



**UNIVERSIDAD
DEL AZUAY**

FACULTAD DE CIENCIA Y TECNOLOGÍA

INGENIERÍA EN ALIMENTOS

**Aplicación de técnicas quimiométricas en la interpretación
de espectros Raman de diversos tipos de cerveza.**

Trabajo de graduación previo a la obtención del título de:

INGENIERA EN ALIMENTOS

Autora:

Soraya Gallegos

Director:

Piercosimo Tripaldi

Cuenca, Ecuador

Enero, 2021

DEDICATORIA.

Este trabajo de investigación se los dedico con mucho amor y gratitud a mis amados padres, hermanos y demás familia por estar conmigo en todo momento y darme el apoyo incondicional en cada situación que he tenido que afrontar, es su ejemplo el que me impulsa a seguir adelante siempre, porque sin ellos hubiese sido difícil culminar esta investigación.

AGRADECIMIENTOS

A Dios por guiarme e iluminarme a cada momento y darme la fortaleza necesaria para llevar a cabo cada paso de este trabajo.

A mis padres, hermanos y demás familia por el apoyo tanto moral como económico brindado de forma incondicional en todo el transcurso de este largo camino.

A la Universidad del Azuay de manera especial a la Carrera de Ingeniería en Alimentos por abrirme las puertas para forjarme como profesional, pues de sus aulas me llevo los mejores aprendizajes y bellos recuerdos.

A mi tutor de tesis Dr. Piercosimo Tripaldi por compartirme sus conocimientos y brindarme la paciencia y apoyo necesario para llevar a cabo este proyecto lo cual hizo posible la consecución del mismo, de igual manera al Dr. Cristian Rojas y al Ing. Andrés Pérez, quienes me brindaron todo el apoyo necesario para el desarrollo de este trabajo de investigación.

Finalmente, a todos mis amigos y compañeros de la carrera de Ingeniería en Alimentos quienes me brindaron su apoyo y amistad durante estos cuatro años de universidad.

Aplicación de técnicas quimiométricas en la interpretación de espectros Raman de diversos tipos de cerveza.

RESUMEN

En el presente trabajo de investigación se desarrolló un modelo matemático utilizando quimiometría para la identificación de parámetros fisicoquímicos en espectros Raman de cervezas. Los espectros fueron previamente pretratados, posteriormente se realizó un análisis exploratorio para luego ser sometidos a un proceso de selección supervisada de variables mediante Algoritmos Genéticos acoplados con el método de clasificación de los k -vecinos más cercanos (k NN). El algoritmo genético identificó los picos de los espectros que pertenecen al grado alcohólico y el amargor de la cerveza. Se obtuvo un modelo de clasificación confiable al tener una tasa de aciertos de (NER) de 0,98 tanto en calibración, validación y predicción.

Palabras clave: cerveza, espectroscopía Raman, algoritmos genéticos, k NN,



Dr. Piercosimo Tripaldi

Director de tesis



Ing. María Fernanda Rosales

**Coordinadora de la Escuela
Ingeniería en Alimentos**



Soraya Elizabeth Gallegos Torres

Autora

APPLICATION OF CHEMOMETRIC TECHNIQUES IN THE INTERPRETATION OF RAMAN SPECTRA OF VARIOUS TYPES OF BEER.

ABSTRACT

In this research a mathematic model using chemometry was developed to identify the physiochemical parameters in Raman spectra of beers. The spectra were previously pretreated. After, an exploratory analysis was carried. Later, the spectra were subject to a supervised selection process of variables though coupled genetic algorithms with the classification method of the k-nearest neighbors (kNN). The genetic algorithm identified the peaks of the spectra that belonged to the alcoholic grade and the bitterness of the beer. A trustworthy classification model was obtained by having a hit rate of (NER) from 0.98 in both calibration, validation and prediction.

Key words: beer, Raman spectroscopy, genetic algorithms, kNN.



Dr. Piercosimo Tripaldi

**Thesis Director
Coordinator**



Ing. María Fernanda Rosales

Food Engineering Faculty



Soraya Elizabeth Gallegos Torres

Author



ÍNDICE DE CONTENIDOS

DEDICATORIA.....	ii
AGRADECIMIENTOS	iii
RESUMEN	iv
ABSTRACT	v
ÍNDICE DE CONTENIDOS.....	vi
ÍNDICE DE TABLAS.....	viii
ÍNDICE DE FIGURAS.....	ix
INTRODUCCIÓN.....	1
Cerveza.....	1
Producción de la cerveza.....	1
Tipos de cerveza.....	1
Tipos de cerveza según su fermentación	2
Calidad de la cerveza.....	4
Espectroscopía Raman	4
Fundamento de la espectroscopía Raman	5
Dispositivo Raman.....	5
Quimiometría.....	7
METODOLOGÍA	8
Construcción de base de datos	8
Curado de la base de datos	8
Análisis exploratorio.....	8
Análisis de conglomerados.....	9
Análisis de componentes principales	9
Modelos supervisados.....	10
Modelos de clasificación (<i>k</i> NN).....	10
Selección de variables	10
Validación del Modelo	11
Validación cruzada	11
Validación externa	11
RESULTADOS	12
Generación de la base de datos	12
Análisis de componentes principales	12
Análisis de conglomerados.....	13

Modelo de clasificación.....	15
DISCUSIONES.....	16
CONCLUSIONES.....	17
BIBLIOGRAFÍA.....	18
ANEXOS.....	20
Anexo 1. Muestras de 118 cervezas procedentes de 20 países con sus respectivos parámetros fisicoquímicos	20
Anexo 2. Variables seleccionadas por lo algoritmos genéticos	26

ÍNDICE DE TABLAS

Tabla 1. Diferentes tipos de cerveza según su fermentación.....	2
Tabla 2. Parámetros de calidad del modelo obtenidos con el método de clasificación kNN	15

ÍNDICE DE FIGURAS

Figura: 1 Espectro Raman de una cerveza (azul), espectro Raman del agua de referencia (rojo)	5
Figura: 2 Esquema de un espectroscopio raman	6
Figura: 3 Resultados del análisis del componente principal de los espectros Raman	13
Figura: 4 Dendrograma de similitud de los espectros de las cervezas; se notan los cuatro grupos como en el análisis PCA.	14
Figura: 5 Grafica de los centroides de las longitudes de onda seleccionadas, de sus derivadas segunda y de las características de las cervezas ABV y IBU.	14

Soraya Elizabeth Gallegos Torres

Trabajo de Titulación

Dr. Piercosimo Tripaldi

Enero, 2021

Aplicación de técnicas quimiométricas en la interpretación de espectros Raman de diversos tipos de cerveza.

INTRODUCCIÓN

Cerveza

La cerveza es la bebida resultante de la fermentación alcohólica, mediante levadura seleccionada, de un mosto procedente de malta de cebada, solo o mezclado con otros productos amiláceos transformables en azúcares por digestión enzimática (malta de otros cereales, granos crudos que contengan féculas, así como azúcares y féculas, siempre que estas sustancias añadidas no excedan del 50 por ciento en masa de la materia prima empleada), al cual se agrega lúpulo y/o sus derivados y se somete a un proceso de cocción (IICA, 1999).

Producción de la cerveza

En la actualidad, la cerveza se produce en todo el mundo, el número de conocedores y aficionados no deja de crecer, y sus exigencias obligan a los proveedores y fabricantes de cerveza a una especialización cada vez mayor y a una búsqueda constante en mercados lejanos o poco habituales.

La producción mundial de cerveza, industrial y artesanal, es de aproximadamente 1500 millones de hectolitros cada año, y lo que realmente ha supuesto un impulso para el mercado ha sido la producción de cervezas artesanales de calidad. El crecimiento de estas pequeñas empresas cerveceras es constante, lo que permite emprender nuevos caminos conjugando las innovaciones con el saber tradicional para obtener así nuevos resultados (Freixes & Punsola, 2014)

Tipos de cerveza

En el mundo existen diversas clases de cerveza, cada una con un particular aroma, sabor, color y cuerpo. Muchas veces llevan el nombre de los pueblos de los cuales son originarias. Si bien todas se fabrican con los mismos ingredientes, cebada malteada, lúpulo, levadura y agua, lo que establece la diferencia entre una y otra son las variaciones de las materias primas y el tipo de fermentación experimentada. En la Tabla 1. se muestran diferentes tipos de cerveza de acuerdo al tipo de fermentación.

Tipos de cerveza según su fermentación

Fermentación Baja

Son las llamadas lager, el proceso de fermentación se realiza a temperaturas bajas (10 a 12°C), y en él la levadura se mantiene al fondo del estanque permitiendo que el lúpulo y la cebada malteada dominen el aroma y sabor del producto. Suelen ser ligeras, espumosas, suaves, de color ambarino o negro (De Clerck, 1957).

Los mostos para baja fermentación tradicionalmente se han macerado por infusión escalonada o decocción con malta menos modificada, con niveles altos de proteínas

Fermentación alta

Son las denominadas Ale, la fermentación de esta cerveza ocurre de manera más rápida y a temperaturas de 20°C aproximadamente, actuando la levadura en las superficies del mosto. Los mostos se maceran por infusión con malta muy modificada, con niveles bajos de proteínas. Las altas temperaturas favorecen la producción de ésteres, que pueden añadir notas frutales, pero también de alcoholes superiores y otros subproductos no deseables. Tiene un elevado porcentaje de alcohol y son muy aromáticas (De Clerck, 1957).

La cerveza Ale se diferencia de la cerveza lager por la disminución más rápida del extracto de azúcar en la etapa de fermentación, causado por el uso de la levadura *Saccharomyces cerevisiae*, que permanece en suspensión, y por las altas temperaturas utilizadas (20-23°C) (Knudsen, 1977).

Tabla 1. Diferentes tipos de cerveza según su fermentación

LOS PRINCIPALES TIPOS DE CERVEZA		
Fermentación alta	Fermentación baja	Fermentación espontánea
ABBAYE (de abadía)	LAGER	LAMBIC
Dubbel	American lager	Faro
Tripel	Golden lager	Gueuze
Quadrupel	Dark lager	Cassis
TRAPPISTE (trapense)	Strong lager y malt liquor	Framboise
Dubbel	Munchner helles	Pécheresse
Tripel	Munchner dunkel	Kriek
ALE	Light beer	
Belgian strong ale	Icebeer	
Belgian red ale	Eisbier	
Dark ale	Exóticas	
Irish red ale	Kellerbier	
Scotch ale	Kulmbacher	
Mild ale	BOCK	
Old ale	Doppelbock	

Pale ale	Eisbock	
English brow ale	Hellesbock	
Bitter ale (strong bitter)	Maibock	
Cream ale	PILSNER O PILS	
Indian pale ale	Deutscher pilsner	
Iber ale	American pilsner o premium lager	
British ale	Dortmunder o export	
PORTER	Biere de garde	
STOUT	Biere de mars-marzen-oktoberfestbier	
Irish stout	Rauchbier	
Imperial stout	Schwarzbier	
Oatmeal stout	Vienna	
Sweet o milk stout	Kellerbier	
Stout chocolate		
Dry stout		
Stout ale		
BIERE BLANCHE WEIZEN – WEISSBIER (cerveza blanca)		
Hefe weizen		
Kristall weizwn		
Weizenbock		
Berliner weizen		
Weizen dunkel		
Weizen gose		
Wheat beer		
Roggenbier		
Altbier		
kolschbier		
Barley wine – vintage		
Saison		
Steam beer		
Steinbier		

Fuente: (IICA, 1999) .

Calidad de la cerveza

Grados de alcohol

El alcohol se forma durante en la etapa de fermentación del mosto (proceso anaeróbico), mediante el cual la levadura convierte la glucosa en etanol y dióxido de carbono acorde a la siguiente fórmula $C_6H_{12}O_6 + 2C_2H_5OH + 2CO_2 + \text{calor}$. Los principales productos de fermentación son etanol y CO_2 ; sin embargo, también se forman numerosos subproductos del crecimiento de levaduras, que contribuyen de forma importante al perfume y aroma de la cerveza.

El contenido de alcohol (volumen) de las bebidas alcohólicas se refiere al volumen de alcohol (etanol) contenido en 100 volúmenes de productos medidos a una temperatura de 20 °C. Es una medida del porcentaje de concentración de volumen ABV.

Amargor

El sabor amargo característico de la cerveza, es la contrapartida del dulzor conferido por la malta. La intensidad y la calidad varían mucho de acuerdo al estilo. El amargor es el resultado de los iso- α -ácidos de los lúpulos en el mosto durante la cocción. La intensidad y la calidad del amargor se manipula a través de la elección de la variedad de los lúpulos, el contenido de α -ácidos o a veces la cantidad de β -ácidos oxidados que se cree que dan un tipo de amargor suave, la duración de la cocción de los lúpulos, y el contenido mineral del licor. El parámetro más utilizado para clasificar las cervezas en función de su amargor es el IBU (Internacional Bittering Units), unidades internacionales de amargor. El IBU es una medida de concentración que se expresa en miligramos de iso- α -ácidos en un litro de mosto o cerveza o en partes por millón (ppm)

Espectroscopía Raman

La espectroscopía Raman ha cautivado a la comunidad científica por la extraordinaria información que proporciona a cerca del material que se analiza. Es una técnica fotónica de alta resolución que proporciona en pocos segundos información química y estructural de cualquier material o compuesto orgánico y/o inorgánico, permitiendo así su identificación. Se trata de una técnica analítica no invasiva que se realiza directamente sobre el material, sin apenas requerir preparación previa de la muestra, lo que otorga una importancia especial de cara a sus posibles aplicaciones (Medina, Sato, & Frausto, 2003).

El análisis mediante espectroscopía Raman se basa en irradiar un haz de luz con frecuencia monocromática sobre una muestra cuyas propiedades moleculares se van determinar, y examinar la luz dispersada por dicha muestra. La mayor parte de luz dispersada presenta la misma frecuencia que la luz incidente pero una pequeña parte muestra un cambio de frecuencia, que es el resultado de la interacción de la luz con el material. La luz que mantiene la misma frecuencia que la luz incidente se denomina dispersión Rayleigh, la cual no proporciona ninguna información sobre la composición de la muestra analizada, mientras que la luz dispersada que

presenta frecuencias distintas a la de la radiación incidente, es la que proporciona información sobre la composición molecular de la muestra y es la que se conoce como dispersión Raman (Pelletier, 1999).

Los espectros Raman pueden proporcionar multitud de información. Principalmente, cada molécula o especie química tiene su propio espectro Raman, proporcionando su correspondiente información estructural. Cada pico del espectro está relacionado unívocamente con un modo de vibración, rotación o cualquier otro modo de baja frecuencia, independientemente de la longitud de onda del láser empleado. En la Figura: 1 se presenta un espectro Raman de una cerveza utilizada en esta investigación.

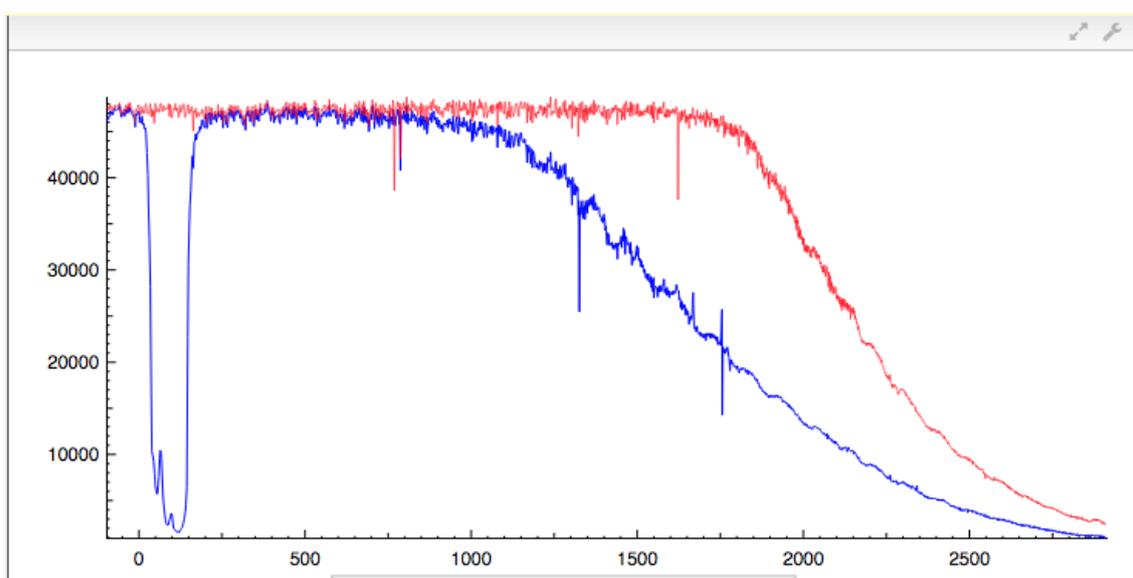


Figura: 1 Espectro Raman de una cerveza (azul) y espectro Raman del agua de referencia (rojo)

Fundamento de la espectroscopía Raman

El espectro Raman se obtiene irradiando una muestra con una potente fuente láser de radiación monocromática visible, ultravioleta o infrarroja, y con un espectrómetro adecuado se registra la radiación dispersada a un cierto ángulo. Para que una radiación sea activa en el efecto Raman debe cambiar de polarizabilidad de la molécula. Es decir, la nube electrónica de la molécula debe deformarse más fácilmente en un extremo que en el otro durante la vibración o rotación (Valderrama, 1999). Sus dos principales ventajas, frente a otras técnicas de análisis químico y estructural, son el carácter no destructivo y alta resolución espacial (Medina, Sato, & Frausto, 2003).

Dispositivo Raman

En las últimas décadas los equipos Raman han evolucionado a pasos agigantados, en particular desde el descubrimiento del láser en los años 60, y con la aparición de la fibra óptica y los CCDs

(Dispositivos de cargas acopladas) en los 80 (Willard, 1988). El dispositivo utilizado en esta investigación para la obtención del espectro Raman fue el FirstDefender RMX, que es un instrumento robusto con tecnología Raman de última generación para la identificación rápida y precisa de productos químicos desconocidos.

En la Figura: 2 se presenta el esquema típico de un espectrómetro Raman. Generalmente el proceso empieza con un haz de luz láser transmitido por fibra óptica hasta el cabezal óptico, el cual proviene de una fuente láser de onda continua. En el cabezal existe un filtro interferencial para depurar frecuencias y focalizar sobre el objetivo, y un filtro “*notch*” que elimina la frecuencia Rayleigh.

A continuación, la radiación dispersada es captada por el cabezal y pasa directamente de la fibra óptica al espectrómetro, atravesando una rendija estrecha de entrada que corta este primer haz. Seguidamente llega al monocromador que se encarga de separar espacialmente todas las frecuencias distintas. En la mayoría de estos dispositivos, la luz divergente se pone en paralelo después de pasar a través de una lente cóncava, y se dirige hacia una rejilla. La rejilla divide gradualmente los componentes espectrales de la luz en diferentes ángulos, después estos se reenfojan usando de nuevo espejo cóncavo que proyecta la imagen en el detector.

Una vez que la luz produce una imagen en el detector, los fotones se convierten en electrones, que luego se digitalizan y leen a través de la entrada USB. Luego el programa se basa en el número de píxeles del detector y dispersión lineal de la rendija de difracción para crear una calibración que permita representar los datos obtenidos en función del desplazamiento Raman en todo el rango espectral. Aquí, el espectro se puede comparar con otros espectros de referencia y también se puede manipular y procesar a conveniencia del propósito experimental gracias a un software específico (SKoog, Holler, & Crouch, 2006)

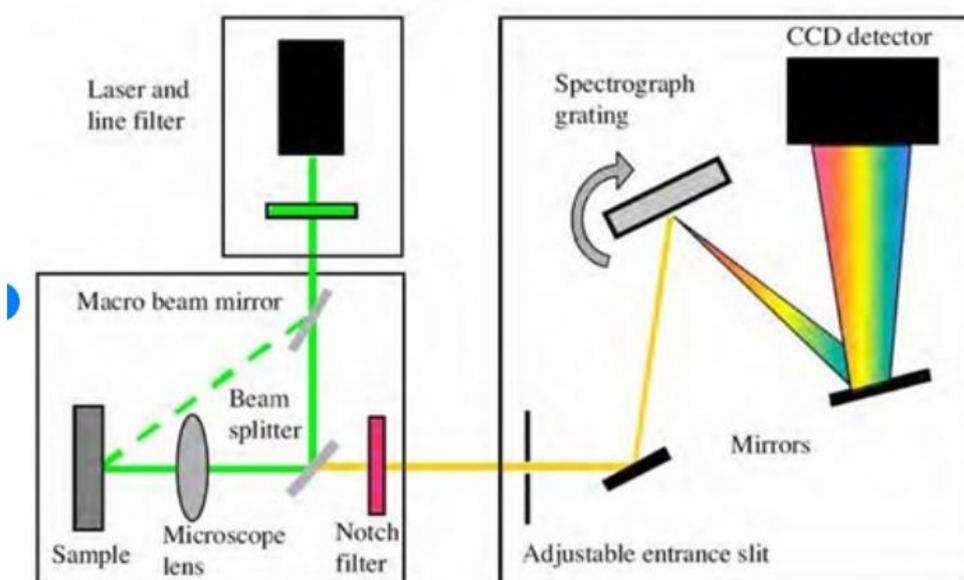


Figura: 2 Esquema de funcionamiento de un espectrómetro Raman

Quimiometría

La quimiometría es la disciplina científica que involucra análisis multivariado de los datos de origen químico (por ejemplo, de la química analítica), sea desde el punto de vista de exploración como del desarrollo de modelos matemáticos. Algunas de las áreas de aplicación más importantes de la quimiometría incluyen calibración, verificación y pruebas de importancia, optimización de mediciones químicas y procedimientos experimentales, y extracción de la máxima información química de datos analíticos (Gemperline, 2006). Con ayuda de la quimiometría se puede afrontar, con mayores posibilidades de éxito, las dificultades que surgen al emplear técnicas espectroscópicas para el control de calidad de las cervezas, como son las señales no resueltas, la gran cantidad de componentes químicos en las cervezas y la complejidad de los espectros Raman.

Los espectros Raman contienen información química y física de todos los componentes de la muestra; sin embargo, debido a la complejidad de los mismo, en la mayoría de los casos no es posible diferenciar la información multivariable que contienen. La quimiometría es capaz de extraer los parámetros de interés de los espectros a través de métodos matemáticos y estadísticos, relacionando las variables espectrales con las propiedades de los analitos y de las muestras, extrayendo información relevante y reduciendo la contribución de información irrelevante (Brereton, 2007).

Los principales métodos quimiométricos son:

- Pretratamientos espectrales: para mejorar la información que interesa y disminuir la influencia de la información secundaria.
- Métodos de clasificación: para modelar grupos de muestras según su espectro.
- Métodos de regresión, para relacionar el espectro con las propiedades cuantificables de las muestras.

CAPÍTULO I

METODOLOGÍA

En el presente capítulo se describirá cada uno de los procesos realizados para el desarrollo del modelo, empezando por la construcción de la base de datos, seguido por el curado de la base de datos, análisis exploratorio y finalmente obtención de los modelos supervisados.

1.1. Construcción de base de datos

Para la elaboración de la base de datos se utilizó la información proporcionada por el grupo de investigación Beer DeCoded de Suiza disponible en su página web <https://hack.opendata.ch/project/71>. Aquí encontramos información de diversos parámetros fisicoquímicos de las cervezas. En esta página existen 118 muestras de cervezas de diferentes partes del mundo, con sus respectivos espectros necesarios para el desarrollo del modelo.

1.2. Curado de la base de datos

El pretratamiento de la información es la modificación (estandarización) de los datos antes de construir un modelo matemático o antes de aplicar alguna técnica de análisis quimiométrico. La finalidad es hacer lineal la respuesta de las variables y eliminar las fuentes extrañas de variación que no son de interés en el análisis (Hendriks, Cruz-Juarez, Bont, & Hall, 2005). Los pretratamientos espectrales utilizados en esta investigación son la obtención de la primera y segunda derivada, el escalado y el centrado.

Para obtener las derivadas de los espectros de las cervezas se utilizó el programa OriginPro 8.0. Esta operación permite aumentar las diferencias entre las bandas anchas solapadas de los espectros; también se utilizan para corregir los efectos de línea base. La primera derivada elimina los desplazamientos de línea base paralelos al eje de las abscisas, mientras que la segunda derivada elimina los términos que varían linealmente con la longitud de onda (Norris & Williams, 1984). El escalado de los datos se realizó utilizando el programa de MATLAB (R2016a), este pretratamiento ayuda a evitar la presencia de valores extremos en la escala de los datos en algunas muestras de origen natural y se prefiere su uso cuando se trata de aplicaciones cuantitativas (Fan, Cheng, Ye, Lin, & Qian, 2006).

1.3. Análisis exploratorio

Una vez realizado el curado, los datos se ingresaron al programa de MATLAB (R2016a) para realizar el análisis de conglomerados y el análisis de componentes principales mediante el PCA toolbox (Ballabio, 2015).

1.3.1. Análisis de conglomerados

El objetivo del análisis de conglomerados es detectar las similitudes entre las muestras, con las cuales se busca identificar grupos consistentes en los datos. Las similitudes entre las muestras se estiman por medio de distancias: muestras similares se caracterizan por pequeñas distancias y muestras distintas por grandes distancias. La mejor interpretación gráfica del análisis de agrupamientos se representa mediante un dendrograma, el cual ilustra la información en forma gráfica de árbol. La división de las muestras se obtiene cortando el dendrograma en el nivel deseado de similitud (Ballabio, 2015).

Los métodos de agrupamiento jerárquicos usan las medidas de distancia para cuantificar las similitudes entre grupos de muestras. Los métodos jerárquicos aglomerativos requieren que se defina una medida de uso para el cálculo de la matriz de distancias y posterior definición de la correspondiente matriz de similitudes. El algoritmo inicia con un número de grupos igual al número de observaciones. Posteriormente, se identifican los dos objetos más similares y son unidos en un único nuevo grupo a un cierto nivel de similitud. Seguidamente, se recalculan las similitudes de este nuevo grupo con respecto a los objetos restantes y se une al objeto con mayor similitud. El proceso continúa hasta que se genera un único grupo que contiene a todas las observaciones (Frank & Todeschini, *The data analysis handbook*, 1994).

Los métodos de agrupamiento se pueden utilizar también para rellenar datos faltantes de forma simple, reemplazando el valor del dato faltante con el del centroide para las variables en consideración.

1.3.2. Análisis de componentes principales

El análisis de componentes principales (PCA) (Bro & Smilde, 2014) es el método quimiométrico mayormente usado para mostrar la estructura multivariable de los datos. Esta técnica fue descrita por primera vez por Karl Pearson y desarrollada formalmente por Harold Hotelling. En esta técnica las variables que describen los datos se transforman en nuevas variables denominadas Componentes Principales (PCs) que resultan de combinaciones lineales de las variables originales y su principal característica es de ser ortogonales entre sí. La primera componente tiene la máxima varianza, la segunda componente tiene la segunda máxima varianza y así sucesivamente. De esta forma es posible tener un número de componentes reducido en comparación al número de variables originales (reduce la dimensión de los datos). El número de componentes significativas se puede elegir en base a la varianza explicada por cada componente. Mediante esta técnica es posible (Rasmus & Ange k, 2014):

- Evaluar las correlaciones entre las variables y su relevancia
- Ver objetos (identificación de valores atípicos, clases, etc.)
- Resumir la descripción de los datos (eliminación de ruido)
- Reducir el tamaño de los datos
- Buscar propiedades principales

- Definir un modelo de representación de datos en un espacio ortogonal.

1.4. Modelos supervisados

Para está estudio se utilizó el método de clasificación k NN combinado con los algoritmos genéticos.

1.4.1. Modelos de clasificación (k NN)

En Matlab se utilizó el método de los k -vecinos más cercanos (k NN) (Cover & Hart, 1967), el cual es un método de clasificación no lineal y no paramétrico, es decir, que no considera el conocimiento *a priori* de las distribuciones estadísticas que siguen las variables. Este método clasifica en función de analogías, es decir, un objeto es clasificado según las clases a las que pertenecen la mayoría de los k objetos más cercanos en el espacio de los datos. Para este propósito, k NN calcula y analiza la matriz de distancias entre los objetos, para la cual normalmente se usa la distancia euclidiana y se selecciona un número k de vecinos cercanos. El valor óptimo de k se obtiene por validación cruzada, de tal forma que genere la menor tasa de aciertos en clasificación (NER: Non Error Rate). Los resultados se expresan en NER, sensibilidad (S_n) y especificidad (S_p).

La especificidad se refiere a la pureza de la clase es decir la capacidad del modelo para aislar los espectros de una clase de las otras; la sensibilidad es la capacidad de representar los espectros en cada clase, mientras que la precisión es la habilidad que tiene un modelo de clasificación de no incluir en una clase muestras de las demás clases.

1.4.2. Selección de variables

1.4.2.1. Algoritmos genéticos

La selección de las variables para la construcción del modelo se realizó mediante los algoritmos genéticos incorporados en el GA toolbox MATLAB (Leardi, 2009; Leardi & Gonzalez, 1998). Esta técnica se basa en la teoría de la evolución de Darwin, es decir se seleccionan individuos al azar de la población inicial para ser padres y producir los hijos para la próxima generación de tal forma que, durante generaciones sucesivas, la población “evoluciona” hacia una solución óptima. Los algoritmos genéticos se aplican para resolver una variedad de problemas de optimización que no son adecuados para los algoritmos de optimización estándar, en los que la función objetivo es discontinua, no diferenciable, estocástica o altamente no lineal. Los algoritmos genéticos utilizan dos tipos principales de reglas en cada paso para crear la siguiente generación a partir de la población actual (proceso evolutivo):

- Reproducción (crossover); combinan a dos padres para formar hijos para la próxima generación pueden ser codificados por una secuencia de 0s y 1s.
- Mutación; aplican cambios aleatorios a los padres individuales para formar hijos.

Cuando los algoritmos genéticos se aplican a bases de datos que contiene grandes cantidades de variables, se pueden correr el riesgo de que los modelos se sobre ajusten, para ello, se realiza un número pequeño de corridas independientes a partir de varias poblaciones iniciales registrando la frecuencia de selección de variables. Entonces, el modelo se construye adicionando la variable más frecuente de cada población y agregando posteriormente las demás variables en función de su frecuencia de selección.

1.5. Validación del Modelo

Para la validación del modelo de clasificación KNN se realizó una validación interna o cruzada y una validación externa:

1.5.1. Validación cruzada

La validación cruzada consiste en definir k -grupos a través de una secuencia lógica que permite la partición del grupo de calibración, es decir, ayuda a dividir dicho grupo en k -grupos de validación, los que se excluyen uno a la vez, recalibrar el modelo y posteriormente realizar la predicción de las moléculas excluidas. Un enfoque con ventanas venecianas, donde cada objeto del grupo de validación se selecciona a partir del primer objeto del grupo de calibración y los subsiguientes cada k -ésimo objeto (Stuart, 2004).

1.5.2. Validación externa

Para validar el modelo se construyó un conjunto de calibración y predicción, esta división se debe hacer de tal forma que se mantenga una relación de balance entre la numerosidad de las clases.

Grupo de calibración: conjunto de datos utilizados para ajustar el modelo. Estos datos se utilizan para entrenar el modelo, es decir el modelo ve y aprende de los mismos.

Grupo de predicción: se utiliza para evaluar la capacidad predictiva de un modelo: Este grupo no debe ser considerado durante la calibración del modelo.

La partición se realizó de forma aleatoria y proporcional a la numerosidad de las clases con un algoritmo programado en MATLAB (8.0).

CAPITULO II

RESULTADOS

2.1. Generación de la base de datos

El conjunto de datos iniciales está constituido de 118 muestras de cerveza procedentes de 20 países, dentro de los cuales constan los siguientes: Estados Unidos, México, Canadá, Bélgica, Francia, Suiza, Japón, Australia, Argentina, Jamaica, Italia, España, Reino Unido, Alemania, Irlanda, Tailandia, China, Vietnam, como se muestra en el Anexo 1. Así mismo se reflejan los grados de alcohol (ABV), amargor (IBU), color (SMR) y calorías de algunas cervezas. Se identificó tres tipos de cerveza de las cuales 68 son tipo Ale, 13 son lager, 8 son a base de trigo y 29 cervezas se encuentran sin identificar.

Cada cerveza consta con sus respectivos espectros Raman, los cuales se utilizaron para la obtención del modelo. Los datos fueron sometidos a los pretratamientos de la primera y segunda derivada, escalado, con la finalidad de obtener espectros que puedan ser fácilmente digitalizados y así obtener modelos ajustables que permitan identificar los parámetros fisicoquímicos de las cervezas.

2.2. Análisis de componentes principales

El PCA se empleó para visualizar las muestras de las cervezas a partir de los espectros Raman. En la Figura: 3, se puede observar el gráfico de puntuaciones para la primera y segunda componente principal (PC1 con 58,89% de varianza y PC2 con un 20,51%), donde la clase 1 está compuesta por cervezas de baja fermentación, mientras que las clases 2,3,4 son de fermentación alta con grado alcohólico alto, medio y bajo respectivamente. Este análisis permitió observar los cuatro conglomerados de cervezas, esto proporciona una indicación cualitativa de que aparentemente será posible construir un modelo de clasificación.

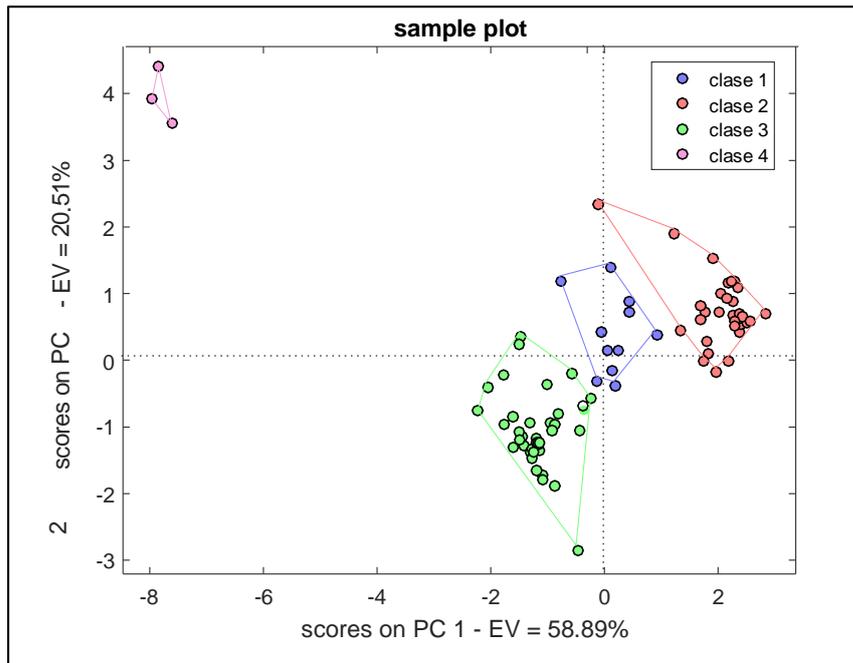


Figura: 3 Resultados del análisis del componente principal de los espectros Raman.

2.3. Análisis de conglomerados

El análisis de conglomerados se realizó con la finalidad de rellenar los datos faltantes al remplazar por el centroide de la variable utilizada. En la Figura: 4 se representa el dendrograma en el que se utilizó euclidiana como distancia, donde se puede observar que, para un valor de corte mínimo de 55, todas las muestras se clasificaron en cuatro grupos principales, correspondientes a muestras con características comunes como grado de alcohol, tipo de fermentación (alta y baja), y grados IBU. El primer grupo está compuesto por cervezas de fermentación baja con grado alcohólico de 4.88% y 24,4 IBU, el segundo grupo es de fermentación alta con grado alcohólico de 5,79 y 30.35 IBU, el tercer grupo es de fermentación alta con 5,59% de grado alcohólico y 31 IBU y el cuarto grupo es de fermentación alta con grado alcohólico alto de 7% y 27,8 IBU.

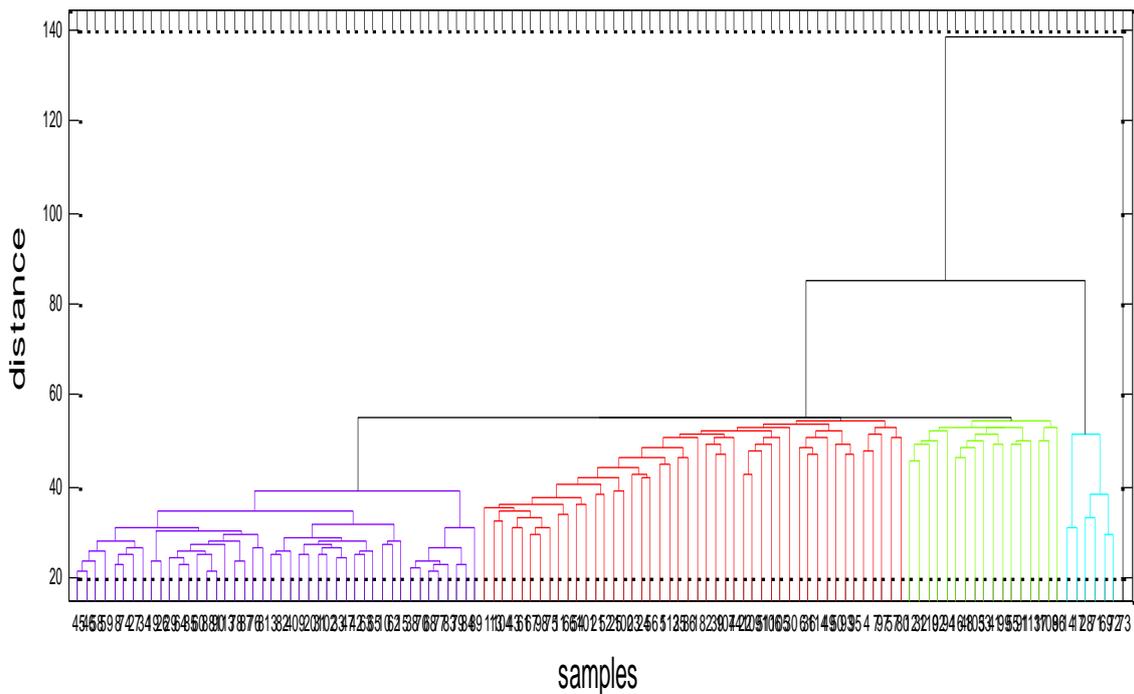


Figura: 4 Dendrograma de similitud de los espectros de las cervezas; se notan los cuatro grupos como en el análisis PCA.

La gráfica de los centroides de las tres longitudes de ondas, de sus respectivas derivadas segundas y los valores respectivos de ABV y IBU se presenta en la Figura: 5.

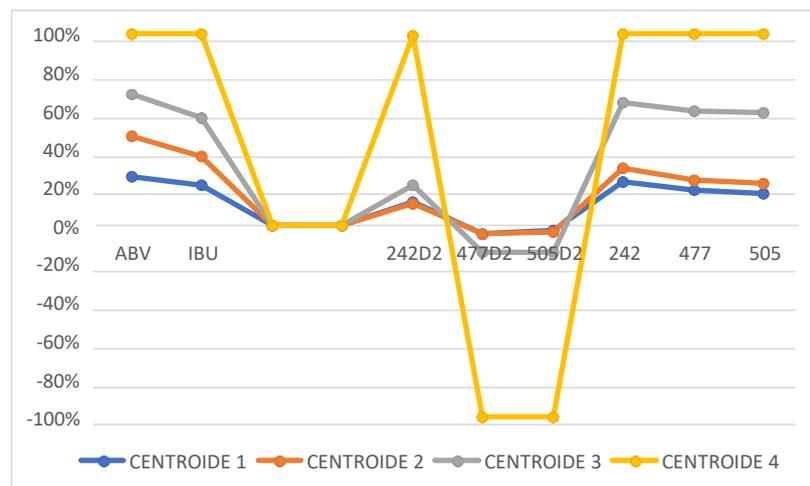


Figura 5: Centroides de las longitudes de onda seleccionadas, de sus derivadas segunda y de las características de las cervezas ABV y IBU.

Se nota inmediatamente que las cervezas de la clase 4 (ABV y IBU altos) tienen los valores más altos de las variables seleccionadas, mientras que sus valores de derivada segunda son las más bajas.

Las cervezas de la clase 3 (ABV y IBU medios altos), 2 (ABV y IBU medios) y 1 (ABV y IBU bajos) están reflejados en los valores de las 3 longitudes de onda seleccionadas. Los valores de las

derivadas segundas de las longitudes de onda 477D2 y 505D2 presentan una inversión con respecto al comportamiento de sus valores originales.

2.3. Modelo de clasificación

Para el desarrollo del modelo de clasificación *k*NN se usaron los datos autoescalados y la distancia euclidiana como tipo de medida. El modelo se determinó utilizando 83 espectros sea de longitudes de onda que, de sus respectivas derivadas segundas, mientras que los remanentes 35 se incluyeron en el grupo de predicción para la validación externa del modelo. Los espectros del conjunto de validación se extrajeron respetando la numerosidad de las clases.

Como primer paso se procedió a extraer de las 2048 longitudes de onda, mediante técnicas de algoritmos genéticos las variables más importantes para producir la mejor clasificación.

Las variables importantes obtenidas han sido: las longitudes de onda de: 242, 477, 505 de diferencia positiva (líneas tipo Stokes) con respecto a la luz Laser de excitación (785 nm) como se muestra en el Anexo 2.

La calidad del modelo final fue evaluada considerando la tasa de aciertos (NER), especificidad (Sp), sensibilidad (Sn) y precisión de las clases, tal como se indica en la Tabla 2.

Tabla 2. Parámetros de calidad del modelo obtenidos con el método de clasificación *k*NN

	NER	Sn				Sp				Prec.			
		Clase 1	Clase 2	Clase 3	Clase 4	Clase 1	Clase 2	Clase 3	Clase 4	Clase 1	Clase 2	Clase 3	Clase 4
Calibración	0.98	0.91	1	1	1	1	1	0.98	1.00	1	1	0.97	1
Validación	0.98	0.91	1	1	1	1	1	0.98	1.00	1	1	0.97	1
Predicción	0.98	1	1	0.94	1	0.8	1	1	1	0.8	1	1	1

La Tabla 2. indica que la validación interna del modelo muestra buena estabilidad del mismo con un ($NER_{cv}=0.98$), de igual manera: el modelo presenta buena capacidad predictiva ($NER_{pred}=0.98$), es decir, presenta 98% de probabilidad de predecir apropiadamente la clase a la cual pertenecen los nuevos espectros y por lo tanto establecer a cuáles clases se pueden asignar las cervezas.

CAPITULO III

DISCUSIONES

Los espectros Raman contienen información de los parámetros fisicoquímicos de las cervezas, sin embargo, debido a la complejidad de los mismos, en la mayoría de los casos no es posible diferenciar la información multivariable que contienen. La quimiometría es capaz de extraer los parámetros de interés de los espectros a través de métodos matemáticos y estadísticos, relacionando las variables espectrales con los parámetros fisicoquímicos, extrayendo información relevante y reduciendo la contribución de variables irrelevantes.

El curado del conjunto de datos inicial permitió minimizar las contribuciones no deseadas, presentes en los espectros, que disminuyen la reproducibilidad y pueden provocar no linealidades u otros efectos que darían lugar a estimaciones menos sólidas (Da-Wen, 2009).

El análisis de componentes principales proporciona un panorama general para explicar las pautas de las muestras estudiadas, sin embargo, para la asignación de nuevas muestras a una clase puede ser una tarea difícil mediante este método, ya que no calcula una regla que defina fronteras o regiones para diferenciar entre grupos obtenidos a partir de las muestras en estudio. Por lo tanto, se emplearon métodos supervisados con el fin de lograr una clasificación más objetiva.

La selección de variables mediante la aplicación de algoritmos genéticos permitió seleccionar tres variables que describen las características de las cervezas en el presente estudio. Esto evidencia que los algoritmos genéticos son una herramienta potente para el desarrollo de modelos supervisados.

El análisis de los centroides permite observar el comportamiento de las variables y su aporte a las clases. Esto se puede observar en la figura 5, donde a la clase 4 se asignan aquellas cervezas cuyos valores de grados alcohólicos y amargor es alto, para la clase 3 aquellos con grado alcohólico y amargor son medios altos, en la clase 2 cuyo grado alcohólico y amargor son medios y clase 1 ABV y IBU bajos. Las cervezas cuyo grado alcohólico es alto (mayor al 5% ABV) y un contenido alto de amargor pertenecen a las cervezas tipo lager; siendo estas en este estudio las de clase 4,3 y 2, mientras que la clase 1 pertenecen a las cervezas tipo ale con un grado alcohólico bajo (4%ABV).

Los resultados presentados en la tabla 2 indican que el modelo que se obtuvo mediante algoritmos genéticos (GAs) acoplados con el método de clasificación de los k -vecinos más cercanos (k NN) son satisfactorios, por lo que es factible utilizar el modelo para predecir la pertenencia de nuevas cervezas a una de las 4 clases establecidas en este estudio. Por ejemplo, obtener el grado alcohólico o el nivel de amargor de una cerveza nueva a partir del espectro Raman.

CONCLUSIONES

En este trabajo de investigación se ha desarrollado un modelo de clasificación con la finalidad de identificar parámetros fisicoquímicos de las cervezas a partir de una base de datos de 118 espectros. El método de clasificación de los k -vecinos más cercanos (k NN) acoplados a los algoritmos genéticos (GAs) permitieron obtener un modelo con buena capacidad de ajuste, validación y predicción, con un NER de 0,98. En consecuencia, el modelo permite predecir la clase a la cual pertenecen los nuevos espectros y, por lo tanto, las nuevas cervezas examinadas.

BIBLIOGRAFÍA

- Bagheri, S., Omidikia, N., & Kompany-Zareh, M. (2013). Unsupervised Selection of Informative Descriptors in QSAR Study of Anti-HIV Activities of HEPT Derivatives. *Chemometrics and Intelligent Laboratory Systems*, 128, 135-143.
- Ballabio, D. (2015). A MATLAB toolbox for Principal Component Analysis and. *Chemom. Intell. Lab. Syst*, 149, 1-9.
- Brereton, R. (2007). *Applied Chemometrics for Scientist*. Great Britain: John Wiley & Sons.
- Bro, R., & Smilde, A. (2014). Principal component analysis. *The Royal Society of Chemistry*.
- Consonni, V., Ballabio, D., Manganaro, A. M., & Todeschini, R. (2009). Canonical Measure of Correlation (CMC) and Canonical Measure of Distance (CMD) between Sets of Data: Part 2. Variable Reduction. *Analytica Chimica Acta*, 52-59.
- Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification . *IEEE transactions on information* , 13, 21-27.
- Da-Wen, S. (2009). *Infrared Spectroscopy for Food quality Analysis and Control*. Dublin: Elsevier Inc.
- De Clerck, J. (1957). *A Textbook of Brewing* (Vol. Two). London: Chapman and Hall.
- De Clerck, J. (1957). *A Textbook of Brewing*. (Vol. One). London: Chapman and Hall.
- Fan, X., Cheng, Y., Ye, Z., Lin, R., & Qian, Z. (2006). Multiple chromatographic fingerprinting and its application to the quality control of herbal medicines. *Elsevier* .
- Frank, I., & Todeschini, R. (1994). *The data analysis handbook*. Elsevier, Amsterdam.
- Frank, I., & Todeschini, R. (2006). The Data Analysis Handbook. *Elsevier Science*, 334-350.
- Freixes, S., & Punsola, A. (2014). *El mundo de la cerveza artesanal*. Grupo Anaya Comercial.
- Gemperline, P. (2006). *Practical Guide to Chemometrics*. Boca Raton: CRC Press.
- Hastie, T., Tibshirani, R., & Friedman, J. (2011). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (Second ed.). Germany: Springer-Verlag.
- Hendriks, M., Cruz-Juarez, L., Bont, D., & Hall, R. (2005). Preprocessing and exploratory analysis of chromatographic profiles of plant extracts. *Elsevier*, 53-64.
- IICA. (1999). *Industria de la Cerveza Guia para la Aplicacion del Sistema de Analisis*. IICA.
- Knudsen, F. (1977). *El Cerveceros en la práctica* . Madison: Asociación de Maestros cerveceros de las Américas.
- Leardi, R. (2009). Genetic Algorithms Comprehensive Chemometrics. 631-653.
- Leardi, R., & Gonzalez, A. (1998). Genetic algorithms applied to feature selection in PLS regression: how and when to use them. *Chemometrics and Intelligent Laboratory Systems*, 195-207.
- Martens, H., & Naes, T. (1989). *Multivariate Calibration*. Great Britain: Wiley & Sons.

- Medina, J., Sato, R., & Frausto, C. (2003). La Espectroscopia Raman molecular y su aplicación ambiental. *Conciencia Tecnológica*.
- Norris, K. H., & Williams, P. (1984). Optimization of Mathematical Treatments of Raw Near-Infrared Signal in the Measurement of Protein in Hard Red Spring Wheat. I. Influence of Particle Size. *Cereal Chemistry*, 61,158-165.
- Pelletier, M. (1999). *Analytical applications of Raman spectroscopy*. Blackwell Science.
- Pilla, S., & Vinc, G. (2012). *Cervezas de todo el mundo* . México: De Vecchi Ediciones, S.A} .
- Rasmus, B., & Ange k, S. (2014). Principal component analysis.
- Rencher, A., & Schaalje, G. (2008). *Linear models in statistics*. USA: John Wiley & Sons, Inc.
- SKoog, D., Holler, J., & Crouch, S. (2006). *Principles of instrumental analisis* (6 ed.). Cengage Learning.
- Stuart, B. (2004). *Infrared Spectroscopy: Fundamentals and Applications*. England: Jhon Wiley and Sons Ltda.
- Valderrama, J. (1999). La espectroscopia Raman y la espectroscopia de Resonancia Raman. *Información Tecnológica* .
- Willard, H. H. (1988). *Instrumental methods of analysis* (7th ed.). United States.: CBS Publisher & Distributors.
- Wold, S., Sjostrom, M., & Eriksson. (2001). PLS-regression: a basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems*.

ANEXOS

Anexo 1. Muestras de 118 cervezas procedentes de 20 países con sus respectivos parámetros fisicoquímicos.

Raman Spectra	Brewery	Beer-name	Beer country	Style, ratebeer	ABV	IBU	SRM	Calories, ratebeer
Analyse002	Adnams	Adnams Broadside	UK	English Strong Ale	6.3	43	na	189
Analyse003	Brasserie Du Jorat	Ambrée des Brigands du Jorat	Switzerland	Smoked	6	14	10	180
Analyse004	Anchor	Anchor Porter	USA	Porter	5.6	27	20	168
Analyse005	Anchor	Anchor Steam	USA	California Common	4.9	37	6	147
Analyse006	Anderson Valley	Heelch O'hops Double IPA	USA	IPA	8.7	100	na	261
Analyse007	Asahi	Asahi super dry	Japan	Lager	5	12	3	150
Analyse008	BFM	La Salamandre	Switzerland	Weissbier	6.3	15	na	189
Analyse009	BFM	La Torpille	Switzerland	Belgian Strong Ale	7.5	40	na	225
Analyse010	Brasserie Du Jorat	Blanche du Jorat	Switzerland	Weissbier	5.5	15	na	180
Analyse011	Brooklyn Brewery	Brooklyn Lager	USA	Amber Lager	5.2	33	na	156
Analyse012	Black Sheep	Monty Python's Holy Grail	UK	Premium Bitter	4.7	na	na	141
Analyse013	Thai Beverages	Chang	Thailand	Pale Lager	5	na	na	150
Analyse014	Chimay	Chimay Red Cap	Belgium	Abbey Dubbel	7	20	14	210
Analyse015	Innis & Gunn	Innis & Gunn Rum Finish	UK	English Strong Ale	7.4	21	28	222
Analyse016	Brasserie Castelain	Ch'Ti Blonde	France	Bière de Garde	6.4	24	12	192
Analyse017	Brewdog	Cocoa Psycho	UK	Imperial Stout	10	85	na	300
Analyse018	Coopers Brewery	Coopers Original Pale Ale	Australia	Golden Ale	4.5	22	5	135
Analyse019	Brouwerij Huyghe	Delirium Red	Belgium	Fruit Beer	8	10	na	240
Analyse020	Brouwerij Huyghe	Delirium Tremens	Belgium	Belgian Strong Ale	8.5	50	16	255
Analyse021	Verhaeghe	Duchesse de Bourgogne	Belgium	Sour Red/Brown	6.2	6	20	186

Analyse022	Duvel Moortgat	Duvel	Belgium	Belgian Strong Ale	8.5	50	16	255
Analyse024	Unibroue (Sapporo)	La Fin du Monde	Canada	Abbey Tripel	9	19	5.5	270
Analyse025	Flying Dog Brewery	Flying Dog Pale Ale	USA	American Pale Ale	5.5	35	6	165
Analyse026	Flying Dog Brewery	Flying Dog Snake Dog IPA	USA	IPA	7.1	60	11	213
Analyse027	Founders Brewing Company	Founders All Day IPA	USA	IPA	4.7	42	na	141
Analyse028	Founders Brewing Company	Founders Centennial IPA	USA	IPA	7.2	65	11	216
Analyse029	Founders Brewing Company	Founders Dirty Bastard	USA	Scotch Ale	8.5	50	13	255
Analyse030	Frankenheim Privatbrauerei	Frankenheim Alt	Germany	Altbier	4.8	11	11	144
Analyse031	St. James's Gate (Diageo)	Guinness Spec. Export	Ireland	Foreign Stout	8	50	35	240
Analyse032	Brouwerij Van Steenberge	Gulden Draak	Belgium	Belgian Strong Ale	10.5	28	17	315
Analyse033	Brouwerij Bosteels	Tripel Karmeliet	Belgium	Abbey Tripel	8.4	30	6	252
Analyse034	Brouwerij Bosteels	Pauwel Kwak	Belgium	Ale	8.4	50	16	252
Analyse035	Liefmans (Duvel-Moortgat)	Liefmans Fruitesse	Belgium	Fruit Beer	3.8	7	na	114
Analyse036	John Martin	Martin's IPA	Belgium	IPA	6.9	36	na	207
Analyse037	Charles Wells	McEwan's Export	UK	Bitter	4.5	24	na	135
Analyse038	Moosehead Brewery	Moosehead Lager	Canada	Pale Lager	5	16	na	150
Analyse039	Nøgne Ø (Hansa Borg)	Nøgne Ø Pale Ale	Norway	American Pale Ale	6	40	na	180
Analyse040	Officina della birra	Lisbeth	Switzerland	American Pale Ale	7.5	47	na	225
Analyse041	Brasserie d'Orval	Orval	Belgium	Belgian Ale	6.2	32	na	186
Analyse042	Pietra	Pietra	Corse	Amber Lager	6	26	na	180
Analyse043	Brewdog	Punk	UK	IPA	5.6	40	7.6	168
Analyse044	Cerveceria Malteria Quilmes SAICAY	Quilmes	Argentina	Pale Lager	4.9	15	na	147

Analyse045	Desnoes and Geddes (Diageo PLC)	Red Stripe	Jamaica	Pale Lager	4.7	17	na	141
Analyse046	Brasserie Rochefort	Rochefort Trappistes 10	Belgium	Abt/Quadrupel	11.3	30	na	339
Analyse047	Sabeco	Saigon Export	Vietnam	Pale Lager	4.9	na	na	147
Analyse048	Sierra Nevada Brewing Company	Sierra Nevada Pale Ale	USA	American Pale Ale	5.6	37	na	168
Analyse049	Sierra Nevada Brewing Company	Sierra Nevada Torpedo Extra IPA	USA	IPA	7.2	65	na	216
Analyse050	Boon Rawd Brewery	Singha	Thailand	Premium Lager	5	na	na	150
Analyse051	John Martin	Timmermans Tradition Blanche	Belgium	Witbier	4.5	5	na	135
Analyse052	Trois Dames	La Fraicheur	Switzerland	Witbier	4.8	20	na	144
Analyse053	na	Tsing Tao	China	Lager	na	na	na	na
Analyse054	Brouwerij der Trappisten van Westm	Westmalle tripel	Belgium	Trappiste	9.5	39	na	285
Analyse055	Verhaeghe	Barbe Rufa	Belgium	Belgian Strong Ale	8	30	na	255
Analyse056	La cotta	La cotta bionda	Italy	Golden Ale	4.8	75	na	144
Analyse057	Birrificio Trunasse	Hoppy Couple	Italy	IPA	6	na	na	na
Analyse058	Trumer	Waldbier 2014 Wienerwald	Austria	Spice/Herb/Vegetable	8.7	15	12.7	246
Analyse059	na	Brick By Brick Bosun's Moustache	Lithuania	IPA	6	60	na	na
Analyse060	Grupo Modelo	Corona Extra	Mexico	Pale Lager	4.6	18	3	138
Analyse061	Hofbrau	Weisse Alkoholfrei	Germany	Low Alcohol	0.5	12	na	80
Analyse062	Valaisanne (Feldschlösschen)	Valaisanne Ämrich	Switzerland	Weissbier	5.5	10	5.6	165
Analyse063	Brasserie Gessienne	Blanche	France	Weissbier	5.6	na	na	na
Analyse064	Mateo & Bernabé	Mateo 21	Spain	Wheat Ale	5	13	na	150
Analyse065	Oland Brewery (AB-InBev)	Alexander Keiths India Pale Ale	Canada	IPA	5	20	na	150
Analyse066	Sleeman	Cream Ale	Canada	Cream Ale	5	na	na	150

Analyse067	Sleeman	Original draught	Canada	na	5	na	na	na
Analyse068	Mill Street	Organic Lager	Canada	na	4.2	na	na	na
Analyse069	Labatt	Budweiser	Canada	na	5	na	na	na
Analyse070	Coors Brewing	Coors Light	Canada	na	4	na	na	na
Analyse071	Wellington	S.P.A	Canada	Pale Ale	4.5	22	na	150
Analyse072	Wellington	County	Canada	Ale	5	na	na	na
Analyse073	Wellington	Arkell	Canada	Ale	4	na	na	na
Analyse074	Wellington	Imperial	Canada	Stout	8	42	na	240
Analyse075	Bitchin' Brewery	Stout	UK	Stout	6.4	na	na	na
Analyse076	Bitchin' Brewery	Coffee Stout	UK	Stout	6.4	na	na	na
Analyse077	Windsor & Eton	Magna Carta	UK	Ale	7.2	na	na	na
Analyse078	Hackuarium	Fakufaku	CH	Stout	na	na	na	na
Analyse079	Frères Papinot	Calvinus Blanche	CH	Weissbier	5.2	15	13	156
Analyse080	Frères Papinot	Calvinus Blonde	CH	Pale Lager	4.8	18	7	156
Analyse081	Les Murailles	Meynite	CH	Belgian Ale	5	na	na	150
Analyse082	Les Murailles	Sorcière	CH	na	na	na	na	na
Analyse083	Les Murailles	Pieuse	CH	na	na	na	na	na
Analyse084	Homebrew	Roter Baron	CH	na	na	na	na	na
Analyse085	Docteur Gabs	Houleuse	CH	Witbier	5	na	na	150
Analyse086	Docteur Gabs	Pepite	CH	na	na	na	na	na
Analyse087	Docteur Gabs	Tempête	CH	na	na	na	na	na
Analyse088	La Nébuleuse	Embuscade	CH	IPA	6.7	65	8.2	na

Analyse089	La Nébuleuse	Malt Capone	CH	Robust Porter	7	30	50	na
Analyse090	La Nébuleuse	Stirling	CH	California Common	5.3	38	7.3	na
Analyse091	Bieraria Tschlin	BE	CH	na	na	na	na	na
Analyse092	Mongozo	Banana	Belgium	na	na	na	na	na
Analyse093	Chimay	Tripel	Belgium	na	na	na	na	na
Analyse094	Andechs	Weissbier hell	Belgium	na	na	na	na	na
Analyse095	Warsteiner	Konig Ludwig Dunkel	Germany	na	na	na	na	na
Analyse096	Feldschlosschen	Braufrisch	CH	na	na	na	na	na
Analyse097	Brasserie DuPont	BioLegère	Belgium	na	3.5	na	na	na
Analyse102	Coudres	Pale Ale	CH	na	na	na	na	na
Analyse103	Unserbier	Amber	CH	na	na	na	na	na
Analyse104	Brasserie du Mont Blanc	Brassin d'Hiver	France	na	na	na	na	na
Analyse105	Ottakringer	Wiener Original	Austria	na	na	na	na	na
Analyse106	Homebrew	Amber Ale	CH	na	na	na	na	na
Analyse107	Homebrew	SquareBeer	CH	White IPA	na	na	na	na
Analyse110	La Mule	Browney	CH	na	na	na	na	na
Analyse111	Bartis	La Koyot	CH	na	na	na	na	na
Analyse112	Brasserie Montheysanne	La Montheysanne	CH	na	na	na	na	na
Analyse113	Brasserie du Griffon	La Fourbe	CH	na	na	na	na	na
Analyse114	Brasserie Sierrvoise	Noire	CH	na	na	na	na	na
Analyse115	Brasseurs de Vollèges	La Tourbillon	CH	na	na	na	na	na
Analyse116	Brasserie du Vieux Chemin	La Prudencia	CH	na	na	na	na	na

Analyse117	Brasserie Tardiv	Tardiv	CH	na	na	na	na	na
Analyse118	Brasserie des 5 quatre mille	Biere de Zinal	CH	Blanche au seigle	na	na	na	na
Analyse119	Brasserie d'Ayent	Celsius Folamour	CH	na	na	na	na	na
Analyse120	Coudres	Blonde	CH	na	na	na	na	na
Analyse121	Leffe	Vieille Cuvée	Belgium	na	na	na	na	na
Analyse122	La Nébuleuse	ChichaBeer experimental	CH	na	na	na	na	na
Analyse123	Muller Brau	Pale Ale	CH	na	na	na	na	na
Analyse124	Docteur Gabs	Ipanema	CH	na	na	na	na	na
Analyse125	Brewdog	5AM Saints	UK	Amber Ale	5	35	na	150
Analyse126	Boxer Old	CH	na	na	5.2	na	na	na

Anexo 2. Variables seleccionadas por lo algoritmos genéticos

242	477	505
0.1895244	-0.0219518	-0.0413785
0.29861738	-0.03509252	-0.04031291
-0.09249936	-0.02971958	-0.04264878
-0.08744531	-0.04450223	-0.03685237
0.00958174	-0.01592953	0.00961204
0.31354366	-0.04186126	-0.0427444
-0.01873301	-0.00316191	0.03538557
0.36917072	-0.04429676	-0.04115275
-0.09244272	-0.03979735	-0.04400919
0.02688872	0.00419812	-0.013709
-0.11606648	-0.04439562	-0.04126417
0.24226267	-0.03453783	-0.03068381
0.01062421	-0.03695812	0.02209432
0.27727488	-0.03058802	-0.04241866
0.17074029	-0.04235537	-0.05111814
-0.02145032	-0.02893674	0.00443687
-0.10164025	-0.03501509	-0.03706045
-0.01694056	-0.04010037	0.00882276
0.05160842	-0.01194466	-0.06735051
-0.01054438	-0.03647159	-0.01499069
0.04373622	-0.00149257	-0.04942545
0.02370291	0.00795076	-0.02867766
-0.0158884	-0.04414955	-0.0021698
-0.11434651	-0.03658827	-0.02973306
-0.11163087	-0.03136744	-0.04631374
0.16179397	-0.04729072	-0.04332197
0.27345751	-0.02765893	-0.04728229
-0.06838302	-0.03633829	-0.03633335
0.22508501	-0.02808701	-0.04964429
0.18207464	-0.04803465	-0.05200823
-0.10431863	-0.02917452	-0.03947695
-0.08464442	-0.03363219	-0.02602608
0.10733813	-0.05164557	-0.05413353
-0.07051217	-0.03282401	-0.03508448
-0.10119606	-0.03495502	-0.03262675
-0.12337074	-0.03509766	-0.03823655
0.01722641	0.00591109	-0.02002661
-0.12308286	-0.03978871	-0.0413324
-0.08110383	-0.03064702	-0.03782027
0.02564038	0.00899457	-0.0070327
-0.11752773	-0.029991	-0.03848865
-0.55463015	-0.03041403	-0.04128545
-0.02465589	0.02511202	-0.00085932
-0.00957569	0.032891	-0.04088751