD, B. y TRYSTRAM, G., 2008. Prediction of brightness and surface area kinetics during coffee roasting. Journal of Food Engineering. DOI: 10.1016 / j.jfoodeng.2008.04.026

HERNÁNDEZ, Juan Luis. Medidas de Posición: Cuantiles. [sitio web]. Disponible en: <https://www.tuveras.com/estadistica/estadistica02.htm>

Iberdrola, S.A. Qué es el 'machine learning'. [sitio web]. Disponible en: [https://www.iberdrola.com/innovacion/machine-learning-aprendizaje-automatico#:~:text=El%20Machine%20Learning%20es%20una,elaborar%20predicciones%20\(an%C3%A1lisis%20predictivo\)](https://www.iberdrola.com/innovacion/machine-learning-aprendizaje-automatico#:~:text=El%20Machine%20Learning%20es%20una,elaborar%20predicciones%20(an%C3%A1lisis%20predictivo))

JUÁREZ, Marc Gisbert. Medidas estadísticas. [sitio web]. Disponible en: <https://www.probabilidadyestadistica.net/medidas-estadisticas/>

KEMSLEY, E. K; RUAULT, S y WILSON, R.H, 1995. Discrimination between Coffea arabica and Coffea canephora variant robusta beans using infrared spectroscopy. Food Chemistry. DOI: doi.org/10.1016/0308-8146(95)00030-M

Labster Theory, Espectro electromagnético. [sitio web]. Disponible en: <https://theory.labster.com/electromagnetic-spectrum-es/>

LOHUMI, Santosh, et al, 2015. A review of vibrational spectroscopic techniques for the detection of food authenticity and adulteration. Trends in Food Science & Technology. DOI: <https://doi.org/10.1016/j.tifs.2015.08.003>

LYMAN, Donald. J, et al, 2003. FTIR-ATR analysis of brewed coffee: Effect of roasting conditions. Journal of Agricultural and Food Chemistry. DOI: 10.1021/jf0209793

LYMAN, Donald. J, et al, 2003. FTIR-ATR analysis of brewed coffee: Effect of roasting conditions. Journal of Agricultural and Food Chemistry. DOI: 10.1021/jf0209793

MANUEL FORTÚN, 2012. Importancia de la estadística. [Blog]. Disponible en: http://materiaestadistica.blogspot.com/2012/01/importancia-de-la-estadistica_15.html

MARTÍNEZ BENCARDINO, Ciro. Estadística y muestreo. Bogotá: Ecoe Ediciones, 2012. ISBN 978-958-648-702-3

MathWorks, Inc. App Classification Learner. Choose Classifier Options. [sitio web]. Disponible en: https://www-mathworks-com.translate.goog/help/stats/choose-a-classifier.html? x tr sl=en& x tr tl=es& x tr hl=es& x tr_pto=sc

MathWorks, Inc. App Classification Learner. fitcnet [sitio web]. Disponible en: https://www-mathworks-com.translate.goog/help/stats/fitcnet.html? x tr sl=en& x tr tl=es& x tr hl=es& x tr_pto=sc#mw_a9fa4524-aecf-4d9b-95ef-58914f7780f3

MathWorks, Inc. Classification Learner. [sitio web]. Disponible en: <https://la.mathworks.com/help/stats/classificationlearner-app.html>

MathWorks, Inc. Estructura red neuronal. [sitio web]. Disponible en: https://www-mathworks-com.translate.goog/help/stats/fitcnet.html? x tr sl=en& x tr tl=es& x tr hl=es& x tr_pto=sc#mw_a9fa4524-aecf-4d9b-95ef-58914f7780f3

MathWorks, Inc. Evaluación del rendimiento de los modelos de clasificación de machine learning. [sitio web]. Disponible en: <https://la.mathworks.com/discovery/roc-curve.html>.

MathWorks, Inc. Machine learning. [sitio web]. Disponible en: <https://la.mathworks.com/discovery/machine-learning.html>

MONDRAGÓN CORTEZ, Pedro. Espectroscopia de infrarrojo para todos. Centro de Investigación y Asistencia en Tecnología y Diseño del Estado de Jalisco, A. C., 2017. ISBN: 978-607-97548-4-6

NAVARRO FRÓMETA, Amado Enrique. Medidas de forma. [sitio web]. Disponible en: <https://navarrof.orgfree.com/Docencia/MatematicasIII/M3UT3/m3ut3mf.htm>

OPS, Clasificación de los alimentos y sus implicaciones en la salud, [sitio web]. Disponible en: <https://www3.paho.org/ecu/1135-clasificacion-alimentos-sus-implicaciones-salud.html>

PARADKAR, Manish M y IRUDAYARAJ, Joseph, 2004. Rapid determination of caffeine content in soft drinks using FTIR–ATR spectroscopy. Food Chemistry. DOI: doi.org/10.1016/S0308-8146(02)00116-4

PETTERI ROUHIAINEN, Lasse, 2018. Inteligencia artificial. 101 cosas que debes saber hoy sobre nuestro futuro. España: Editorial Planeta, S.A., 2018. ISBN: 978-84-17568-08-5. [En línea]. Disponible en: https://static0planetadelibroscom.cdnstatics.com/libros_contenido_extra/40/39308_Inteligencia_artificial.pdf

REIS, Nadia; FRANCA, Adriana S y OLIVEIRA, Leonardo S, 2013a. Discrimination between roasted coffee, roasted corn and coffee husks by Diffuse Reflectance Infrared Fourier Transform Spectroscopy. Food Science and Technology. DOI: 10.1016/j.lwt.2012.07.016

Satapathi lab. Fourier-Transform Infrared Spectroscopy (FTIR). [sitio web]. Disponible en: <https://satapathilab.iitr.ac.in/facilities/>

STATISTA, Consumo de café en Colombia en la temporada de cultivo 2022, por tipo. [sitio web]. Colombia, 2022, Disponible en: <https://es.statista.com/estadisticas/1284331/colombia-consumo-de-cafe-por-tipo/#:~:text=Colombia%3A%20consumo%20de%20caf%C3%A9%202021%2D2021%2C%20por%20tipo&text=Se%20pronostic%C3%B3%20que%20el%20consumo,del%20pa%C3%ADs%20sudamericano%20ese%20a%C3%B1o>

Tecno Soluciones, 2022. Espectroscopia infrarroja por transformada de Fourier y sus aplicaciones en la industria. [sitio web]. Disponible en: <https://tecnosolucionescr.net/blog/635-espectroscopia-infrarroja-por-transformada-de-fourier-y-sus-aplicaciones-en-la-industria>

TÉLLEZ MESA, Clara. Aplicaciones de la espectroscopía infrarroja en el análisis de alimentos. Facultad de farmacia universidad de Sevilla, 2019. 38 p.

the infrared spectrum between defective and non-defective

VALERIO, Yker, Café comercial versus café de especialidad: ¿cuál es la diferencia?. [en línea]. Era of We, Diciembre 11, 2021. Disponible en: <https://www.eraofwe.com/coffee-lab/es/articles/cafe-comercial-versus-cafe-de>

VANEGAS, René. Aplicaciones de inteligencia artificial para la clasificación automatizada de propósitos comunicativos en informes de ingeniería. [En línea]. Revista Signos, 2021. Disponible en: <https://www.semanticscholar.org/paper/Aplicaciones-de-inteligencia-artificial-para-la-de-Venegas/46a3502e80d904808a4427ae5154973fb8f3e30c>

VILLASANTE, Paula. Las medidas de dispersión en estadística. [sitio web]. Disponible en: <https://lamenteesmaravillosa.com/las-medidas-de-dispersion-en-estadistica/>

WANG, Jun, et al, 2009. Fourier transform infrared spectroscopy for kona coffee authentication. Journal of Food Science. DOI: 10.1111 / j.1750-3841.2009.01173.x Wikipedia, 26 de enero 2023. Matlab. [sitio web]. Disponible en: <https://es.wikipedia.org/wiki/MATLAB>

Zelad, Carlos. Evaluación de modelos de clasificación. [en línea]. Disponible en: <https://rpubs.com/chzelada/275494>

ZELAD, Carlos. Evaluación de modelos de clasificación. [en línea]. Disponible en: <https://rpubs.com/chzelada/275494>

ANEXOS

Anexo A. Algoritmo de clasificación.

```
% Se crea un almacen de datos con todos los archivos .csv
ds_graficas = datastore("*.csv");

numeroDeArchivos = length(ds_graficas.Files); % Variable que guarda la catidad
de archivos

% Arreglo de celdas para los diferentes compuestos
datosAbsorbanciaAgua = {};
datosAbsorbanciaCafeinalipidos = {};
datosAbsorbanciann = {};
datosAbsorbanciann2 = {};
datosAbsorbanciaLipido = {};
datosAbsorbanciaCafeina = {};
datosAbsorbanciaTrigonelina = {};
datosAbsorbanciaACG = {};
datosAbsorbanciaPolisacaridos = {};
datosAbsorbanciaAQuinico = {};
datosAbsorbanciaCarbohidratos = {};
%%
for i=1:1:numeroDeArchivos
    % Con la función "read" se obtiene el primer archivo del almacen de
    % datos, en en el siguiente ciclo se obtiene el segundo y así
    % sucesivamente
    grafica = read(ds_graficas);

    % Se obtienen los datos del eje Y segun el compuesto
    ejeYaAbsorbanciaAgua = grafica.YLabelAbsorbance(606:807);
    ejeYaAbsorbanciaCafeinalipidos = grafica.YLabelAbsorbance(552:605);
    ejeYaAbsorbanciann = grafica.YLabelAbsorbance(404:552);
    ejeYaAbsorbanciann2 = grafica.YLabelAbsorbance(297:405);
    ejeYaAbsorbanciaLipido = grafica.YLabelAbsorbance(278:287);
    ejeYaAbsorbanciaCafeina = grafica.YLabelAbsorbance(257:278);
    ejeYaAbsorbanciaTrigonelina = grafica.YLabelAbsorbance(176:257);
    ejeYaAbsorbanciaACG = grafica.YLabelAbsorbance(149:176);
    ejeYaAbsorbanciaPolisacaridos = grafica.YLabelAbsorbance(128:149);
    ejeYaAbsorbanciaAQuinico = grafica.YLabelAbsorbance(69:128);
    ejeYaAbsorbanciaCarbohidratos = grafica.YLabelAbsorbance(1:69);

    % La función reshape cambia el tamaño y la forma de un arreglo, se cambia
    % de columna a fila
    ejeYaAbsorbanciaAgua = reshape(ejeYaAbsorbanciaAgua,[1,202]);
    ejeYaAbsorbanciaCafeinalipidos =
    reshape(ejeYaAbsorbanciaCafeinalipidos,[1,54]);
    ejeYaAbsorbanciann = reshape(ejeYaAbsorbanciann,[1,149]);
    ejeYaAbsorbanciann2 = reshape(ejeYaAbsorbanciann2,[1,109]);
    ejeYaAbsorbanciaLipido = reshape(ejeYaAbsorbanciaLipido,[1,10]);
    ejeYaAbsorbanciaCafeina = reshape(ejeYaAbsorbanciaCafeina,[1,22]);
```

```

ejeYaAbsorbanciaTrigonelina = reshape(ejeYaAbsorbanciaTrigonelina,[1,82]);
ejeYaAbsorbanciaACG = reshape(ejeYaAbsorbanciaACG,[1,28]);
ejeYaAbsorbanciaPolisacaridos =
reshape(ejeYaAbsorbanciaPolisacaridos,[1,22]);
ejeYaAbsorbanciaAQuinico = reshape(ejeYaAbsorbanciaAQuinico,[1,60]);
ejeYaAbsorbanciaCarbohidratos =
reshape(ejeYaAbsorbanciaCarbohidratos,[1,69]);

% Se agrega de manera iterativa los valores correspondientes a cada
% arreglo de compuestos.
datosAbsorbanciaAgua{i, 1} = ejeYaAbsorbanciaAgua;
datosAbsorbanciaCafeinalipidos{i, 1} = ejeYaAbsorbanciaCafeinalipidos;
datosAbsorbanciann{i, 1} = ejeYaAbsorbanciann;
datosAbsorbanciann2{i, 1} = ejeYaAbsorbanciann2;
datosAbsorbanciaLipido{i, 1} = ejeYaAbsorbanciaLipido;
datosAbsorbanciaCafeina{i, 1} = ejeYaAbsorbanciaCafeina;
datosAbsorbanciaTrigonelina{i, 1} = ejeYaAbsorbanciaTrigonelina;
datosAbsorbanciaACG{i, 1} = ejeYaAbsorbanciaACG;
datosAbsorbanciaPolisacaridos{i, 1} = ejeYaAbsorbanciaPolisacaridos;
datosAbsorbanciaAQuinico{i, 1} = ejeYaAbsorbanciaAQuinico;
datosAbsorbanciaCarbohidratos{i, 1} = ejeYaAbsorbanciaCarbohidratos;

end
%%

% Se llama la función que calcula las medidas estadísticas y las guarda en
% una tabla
tablaDatosEstadisticos = medidasEstadisticas(datosAbsorbanciaAgua,
datosAbsorbanciaCafeinalipidos, ...
datosAbsorbanciann, datosAbsorbanciann2, datosAbsorbanciaLipido,
datosAbsorbanciaCafeina, ...
datosAbsorbanciaTrigonelina, datosAbsorbanciaACG,
datosAbsorbanciaPolisacaridos, datosAbsorbanciaAQuinico, ...
datosAbsorbanciaCarbohidratos);

% Se obtiene la clasificación de cafeina (Alto - Bajo)
clasificacionCafeina = readcell('clasificacion.csv','Range','B2:B61');
clasificacionCafeina = categorical(clasificacionCafeina);

% Se obtiene la clasificación del tipo de cafe (Instantaneo - Tostado molido)
clasificacionTipoCafe = readcell('clasificacion.csv','Range','C2:C61');
clasificacionTipoCafe = categorical(clasificacionTipoCafe);

% Se crea una tabla que contiene predictores y clasificación de cada
% muestra
tablaCafeina = tablaDatosEstadisticos;
tablaCafeina.class = clasificacionCafeina;

tablaTipoCafe = tablaDatosEstadisticos;
tablaTipoCafe.class = clasificacionTipoCafe;

%Se comienza a realizar el entrenamiento

```

```

n = length(tablaCafeina.class);
particionDeDatos = cvpartition(n,"Holdout",0.2);
%save('particionDeDatos')
%%
% Entrenamiento nivel de cafeina
idEntrenamiento = training(particionDeDatos);
datosEntrenamientoCafeina = tablaCafeina(idEntrenamiento,:);
fprintf("Datos de entrenamiento\n")
summary(datosEntrenamientoCafeina.class)
%%
% Entrenamiento tipo de cafe
idEntrenamiento = training(particionDeDatos);
datosEntrenamientoTipoCafe = tablaTipoCafe(idEntrenamiento,:);
fprintf("Datos de entrenamiento\n")
summary(datosEntrenamientoTipoCafe.class)
%%
% Testeo nivel de cafeina
idTest = test(particionDeDatos);
datosTestCafeina = tablaCafeina(idTest,:);
fprintf("Datos de prueba\n")
summary(datosTestCafeina.class)
%%
% Testeo tipo de cafe
idTest = test(particionDeDatos);
datosTestTipoCafe = tablaTipoCafe(idTest,:);
fprintf("Datos de prueba\n")
summary(datosTestTipoCafe.class)

% Se guardan todas las variables del workspace
save('variablesEntrenamiento')

```

Anexo B. Código de la función encargada de calcular las medidas estadísticas

```

function tablaDatosEstadisticos = medidasEstadisticas(grafAbsAgua,
grafAbsCafeiLipi, grafAbsn, grafAbsn2, grafAbsLipi,...
    grafAbsCafei, grafAbsTrigon, grafAbsACG, grafAbsPolisac, grafAbsAcQin,
grafAbsCarbo)

% Se declara una variable vacía de tipo tabla.
tablaDatosEstadisticos = table();

% Se calcula el máximo valor
tablaDatosEstadisticos.maximoAgua =
cellfun(@max,grafAbsAgua,'UniformOutput',true);
tablaDatosEstadisticos.maximoCafeinaLipidos =
cellfun(@max,grafAbsCafeiLipi,'UniformOutput',true);
tablaDatosEstadisticos.maximonn = cellfun(@max,grafAbsn,'UniformOutput',true);
tablaDatosEstadisticos.maximonn2 =
cellfun(@max,grafAbsn2,'UniformOutput',true);

```

```

tablaDatosEstadisticos.maximoLipido =
cellfun(@max,grafAbsLipi,'UniformOutput',true);
tablaDatosEstadisticos.maximoCafeina =
cellfun(@max,grafAbsCafei,'UniformOutput',true);
tablaDatosEstadisticos.maximoTrigonelina =
cellfun(@max,grafAbsTrigon,'UniformOutput',true);
tablaDatosEstadisticos.maximoACG =
cellfun(@max,grafAbsACG,'UniformOutput',true);
tablaDatosEstadisticos.maximoPolisacaridos =
cellfun(@max,grafAbsPolisac,'UniformOutput',true);
tablaDatosEstadisticos.maximoAcidoQuinico =
cellfun(@max,grafAbsAcQin,'UniformOutput',true);
tablaDatosEstadisticos.maximoCarbohidratos =
cellfun(@max,grafAbsCarbo,'UniformOutput',true);

% Se calcula el mínimo valor
tablaDatosEstadisticos.minimoAgua =
cellfun(@min,grafAbsAgua,'UniformOutput',true);
tablaDatosEstadisticos.minimoCafeinaLipidos =
cellfun(@min,grafAbsCafeiLipi,'UniformOutput',true);
tablaDatosEstadisticos.minimonn = cellfun(@min,grafAbsnn,'UniformOutput',true);
tablaDatosEstadisticos.minimonn2 =
cellfun(@min,grafAbsnn2,'UniformOutput',true);
tablaDatosEstadisticos.minimoLipido =
cellfun(@min,grafAbsLipi,'UniformOutput',true);
tablaDatosEstadisticos.minimoCafeina =
cellfun(@min,grafAbsCafei,'UniformOutput',true);
tablaDatosEstadisticos.minimoTrigonelina =
cellfun(@min,grafAbsTrigon,'UniformOutput',true);
tablaDatosEstadisticos.minimoACG =
cellfun(@min,grafAbsACG,'UniformOutput',true);
tablaDatosEstadisticos.minimoPolisacaridos =
cellfun(@min,grafAbsPolisac,'UniformOutput',true);
tablaDatosEstadisticos.minimoAcidoQuinico =
cellfun(@min,grafAbsAcQin,'UniformOutput',true);
tablaDatosEstadisticos.minimoCarbohidratos =
cellfun(@min,grafAbsCarbo,'UniformOutput',true);

% Se calcula la media
tablaDatosEstadisticos.mediaAgua =
cellfun(@mean,grafAbsAgua,'UniformOutput',true);
tablaDatosEstadisticos.mediaCafeinaLipidos =
cellfun(@mean,grafAbsCafeiLipi,'UniformOutput',true);
tablaDatosEstadisticos.mediann = cellfun(@mean,grafAbsnn,'UniformOutput',true);
tablaDatosEstadisticos.mediann2 =
cellfun(@mean,grafAbsnn2,'UniformOutput',true);
tablaDatosEstadisticos.mediaLipido =
cellfun(@mean,grafAbsLipi,'UniformOutput',true);
tablaDatosEstadisticos.mediaCafeina =
cellfun(@mean,grafAbsCafei,'UniformOutput',true);
tablaDatosEstadisticos.mediaTrigonelina =
cellfun(@mean,grafAbsTrigon,'UniformOutput',true);

```



```

tablaDatosEstadisticos.mediaACG =
cellfun(@mean,grafAbsACG,'UniformOutput',true);
tablaDatosEstadisticos.mediaPolisacaridos =
cellfun(@mean,grafAbsPolisac,'UniformOutput',true);
tablaDatosEstadisticos.mediaAcidoQuinico =
cellfun(@mean,grafAbsAcQin,'UniformOutput',true);
tablaDatosEstadisticos.mediaCarbohidratos =
cellfun(@mean,grafAbsCarbo,'UniformOutput',true);

% Se calcula la mediana
tablaDatosEstadisticos.medianaAgua =
cellfun(@median,grafAbsAgua,'UniformOutput',true);
tablaDatosEstadisticos.medianaCafeinaLipidos =
cellfun(@median,grafAbsCafeiLipi,'UniformOutput',true);
tablaDatosEstadisticos.medianann =
cellfun(@median,grafAbsnn,'UniformOutput',true);
tablaDatosEstadisticos.medianann2 =
cellfun(@median,grafAbsnn2,'UniformOutput',true);
tablaDatosEstadisticos.medianaLipido =
cellfun(@median,grafAbsLipi,'UniformOutput',true);
tablaDatosEstadisticos.medianaCafeina =
cellfun(@median,grafAbsCafei,'UniformOutput',true);
tablaDatosEstadisticos.medianaTrigonelina =
cellfun(@median,grafAbsTrigon,'UniformOutput',true);
tablaDatosEstadisticos.medianaACG =
cellfun(@median,grafAbsACG,'UniformOutput',true);
tablaDatosEstadisticos.medianaPolisacaridos =
cellfun(@median,grafAbsPolisac,'UniformOutput',true);
tablaDatosEstadisticos.medianaAcidoQuinico =
cellfun(@median,grafAbsAcQin,'UniformOutput',true);
tablaDatosEstadisticos.medianaCarbohidratos =
cellfun(@median,grafAbsCarbo,'UniformOutput',true);

```

```

% Se calcula la desviación estándar
tablaDatosEstadisticos.desviacionEstandarAgua =
cellfun(@std,grafAbsAgua,'UniformOutput',true);
tablaDatosEstadisticos.desviacionEstandarCafeinaLipidos =
cellfun(@std,grafAbsCafeiLipi,'UniformOutput',true);
tablaDatosEstadisticos.desviacionEstandarnn =
cellfun(@std,grafAbsnn,'UniformOutput',true);
tablaDatosEstadisticos.desviacionEstandarnn2 =
cellfun(@std,grafAbsnn2,'UniformOutput',true);
tablaDatosEstadisticos.desviacionEstandarLipido =
cellfun(@std,grafAbsLipi,'UniformOutput',true);
tablaDatosEstadisticos.desviacionEstandarCafeina =
cellfun(@std,grafAbsCafei,'UniformOutput',true);
tablaDatosEstadisticos.desviacionEstandarTrigonelina =
cellfun(@std,grafAbsTrigon,'UniformOutput',true);
tablaDatosEstadisticos.desviacionEstandarACG =
cellfun(@std,grafAbsACG,'UniformOutput',true);
tablaDatosEstadisticos.desviacionEstandarPolisacaridos =
cellfun(@std,grafAbsPolisac,'UniformOutput',true);

```

```

tablaDatosEstadisticos.desviacionEstandarAcidoQuinico =
cellfun(@std,grafAbsAcQin,'UniformOutput',true);
tablaDatosEstadisticos.desviacionEstandarCarbohidratos =
cellfun(@std,grafAbsCarbo,'UniformOutput',true);

% Se calcula la desviación media absoluta
tablaDatosEstadisticos.desviacionMediaAbsolutaAgua =
cellfun(@mad,grafAbsAgua,'UniformOutput',true);
tablaDatosEstadisticos.desviacionMediaAbsolutaCafeinaLipidos =
cellfun(@mad,grafAbsCafeiLipi,'UniformOutput',true);
tablaDatosEstadisticos.desviacionMediaAbsolutann =
cellfun(@mad,grafAbsnn,'UniformOutput',true);
tablaDatosEstadisticos.desviacionMediaAbsolutann2 =
cellfun(@mad,grafAbsnn2,'UniformOutput',true);
tablaDatosEstadisticos.desviacionMediaAbsolutaLipido =
cellfun(@mad,grafAbsLipi,'UniformOutput',true);
tablaDatosEstadisticos.desviacionMediaAbsolutaCafeina =
cellfun(@mad,grafAbsCafei,'UniformOutput',true);
tablaDatosEstadisticos.desviacionMediaAbsolutaTrigonelina =
cellfun(@mad,grafAbsTrigon,'UniformOutput',true);
tablaDatosEstadisticos.desviacionMediaAbsolutaACG =
cellfun(@mad,grafAbsACG,'UniformOutput',true);
tablaDatosEstadisticos.desviacionMediaAbsolutaPolisacaridos =
cellfun(@mad,grafAbsPolisac,'UniformOutput',true);
tablaDatosEstadisticos.desviacionMediaAbsolutaAcidoQuinico =
cellfun(@mad,grafAbsAcQin,'UniformOutput',true);
tablaDatosEstadisticos.desviacionMediaAbsolutaCarbohidratos =
cellfun(@mad,grafAbsCarbo,'UniformOutput',true);

% Se calcula el percentil 25
tablaDatosEstadisticos.quantile25Agua = cellfun(@(x)
quantile(x,0.25),grafAbsAgua,'UniformOutput',true);
tablaDatosEstadisticos.quantile25CafeinaLipidos = cellfun(@(x)
quantile(x,0.25),grafAbsCafeiLipi,'UniformOutput',true);
tablaDatosEstadisticos.quantile25nn = cellfun(@(x)
quantile(x,0.25),grafAbsnn,'UniformOutput',true);
tablaDatosEstadisticos.quantile25nn2 = cellfun(@(x)
quantile(x,0.25),grafAbsnn2,'UniformOutput',true);
tablaDatosEstadisticos.quantile25Lipido = cellfun(@(x)
quantile(x,0.25),grafAbsLipi,'UniformOutput',true);
tablaDatosEstadisticos.quantile25Cafeina = cellfun(@(x)
quantile(x,0.25),grafAbsCafei,'UniformOutput',true);
tablaDatosEstadisticos.quantile25Trigonelina = cellfun(@(x)
quantile(x,0.25),grafAbsTrigon,'UniformOutput',true);
tablaDatosEstadisticos.quantile25ACG = cellfun(@(x)
quantile(x,0.25),grafAbsACG,'UniformOutput',true);
tablaDatosEstadisticos.quantile25Polisacaridos = cellfun(@(x)
quantile(x,0.25),grafAbsPolisac,'UniformOutput',true);
tablaDatosEstadisticos.quantile25AcidoQuinico = cellfun(@(x)
quantile(x,0.25),grafAbsAcQin,'UniformOutput',true);
tablaDatosEstadisticos.quantile25Carbohidratos = cellfun(@(x)
quantile(x,0.25),grafAbsCarbo,'UniformOutput',true);

```

```

% Se calcula el percentil 75
tablaDatosEstadisticos.quantile75Agua = cellfun(@(x)
quantile(x,0.75),grafAbsAgua,'UniformOutput',true);
tablaDatosEstadisticos.quantile75Cafeinalipidos = cellfun(@(x)
quantile(x,0.75),grafAbsCafeiLipi,'UniformOutput',true);
tablaDatosEstadisticos.quantile75nn = cellfun(@(x)
quantile(x,0.75),grafAbsnn,'UniformOutput',true);
tablaDatosEstadisticos.quantile75nn2 = cellfun(@(x)
quantile(x,0.75),grafAbsnn2,'UniformOutput',true);
tablaDatosEstadisticos.quantile75Lipido = cellfun(@(x)
quantile(x,0.75),grafAbsLipi,'UniformOutput',true);
tablaDatosEstadisticos.quantile75Cafeina = cellfun(@(x)
quantile(x,0.75),grafAbsCafei,'UniformOutput',true);
tablaDatosEstadisticos.quantile75Trigonelina = cellfun(@(x)
quantile(x,0.75),grafAbsTrigon,'UniformOutput',true);
tablaDatosEstadisticos.quantile75ACG = cellfun(@(x)
quantile(x,0.75),grafAbsACG,'UniformOutput',true);
tablaDatosEstadisticos.quantile75Polisacaridos = cellfun(@(x)
quantile(x,0.75),grafAbsPolisac,'UniformOutput',true);
tablaDatosEstadisticos.quantile75AcidoQuinico = cellfun(@(x)
quantile(x,0.75),grafAbsAcQin,'UniformOutput',true);
tablaDatosEstadisticos.quantile75Carbohidratos = cellfun(@(x)
quantile(x,0.75),grafAbsCarbo,'UniformOutput',true);

```

```

% Se calcula el rango intercuartílico
tablaDatosEstadisticos.signalIQRAgua =
cellfun(@iqr,grafAbsAgua,'UniformOutput',true);
tablaDatosEstadisticos.signalIQRCafeinalipidos =
cellfun(@iqr,grafAbsCafeiLipi,'UniformOutput',true);
tablaDatosEstadisticos.signalIQRnn =
cellfun(@iqr,grafAbsnn,'UniformOutput',true);
tablaDatosEstadisticos.signalIQRnn2 =
cellfun(@iqr,grafAbsnn2,'UniformOutput',true);
tablaDatosEstadisticos.signalIQLipido =
cellfun(@iqr,grafAbsLipi,'UniformOutput',true);
tablaDatosEstadisticos.signalIQRCafeina =
cellfun(@iqr,grafAbsCafei,'UniformOutput',true);
tablaDatosEstadisticos.signalIQRTrigonelina =
cellfun(@iqr,grafAbsTrigon,'UniformOutput',true);
tablaDatosEstadisticos.signalIQRACG =
cellfun(@iqr,grafAbsACG,'UniformOutput',true);
tablaDatosEstadisticos.signalIQRPolisacaridos =
cellfun(@iqr,grafAbsPolisac,'UniformOutput',true);
tablaDatosEstadisticos.signalIQRAcidoQuinico =
cellfun(@iqr,grafAbsAcQin,'UniformOutput',true);
tablaDatosEstadisticos.signalIQRCarbohidratos =
cellfun(@iqr,grafAbsCarbo,'UniformOutput',true);

```

```

% Se calcula la asimetría
tablaDatosEstadisticos.asimetriaAgua =
cellfun(@skewness,grafAbsAgua,'UniformOutput',true);
tablaDatosEstadisticos.asimetriaCafeinalipidos =
cellfun(@skewness,grafAbsCafeiLipi,'UniformOutput',true);

```

```

tablaDatosEstadisticos.asimetriann =
cellfun(@skewness,grafAbsnn,'UniformOutput',true);
tablaDatosEstadisticos.asimetriann2 =
cellfun(@skewness,grafAbsnn2,'UniformOutput',true);
tablaDatosEstadisticos.asimetriaLipido =
cellfun(@skewness,grafAbsLipi,'UniformOutput',true);
tablaDatosEstadisticos.asimetriaCafeina =
cellfun(@skewness,grafAbsCafei,'UniformOutput',true);
tablaDatosEstadisticos.asimetriaTrigonelina =
cellfun(@skewness,grafAbsTrigon,'UniformOutput',true);
tablaDatosEstadisticos.asimetriaACG =
cellfun(@skewness,grafAbsACG,'UniformOutput',true);
tablaDatosEstadisticos.asimetriaPolisacaridos =
cellfun(@skewness,grafAbsPolisac,'UniformOutput',true);
tablaDatosEstadisticos.asimetriaAcidoQuinico =
cellfun(@skewness,grafAbsAcQin,'UniformOutput',true);
tablaDatosEstadisticos.asimetriaCarbohidratos =
cellfun(@skewness,grafAbsCarbo,'UniformOutput',true);

% Se calcula la curtosis
tablaDatosEstadisticos.curtosisAgua =
cellfun(@kurtosis,grafAbsAgua,'UniformOutput',true);
tablaDatosEstadisticos.curtosisCafeinaLipidos =
cellfun(@kurtosis,grafAbsCafeiLipi,'UniformOutput',true);
tablaDatosEstadisticos.curtosisnn =
cellfun(@kurtosis,grafAbsnn,'UniformOutput',true);
tablaDatosEstadisticos.curtosisnn2 =
cellfun(@kurtosis,grafAbsnn2,'UniformOutput',true);
tablaDatosEstadisticos.curtosisLipido =
cellfun(@kurtosis,grafAbsLipi,'UniformOutput',true);
tablaDatosEstadisticos.curtosisCafeina =
cellfun(@kurtosis,grafAbsCafei,'UniformOutput',true);
tablaDatosEstadisticos.curtosisTrigonelina =
cellfun(@kurtosis,grafAbsTrigon,'UniformOutput',true);
tablaDatosEstadisticos.curtosisACG =
cellfun(@kurtosis,grafAbsACG,'UniformOutput',true);
tablaDatosEstadisticos.curtosisPolisacaridos =
cellfun(@kurtosis,grafAbsPolisac,'UniformOutput',true);
tablaDatosEstadisticos.curtosisAcidoQuinico =
cellfun(@kurtosis,grafAbsAcQin,'UniformOutput',true);
tablaDatosEstadisticos.curtosisCarbohidratos =
cellfun(@kurtosis,grafAbsCarbo,'UniformOutput',true);

end

```