



**UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL
FACULTAD DE ECONOMÍA Y EMPRESA
CARRERA DE NEGOCIOS INTERNACIONALES**

TITULO:

Aplicación de algoritmo de Machine Learning para clasificar la calidad de un cacao fino aroma de exportación para la exportadora Cocholato S.A.

AUTORES:

Abad Garzón Alejandra Margarita

Hong Luo Huaxing Angela

**Trabajo de Titulación previo a la obtención de título de
LICENCIADO EN NEGOCIOS INTERNACIONALES**

TUTOR:

Ing. Carrera Buri, Félix Miguel, Mgs.

Guayaquil, Ecuador

15 de febrero del 2024



**UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL
FACULTAD DE ECONOMÍA Y EMPRESA
CARRERA DE NEGOCIOS INTERNACIONALES**

CERTIFICACIÓN

Certificamos que el trabajo de integración curricular fue realizado en su totalidad por **Alejandra Margarita Abad Garzón** y **Huaxing Angela Hong Luo**, como requerimiento para la obtención del título de Licenciado en Negocios Internacionales.

TUTOR

f. _____

Ing. Carrera Buri, Félix Miguel, Mgs.

DIRECTORA DE LA CARRERA

f. _____

Ing. Hurtado Cevallos, Gabriela Elizabeth, Mgs.

Guayaquil, 15 de febrero del 2024



UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL

FACULTAD DE ECONOMÍA Y EMPRESA
CARRERA DE NEGOCIOS INTERNACIONALES

DECLARACIÓN DE RESPONSABILIDAD

Nosotras, **Abad Garzón, Alejandra Margarita;**
Hong Luo, Huaxing Angela

DECLARAMOS QUE:

El trabajo de Integración Curricular, **Aplicación de algoritmo de Machine Learning para clasificar la calidad de un cacao fino aroma de exportación para la exportadora Cocholato S.A.**, previo a la obtención del título de **Licenciado en Negocios Internacionales**, ha sido desarrollado respetando derechos intelectuales de terceros conforme las citas que constan en el documento, cuyas fuentes se incorporan en las referencias o bibliografías. Consecuentemente este trabajo es de mi total autoría-

En esta virtud de esta declaración, me responsabilizo del contenido, veracidad y alcance del Trabajo de Integración Curricular referido.

Guayaquil, 15 de febrero del 2024

AUTORES

f. _____

Abad Garzón, Alejandra Margarita

f. _____

Hong Luo, Huaxing Angela



UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL

FACULTAD DE ECONOMÍA Y EMPRESA
CARRERA DE NEGOCIOS INTERNACIONALES

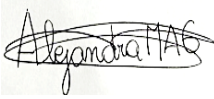
AUTORIZACIÓN

Nosotras, **Abad Garzón Alejandra Margarita;**
Hong Luo Huaxing Angela

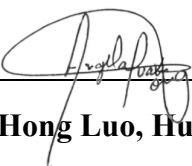
Autorizo a la Universidad Católica de Santiago de Guayaquil a la publicación en la biblioteca de la institución del Trabajo de Integración Curricular, **Aplicación de algoritmo de Machine Learning para clasificar la calidad de un cacao fino aroma de exportación para la exportadora Cocholato S.A.**, cuyo contenido, ideas y criterios son de mi exclusiva responsabilidad y total autoría.

Guayaquil, 15 de febrero del 2024

AUTORES

f.  _____

Abad Garzón, Alejandra Margarita

f.  _____
Hong Luo, Huaxing Angela



UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL

FACULTAD DE ECONOMÍA Y EMPRESA

CARRERA DE NEGOCIOS INTERNACIONALES

REPORTE COMPILATO



Tesis Final Alejandra Abad Garzoën -
Angela Huaxing (1)

1%
Textos
sospechosos

5% Similitudes (ignorado)
0% similitudes entre
comillas
< 1% entre las fuentes
mencionadas (ignorado)
1% Idioma no reconocido

Nombre del documento: Tesis Final Alejandra Abad Garzoën - Angela
Huaxing (1).docx
ID del documento: 2105aec0a5d851e3c0d0b435d95becfe86963a8b
Tamaño del documento original: 2,24 MB

Depositante: Felix Miguel Carrera Buri
Fecha de depósito: 26/1/2024
Tipo de carga: Interface
fecha de fin de análisis: 26/1/2024

Número de palabras: 25.038
Número de caracteres: 171.657

Ubicación de las similitudes en el documento:



Fuentes principales detectadas

N°	Descripciones	Similitudes	Ubicaciones	Datos adicionales
1	repositorio.ucsg.edu.ec http://repositorio.ucsg.edu.ec/bitstream/3317/9365/3/T-UCSG-PRE-MED-634.pdf.txt 3 fuentes similares	< 1%		Palabras idénticas: < 1% (171 palabras)
2	repositorio.ucsg.edu.ec http://repositorio.ucsg.edu.ec/bitstream/3317/16370/1/T-UCSG-PRE-JUR-DER-644.pdf	< 1%		Palabras idénticas: < 1% (151 palabras)
3	www.forbes.com.ec Inteligencia Artificial y Desarrollo Sostenible en Ecuador: Un... https://www.forbes.com.ec/columnistas/inteligencia-artificial-desarrollo-sostenible-ecuador-una-co... 4 fuentes similares	< 1%		Palabras idénticas: < 1% (91 palabras)
4	www.scielo.org.co LA CADENA DE VALOR DEL CACAO EN PERÚ Y SU OPORTUNIDA... http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S0120-63462015000100006 2 fuentes similares	< 1%		Palabras idénticas: < 1% (84 palabras)

TUTOR

f. _____

Ing. Carrera Buri, Félix Miguel, Mgs.

AGRADECIMIENTO

Me gustaría agradecer a la Universidad Católica Santiago de Guayaquil por abrirme las puertas y brindarme la oportunidad de avanzar en mi carrera profesional, por consiguiente, a todos los catedráticos que han sido guías constantes y mentores a lo largo de mi carrera académica, especialmente mi más profundo agradecimiento a mi director de tesis, Ing. Carrera Buri, Félix Miguel Mgs. , su experiencia, comprensión y paciencia contribuyeron a mi vivencia en el complejo y gratificante camino de la investigación, en el campo de la inteligencia artificial.

Este honor no solo me pertenece a mí, yo no estaría aquí si no fuera por personas muy importantes en mi vida, mi otra familia , yo he sido alentada, apoyada y tolerada, no solo por mi compañera de tesis, Angela Hong, mi primera amiga de la universidad, sino por los mejores amigos que una persona pudiera tener, Milton Camacho, Frixon Hurtado, Gabriela Cárdenas, Carlos Cruz, Alexander Cruz, Sabrina Sepa, María José Riofrio, Josselyn Bravo, Daniela García y Alison Garcés, les pido disculpas no siempre he sido la amiga que merecen pero quiero que sepan que a mi manera siempre los he querido, gracias por siempre ser y estar.

- *Alejandra Margarita Abad Garzón*

AGRADECIMIENTO

En primer lugar, deseo agradecer mi gratitud eterna hacia mis padres, quienes han sido una presencia constante y apoyo genuino a lo largo de este viaje. Su amor incondicional, paciencia y respaldo continuo han sido fundamentales en cada paso que he dado. Los sacrificios que han hecho y sus esfuerzos incansables no solo son reconocidos, sino que han sido mi inspiración diaria, guiándome con sabiduría y dándome ánimos incluso en los momentos más difíciles y desafiantes.

A mi prima, Jenny, quiero expresar mi profundo agradecimiento por ser una fuente constante de aliento a lo largo de toda mi carrera. Su constante respaldo ha sido una luz guía, y estoy infinitamente agradecida por tenerla como un pilar en mi vida. Agradezco enormemente tu apoyo inquebrantable y motivación constante.

A mis mejores amigas, Camila y Nicole, les quiero dedicar un agradecimiento especial y sincero. Cada paso en este camino académico ha sido más llevadero gracias a su amistad. Las palabras alentadoras que me han brindado han sido un pilar en los días más difíciles, recordándome que las adversidades son temporales. Agradezco cada viaje compartido, cada consejo y cada risa compartida que, de manera invaluable, han sido mi refugio a lo largo de este recorrido.

A mi compañera de tesis, Alejandra Abad, quiero expresar mi agradecimiento. No solo ha sido una colaboradora indispensable en nuestro proyecto académico, sino que, desde los primeros días de la carrera, se ha convertido mi primera y más cercana amiga de la universidad. Su dedicación y conocimientos han enriquecido nuestro trabajo. Estoy emocionado por seguir compartiendo éxitos y desafíos juntos en el futuro.

Finalmente, deseo expresar mi sincero agradecimiento a todos los profesores que han sido guías y mentores a lo largo de mi carrera académica. Su dedicación, paciencia y sabiduría han sido fundamentales para mi crecimiento intelectual y personal. Cada lección, orientación y apoyo han contribuido significativamente mi formación. Su pasión por la enseñanza ha dejado una huella duradera en mi trayectoria académica.

- *Huaxing Angela Hong Luo*

DEDICATORIA

Gracias Jehová por llenarme de sabiduría, todos mis logros son gracias a ese don que me has dado.

Mi inspiración fundamental viene de la mujer deslumbrante de la que recibí mi nombre y mi sangre, Margarita Garzón, gracias por siempre hacerme pensar que si puedo y llenarme de amor y risas; has sido mi pilar, sin el que no pudiera tenerme en pie durante estos 20 años de vida.

A ti, Angela Hong, porque tu amistad ha traído grandiosos colores a mi vida, ha estado ahí aun en los tiempos más difíciles y soy la persona más afortunada por tener ese regalo.

- *Alejandra Margarita Abad Garzón*

DEDICATORIA

Este logro va dedicado a aquellos, cuyas huellas han dejado una marca imborrable en mi camino académico. A mis padres, a quienes debo todo, por su inquebrantable apoyo; a mi hermana mayor y hermano menor, dos pilares fundamentales que han enriquecido cada capítulo de este viaje; a mi prima confidente y compañera de travesías; y a mi fiel perrito Lucky, quien ha sido testigo silencioso de cada página escrita en esta travesía. Su amor, guía y alegría han sido una inspiración constante.

- *Huaxing Angela Hong Luo*



UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL

FACULTAD DE ECONOMÍA Y EMPRESA
CARRERA DE NEGOCIOS INTERNACIONALES

TRIBUNAL DE SUSTENTACIÓN

Gabriela Hurtado

Ing. Hurtado Cevallos, Gabriela Elizabeth, Mgs.
DIRECTORA DE CARRERA

(NOMBRES Y APELLIDOS)
COORDINADOR DEL ÁREA

(NOMBRES Y APELLIDOS)
OPONENTE



UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL

FACULTAD DE ECONOMÍA Y EMPRESA
CARRERA DE NEGOCIOS INTERNACIONALES

CALIFICACIÓN

Abad Garzón, Alejandra Margarita

Hong Luo, Huaxing Angela

Índice General

Introducción	2
Problemática.....	7
Justificación.....	11
Alcance.....	16
Objetivos	18
Objetivo General.....	18
Objetivos específicos.....	18
Preguntas de investigación	18
Capítulo II	19
Marco Teórico.....	19
Introducción al Aprendizaje Estadístico	19
Inteligencia Artificial	20
Machine Learning	21
Aprendizaje Supervisado	21
Modelo de clasificación	22
Aprendizaje No Supervisado	23
Aprendizaje por Refuerzo	25
Diferencias entre aprendizaje supervisado y no supervisado	26
Técnicas principales de clasificación.....	27
Matriz de confusión.....	27
Árbol de decisión.	27
Bosque Aleatorio.....	28
Marco Conceptual.....	33
Cacao.....	33

Compra de cacao	33
Cadena de suministro	33
Rentabilidad	33
Perfil del cliente	34
Servicio	34
Exportación	34
El mercado mundial del cacao	34
Exportación de cacao ecuatoriano	36
Descripción y principales características del cacao	38
Clasificación botánica del cacao	38
Clasificación comercial del cacao.....	40
Comercio internacional del cacao	41
Principales productores de cacao	41
Estándares de calidad del cacao	42
Calidad del grano fresco en baba	43
Marco Legal.....	45
Capítulo III	46
Metodología.....	46
Pasos para desarrollar	46
Caret.	49
GridExtra.....	49
Ggplot 2.....	49
randomForest.....	50
Tidyverse.....	50
Clúster.....	50

factoExtra.....	51
NbClust.....	51
Resultados	58
Conclusiones	81
Recomendaciones.....	82
Referencias Bibliográficas	83
Anexos.....	88

Índice de Tablas

Tabla 1 Características del grano fermentado	42
Tabla 2 Base de datos tesis	47
Tabla 3 Nivel de importancia	78
Tabla 4 Matriz de Confusión	78
Tabla 5 Predicciones a través de K-means y RandomForest.....	79

Índice de Figuras

Figura 1	Aprendizaje supervisado vs No supervisado	26
Figura 2	Matriz de confusión	27
Figura 3	Árbol de decisión	28
Figura 4	Bosque Aleatorio	29
Figura 5	36
Figura 6	Calidad del cacao	44
Figura 7	Gráfico de normalidad pesticida - días de fermentación	58
Figura 8	Gráfico de normalidad Genética - días de secado	59
Figura 9	Gráfico de normalidad MCultivo - Edad de plantación.....	59
Figura 10	Gráfico de frecuencia Pesticida	60
Figura 11	Gráfico de frecuencia cultivo	61
Figura 12	Gráfico de frecuencia Genética	61
Figura 13	Gráfico de frecuencia Días de fermentación	62
Figura 14	Gráfico de frecuencia Días de secado.....	63
Figura 15	Gráfico de frecuencia Acidez.....	63
Figura 16	Histograma días de fermentación	64
Figura 17	Histograma Días secado	65
Figura 18	Histograma Acidez.....	65
Figura 19	Gráfico evolutivo Edad de plantación vs Pesticida	66
Figura 20	Gráfico evolutivo Edad de plantación vs Cultivo.....	67
Figura 21	Gráfico evolutivo Edad de plantación vs genética	68
Figura 22	Gráfico evolutivo Días de secado vs Pesticida	69
Figura 23	Gráfico evolutivo Días de secado vs Cultivo	70
Figura 24	Gráfico evolutivo Días de secado vs Genética	71
Figura 25	Gráfico evolutivo Acidez vs Pesticida	72
Figura 26	Gráfico evolutivo Acidez vs Cultivo	73
Figura 27	Gráfico evolutivo Acidez vs genética	74
Figura 28	Gráfico de clústeres	75
Figura 29	Clúster conglomerados	76

Figura 30 Gráfico elíptico de tipo euclidiana.....76
Figura 31 Gráfico de aplicación de k-means en los clústeres77

Índice de Anexos

Anexo 1 Script Completo.....	88
Anexo 2 Análisis descriptivo a partir de la función Sumamry	94
Anexo 3 Análisis descriptivo	95
Anexo 4 Base de datos (cacao)	96
Anexo 5 Predicción RandomForest	104

Resumen

En Ecuador la producción y exportación de cacao fino o de aroma representa una gran oportunidad en mercados europeos debido a la tendencia de los consumidores y su preocupación por la sostenibilidad del medioambiente. El proyecto de investigación tiene como propósito analizar la aplicación de algoritmos de machine learning para clasificar la calidad de un cacao fino aroma de exportación para la exportadora Cocholato S.A, una empresa familiar exportadora de cacao, ubicada en el cantón Rio Verde de la provincia de Esmeraldas, la clasificación del cacao es esencial para garantizar la calidad, el sabor y la consistencia de los productos de cacao, así como para facilitar la comercialización y el cumplimiento de estándares de calidad en la industria. Además, puede tener un impacto positivo en los ingresos de los agricultores de cacao y en la sostenibilidad de la producción. Cuya metodología se basó en Restudio con el algoritmo RandomForest. Los resultados arrojados son prometedores no solo para el presente sino también para el futuro sugiere que los hallazgos del estudio tienen aplicaciones prácticas y beneficios a largo plazo para los agricultores.

Palabras claves: Inteligencia de negocios, Machine learning, cacao fino de aroma, exportación, RandomForest, Calidad.

Abstract

In Ecuador, the production and export of fine or aromatic cocoa represents a wonderful opportunity in European markets due to the trend of consumers and their concern for environmental sustainability. The research project aims to analyze the application of machine learning algorithms to classify the quality of a fine aroma cocoa for export for the exporter Cocholato S.A., a family cocoa exporting company, located in the Rio Verde canton of the province of Esmeraldas, the classification of cocoa is essential to ensure the quality, flavor and consistency of cocoa products, as well as to facilitate marketing and compliance with quality standards in the industry. In addition, it can have a positive impact on cocoa farmers' incomes and on the sustainability of production. Whose methodology was based on Restudio with the RandomForest algorithm. The results are promising not only for the present but also for the future, suggesting that the study's findings have practical applications and long-term benefits for farmers.

Key words: Business intelligence, Machine learning, fine aroma cocoa, export, RandomForest, Quality.

Resumé

En Équateur, la production et l'exportation de cacao fin ou aromatique représentent une grande opportunité sur les marchés européens en raison des tendances de consommation et de leur souci de durabilité environnementale. Le projet de recherche vise à analyser l'application d'algorithmes d'apprentissage automatique pour classer la qualité d'un cacao d'arôme fin destiné à l'exportation pour l'exportateur Cocholato S.A., une entreprise familiale d'exportation de cacao, située dans le canton Rio Verde de la province d'Esmeraldas. La classification du cacao est essentielle pour garantir la qualité, le goût et l'uniformité des produits de cacao, ainsi que pour faciliter la commercialisation et le respect des normes de qualité dans l'industrie. Elle peut également avoir un impact positif sur les revenus des producteurs de cacao et sur la durabilité de la production. Dont la méthodologie était basée sur Rstudio avec l'algorithme RandomForest. Les résultats sont prometteurs non seulement pour le présent mais aussi pour l'avenir, ce qui suggère que les conclusions de l'étude ont des applications pratiques et des avantages à long terme pour les agriculteurs.

Mots-clés : Intelligence économique, apprentissage automatique, cacao fin, exportation, RandomForest, qualité.

Introducción

La tecnología ha emergido como una herramienta fundamental en la sociedad moderna, transformando completamente nuestra vida cotidiana, entorno laboral y relaciones interpersonales. Esta influencia no se circunscribe exclusivamente al ámbito personal, sino que también ejerce un impacto significativo en diversos ámbitos empresariales. De manera que, ha transformado significativamente la naturaleza de los procesos empresariales, que va desde la automatización de los procesos de tareas manuales o repetitivas hasta implementar algoritmos capaces de analizar conjuntos de datos complejos o de gran envergadura. En este contexto, la competitividad en el mercado global del sector agrícola busca continuamente formas de mejorar los procesos de control de calidad para ser más eficientes (Pérez González, Villegas Estévez, Cabascango Jaramillo, & Soria Flores, 2023).

Por otra parte, el cacao es el fruto que se origina a partir de un árbol llamado cacaotero (*Theobroma cacao* L.) es uno de los productos agroalimentarios de origen neotropical de mayor penetración en el mercado internacional para la elaboración de chocolate. A pesar de, tener ADN latinoamericano y africano, se cultiva en regiones cálidas y húmedas en más de 50 países ubicados en cuatro continentes (África, América, Asia y Oceanía); se reconocen tres grupos principales de variedades de cacao: Criollo, Forastero y Trinitario. Cada grupo tiene características distintivas y perfiles de sabor únicos. La producción de cacao a nivel mundial ha crecido de forma constante a través de los años, principalmente en el africano. Entre los países latinoamericanos se destaca el Ecuador con un 4% y Brasil el 3% de participación en el mundo en la producción. Los olmecas fueron los primeros en cultivar el cacao en México. En el curso de los siglos, la cultura del cacao se extendió a la civilización Maya (600 A.C.) y a la Azteca (1400 A.C.). Los Mayas utilizaron la

vaina del cacao para crear una bebida, alrededor del año 600 (Alcívar Córdova, Quezada Campoverde, Barrezueta Unda, Garzón Montealegre, & Carvaja Romero, 2021)

Según la INEC (2021) en América latina, específicamente Ecuador es el principal productor y exportador de cacao alcanzado una superficie plantada y un rendimiento de almendra de almendra de 590.579 ha y de 0,62 t ha⁻¹, respectivamente. El cacao ecuatoriano nativo de la región amazónica de América del Sur tiene una larga historia que se remonta a un origen y cultivo precolombino, como los mayas y los aztecas en Mesoamérica, que lo utilizaban para hacer bebidas y ofrendas religiosas. En Ecuador, los pueblos indígenas, como los Huaorani y los Kichwa, también cultivaban cacao antes de la llegada de los europeos, especialmente los españoles, el cacao ecuatoriano se convirtió en una importante mercancía de exportación a Europa por su calidad y sabor único aunque ha experimentado altibajos a lo largo de los años: debido a enfermedades del cacao, la competencia de otras variedades y la disminución de la calidad del cacao Nacional, en las últimas décadas, el cacao ecuatoriano ha ganado reconocimiento y aprecio a nivel internacional por su calidad, sabor y por sus variedades únicas que se cultivan en diversas regiones del país, consolidándose como uno de los principales productores de cacao de alta calidad en la industria chocolatera global. Actualmente, en la Norma Técnica Ecuatoriana NTE-INEN 176, se registra que el 81% del cacao certificado corresponde a cacao fino de aroma con 266.789,23 toneladas exportadas y un 19% a cacao CCN-51 con 64.239,34 toneladas exportadas.

La destacada presencia del cacao fino aroma en mercados internacionales es evidente en los principales destinos para el grano ecuatoriano, según las investigaciones de Macancela (2020) los principales destinos son Estados Unidos, Indonesia, Malasia, Países bajos, México, Bélgica, Italia y Alemania, lo que subraya la diversificación exitosa de los mercados y la sólida demanda global por la calidad excepcional del cacao (p.89).

Ecuador ha logrado posicionarse como el tercer mayor productor y exportador de cacao a nivel mundial, según Merlyn Casanova, directora ejecutiva de la Asociación de Exportadores de Cacao (Anecacao). A pesar de la competencia liderada por Costa de marfil y Ghana, que representan el 50% de las exportaciones globales, el éxito ecuatoriano se atribuye no solo a la excelencia del cacao, sino también a iniciativas locales. Programas de sostenibilidad agrícola y capacitación han contribuido a mejorar la eficiencia y la calidad del producto, consolidando así la reputación de Ecuador como un participante confiable en la industria cacaotera internacional (Franco, 2023)

La posición de Ecuador siendo exportadora de cacao a nivel nacional se fundamenta en varias razones sólidas. La calidad excepcional del cacao ecuatoriano es evidente en el tamaño y peso de las almendras, con características superiores a las del estándar mundial establecido por Ghana. Mientras que en Ghana el peso de la almendra no supera los 1.1 gramos, en Ecuador se alcanzan pesos de hasta 1.6 gramos, lo que representa un diferencial notable en la calidad física del producto final. Este hecho, combinado con la eficiencia del proceso productivo, se traduce en que con 15 o 20 mazorcas se obtenga un kilo de cacao. Además, la apuesta de Ecuador por la calidad se refleja en el liderazgo del país como productor de cacao fino de aroma, una distinción que va más allá del estándar global (Gobierno de la República del Ecuador, 2015).

Este reconocimiento de calidad y propiedades saludables ha impulsado una creciente demanda en mercados internacionales. La sólida trayectoria de la industria cacaotera de Ecuador se refleja en sus cifras de exportación, que ascendieron a \$940 millones FOB en el año 2021, marcando un incremento del 0.5% con respecto al año anterior. Además, las exportaciones del sector al cierre de marzo de 2022 ya alcanzan los \$202 millones FOB, representando un 22% del total exportado en el año previo. Estos datos, respaldados por estadísticas del Banco Central del Ecuador, revelan una tendencia sólida y constante, evidenciando el crecimiento sostenido del sector

a pesar de los desafíos. La contribución del cacao a la economía nacional se consolida, destacando su papel vital en el panorama económico del país. El aumento en las exportaciones no solo destaca la calidad del cacao ecuatoriano, sino también la posición que ha alcanzado en el ámbito internacional, convirtiéndose en un componente esencial para el desarrollo económico del país (Corporación Financiera Nacional, 2022)

La inteligencia artificial es crucial para el éxito y supervivencia de las empresas tanto nacionales como internacionales. Según Rouhiainen (2018) señala que “la inteligencia artificial permite que las maquinas aprendan de datos utilizando algoritmos y utilicen los resultados arrojados para tomar decisiones adecuadas, de manera que se asemeje a la de una persona” (p.17).

Existe gran variedad de métodos para clasificar la calidad del cacao. El aprendizaje automático también conocido como Machine Learning es una rama de la inteligencia artificial que se centra en el estudio científico de los algoritmos y modelos estadísticos que los sistemas informáticos utilizan para realizar una tarea específica sin ser programados explícitamente. Estos algoritmos se utilizan para diversos fines como la minería de datos, el procesamiento de imágenes, el análisis predictivo, etc. La principal ventaja de utilizar el aprendizaje automático es que, una vez que un algoritmo aprende qué hacer con los datos, puede hacer su trabajo automáticamente (Rouhiainen, 2018)

La implementación de la inteligencia artificial dentro de los procesos de las empresas ha experimentado un crecimiento notable a nivel mundial, ya que ha transformado la manera en que diversas industrias abordan desafíos y mejoran sus operaciones internas y externas. Según datos estadísticos recientes, China lidera la adopción de inteligencia artificial dentro de sus compañías con un 58% y un 30% que están considerando la integración de esta. En comparación con Estados Unidos presenta una tasa de adopción más baja, con un 25% de sus empresas ya utilizando inteligencia artificial y un 43% consideran su implementación (Haan, 2023)

Dentro del contexto América Latina y el Caribe, Chile se destaca como líder en la implementación de inteligencia artificial, evidenciado por su destacado puntaje en el índice, alcanzando 72.67 frente al promedio LATAM de 42.61. Este resultado indica que las empresas chilenas hacen un uso significativo de esta. Específicamente en apoyo gubernamental de la inversión en tecnologías y subindicadores de gastos gubernamentales en investigación y desarrollo. En contraste, Ecuador muestra un puntaje de índice de 22.17, situándose por debajo del promedio LATAM. Desde el 2020, el gobierno central y los gobiernos seccionales han empezado a utilizar herramientas IA en salud, turismo, banca, servicios fiscales y movilidad. No obstante, la adopción de la inteligencia artificial en el desarrollo sostenible en Ecuador tiene un gran potencial para abordar los desafíos ambientales y sociales que enfrenta el país, y al mismo tiempo, generar oportunidades de inversión y crecimiento económico (CEPAL, 2023).

A medida que Ecuador camina hacia un futuro más sostenible y tecnológicamente avanzado, la inteligencia artificial desempeñará un papel cada vez más importante en la protección de los recursos naturales, la mejora de la calidad de vida de los ciudadanos y el impulso de la innovación y la competitividad en el ámbito global. La IA (se está utilizando para mejorar la eficiencia y la sostenibilidad de la agricultura para abordar los desafíos actuales y futuros en la producción de alimentos. La inteligencia artificial se presenta como una herramienta clave para ayudar a los agricultores a adaptarse a los efectos del cambio climático y mitigar sus impactos. Al aprovechar la IA, los agricultores pueden tomar decisiones informadas basadas en datos y anticiparse a los cambios climáticos ((CEPAL), 2021).

Problemática

Cocholate S.A. es una empresa familiar orgullosamente ecuatoriana dedicada a la comercialización de cacao en grano, cumpliendo rigurosos estándares de calidad y seguridad alimentaria a través de un trabajo sostenible, socialmente responsable y amigable con el medio ambiente trabajando directamente con agricultores para asegurar confianza, satisfacción y las expectativas de los consumidores en calidad y sabor a los clientes tanto nacionales como internacionales. Su fundador, Ilton Rosado quien a sus 35 decidió apostar por la pepa de oro ecuatoriana, en sus inicios como comercializador y posteriormente junto a su esposa, exportadores.

El desafío principal al comprar cacao de diversos proveedores externos, la empresa trabaja con una cantidad significativa de información, desde detalles precisos de la cosecha hasta prácticas agrícolas específicas. No obstante, la calidad del cacao no solo depende de su cultivo, cosecha u otros factores climáticos y enfermedades de plantas, sino también se ve moldeada por una serie de procesos adicionales que garantizan características distintas en el producto final. Por esta razón, la gestión de la cadena de suministro y la selección de fuentes confiables de cacao son esenciales para garantizar una calidad constante en la producción de chocolate. Es fundamental implementar prácticas de clasificación y selección de cacao adecuadas. Esto implica la capacitación de los trabajadores que realizan la clasificación, fluctuaciones en los precios, la implementación de estándares de calidad y la realización de pruebas de captación para evaluar la calidad del cacao.

En primer lugar, fluctuaciones en los precios, conviene subrayar que los precios internacionales del cacao pueden ser volátiles debido a factores como la oferta y la demanda mundial, las condiciones climáticas, los cambios en la política comercial y las oscilaciones en las tasas de cambio, lo que puede afectar los márgenes de ganancia de la empresa. Concretamente, el cacao es una materia prima que se cotiza en bolsa, debido a su relevancia en la economía global y

su importancia en los mercados internacionales donde los productores, comerciantes y fabricantes de cacao pueden comprar y vender de manera eficiente utilizando contratos de futuros y opciones para cubrir o protegerse contra las variaciones adversas de precios.

El hedging se puede definir como un conjunto de estrategias que se utilizan en el mercado financiero para resguardar una inversión contra cualquier movimiento desfavorable que se pueda presentar en el mercado. Se asegura que, ante un impacto adverso del mercado, se minimice el impacto negativo sobre sus inversiones. (Quiroa, 2023, p.1)

En segundo lugar, la competitividad en la exigencia de cumplir con los rigurosos estándares de calidad y seguridad alimentaria en los mercados internacionales es de suma importancia para Cocholato S.A, de esta manera obtener la aprobación necesaria y poder mantenerse operativa de manera sostenible. Cumplir estos estándares implica inversiones significativas en infraestructura y en sistemas inteligentes de control de calidad. Certificaciones como, GFS (Global Food Safety), FSSC22000, ISO 22000, Halal, Kosher, Orgánico, Non OGM, Gluten free, entre otras. Estas certificaciones contribuyen de manera positiva a la implementación de buenas prácticas agrícolas, aseguran el cumplimiento de los estándares más exigentes en la producción y comercialización de cacao, de igual manera brinda una ventaja competitiva frente a los otros competidores (Perugachi, 2017)

La exigencia de cumplir con los rigurosos estándares de calidad y seguridad alimentaria en los mercados internacionales es de suma importancia para las empresas, ya que estas obtienen la aprobación necesaria y puedan mantenerse operativas de manera sostenible. Cumplir estos estándares implica inversiones significativas en infraestructura y en sistemas inteligentes de control de calidad. Certificaciones como, GFS (Global Food Safety), FSSC22000, ISO 22000, Halal, Kosher, Orgánico, Non OGM, Gluten free, entre otras. Estas certificaciones contribuyen de manera positiva a la implementación de buenas prácticas agrícolas, aseguran el cumplimiento de los

estándares más exigentes en la producción y comercialización de cacao, de igual manera brinda una ventaja competitiva frente a los otros competidores (Perugachi, 2017).

La actual operación de la empresa Cocholate S.A. se apoya en sistemas básico que desempeñan funciones que la gran mayoría de sus competidores utilizan como la recepción de mercadería, facturación, entre otras. Sin embargo, la falta de un sistema de inteligencia artificial deja a la empresa en una posición limitada para tomar decisiones. La carencia de análisis predictivo y procesamiento avanzados obstaculiza la capacidad de anticipar y abordar proactivamente los desafíos con relación a la calidad del cacao aroma fino (Plazas, López, & Corrales, 2017).

La necesidad de implementar un concepto de inteligencia artificial dentro de la empresa se evidencia por los retrasos operativos actuales. La introducción del machine learning no solo mejorará la eficiencia, sino que también brindará a la empresa la capacidad de tomar decisiones más informadas y puedan mantenerse competitivo, enfrentando desafíos del mercado día a día. Al implementarlo dentro de los procesos de fermentación del cacao, los algoritmos pueden indicar un insights sobre la duración óptima de la fermentación, las condiciones ideales y la identificación de problemas durante el proceso (Vite Cevallo, Carvajal Romero, & Barrezueta Unda, 2022).

La transición de sistemas tradicionales a soluciones más avanzadas implica un cambio significativo en la dinámica operativa, lo que puede generar resistencia dentro de la organización. Además, la falta de experiencia previa en el uso de algoritmos de machine learning por parte del personal puede ser un obstáculo. La capacitación del equipo es esencial, pero este proceso puede consumir tiempo y recursos, impactando temporalmente la eficiencia operativa durante la fase de implementación (Camacho Rosero & Reyna Muentes, 2022).

A nivel financiero, la inversión requerida para la implementación de machine learning puede representar también una carga significativa. La adquisición de hardware y software, junto

con la capacitación del personal, implica una asignación de recursos que debe ser cuidadosamente evaluada, especialmente en un entorno empresarial competitivo (Bajaña Morán & Paladines Gallardo, 2021).

Sin embargo, la omisión de su implementación podría acarrear consecuencias significativas para la empresa. Sin inteligencia artificial, la empresa continuaría enfrentándose desafíos en la gestión de variabilidad en la calidad del cacao, lo que afecta negativamente la consistencia de la calidad del cacao. La falta de capacidad para anticipar y abordar proactivamente problemas durante las etapas críticas, como la fermentación, podría resultar en lotes de cacao de calidad inferior, lo que impactara directamente en la satisfacción del cliente y la reputación de la empresa (Camacho Rosero & Reyna Muentes, 2022).

Sin duda, Cocholato SA tiene un enfoque sólido en la industria del cacao, pero también se enfrentan a desafíos relacionados con la fluctuación de precios, estándares de calidad internacionales y la necesidad de adoptar tecnologías avanzadas como la inteligencia artificial (IA) para mejorar sus procesos. La implementación de tecnologías como el aprendizaje automático podría revolucionar sus operaciones. Por ejemplo, en el proceso de fermentación del cacao, el uso de algoritmos de IA podría proporcionar información valiosa sobre la duración óptima de la fermentación y las condiciones ideales, mejorando la calidad del producto final. Sin embargo, la adopción de la IA conlleva desafíos propios, como la resistencia al cambio dentro de la organización, la necesidad de capacitar al personal y la asignación de recursos financieros significativos para la adquisición de tecnología y la formación del equipo.

Justificación

El presente proyecto investigativo tiene como finalidad central abordar las variabilidades que puedan afectar la calidad del cacao destino a la exportación en la empresa Cocholate S.A., considerando el significativo volumen de exportación de granos de cacao que constituyen una actividad económica fundamental para el Ecuador. La relevancia de este estudio radica en el entendimiento de que cualquier disminución en la calidad del cacao puede resultar en pérdidas económicas sustanciales, impactando directamente la confianza y satisfacción de nuestros clientes internacionales.

El cacao fino aroma en Ecuador desempeña un papel crucial en el mercado internacional, gracias a sus características distintivas y su excepcional calidad. Este producto no solo se distingue por su sabor exquisito, sino también por sus propiedades nutritivas, lo que garantiza una demanda significativa en una amplia gama de consumidores (González Rosado, Moreira Sornoza, & Dueñas Rivadeneira, 2022)

Es crucial reconocer que la industria del cacao aún enfrenta desafíos significativos. Van desde las fluctuaciones en los precios internacionales hasta las preocupaciones ambientales y las cambiantes tendencias del mercado. Abordar estas cuestiones se presenta como una tarea importante para asegurar la sostenibilidad a largo plazo del sector cacaotero ecuatoriano

Es crucial destacar que el proceso, desde la compra del cacao hasta la obtención del producto final, es complejo y susceptible a diversas influencias. Identificar y comprender estas variabilidades es esencial para establecer asertivos que garanticen una calidad super capaz de satisfacer las expectativas de nuestros clientes exigente.

La integración efectiva de Inteligencia Artificial en las estrategias empresariales no solo constituye un mero enfoque estratégico, sino que también actúa como un catalizador para la

apertura de nuevas oportunidades de inversión, el impulso de posibilidades de crecimiento y la optimización de procesos internos. Este tipo de inteligencia ya es una realidad, existen varios casos de éxitos palpables en el mercado actual. En América Latina y en diversas partes del mundo la implementación adecuada de la inteligencia empresarial emerge como un salto esencial para cumplir con las demandas del mercado contemporáneo. Aquellas empresas que han sabido implementar estas soluciones tecnológicas han logrado no solo mantenerse al ritmo de las exigencias cambiante, sino también destacarse y prosperar en un entorno empresarial cada vez más dinámico y competitivo.

Por ello, la implementación de esta solución tecnológica dentro del sector agrícola abre un abanico de oportunidades para mejorar la eficiencia y la toma de decisiones informadas. Al hacer uso de avanzadas herramientas analíticas, las empresas del ámbito agrícola pueden analizar datos de manera detallada datos críticos con relación al clima, gestión de cultivos, la cadena de suministro y las tendencias del mercado. Este enfoque permite una planificación más precisa de las siembras, optimización de recursos como agua y fertilizantes, y una gestión más eficiente de la cadena logística. Además, al acceder a información a tiempo real, los agricultores pueden anticipar desafíos y ajustar estrategias para maximizar la productividad.

Dentro del campo de la inteligencia artificial, podemos encontrar varios algoritmos, entre los cuales destaca el Random Forest o Bosque Aleatorio. Este algoritmo se erige como un recurso fundamental para las empresas al proporcionar predicciones precisas y robustas. Su impacto se extiende a áreas críticas como la optimización de la cadena de suministro, el manejo eficiente de datos faltantes, la prevención del sobreajuste y la mejora de la visibilidad operativa. El algoritmo Bosque Aleatorio lo convierte en una herramienta valiosa, adaptable a diversos contextos empresariales, proporcionando así a las empresas una ventaja competitiva a las potencias la eficiencia operativa y la toma de decisiones informadas.

La clasificación del cacao fino permite una selección cuidadosa de granos que cumplen con estándares específicos de calidad. Este enfoque asegura que los productos finales mantengan un sabor consistente y excepcional, satisfaciendo así las expectativas de los consumidores y fortaleciendo la reputación de la exportadora. Dado que numerosos mercados internacionales imponen estándares específicos para el cacao fino, el proceso de clasificación se vuelve indispensable. Esta práctica garantiza que los productos exportados cumplan con los requisitos de importación de diversos países, facilitando así la penetración en mercados internacionales. Asimismo, la comercialización de este cacao no solo posibilita que las exportadoras obtengan precios más elevados por sus productos, como se mencionó anteriormente, sino también les permite destacar en el mercado al proporcionar productos de alta calidad apreciados tanto por chocolateros como fabricantes de productos de cacao de lujo.

La evaluación integral de variables como sabor, aroma, textura, color, origen, procesamiento y más determina si un cacao se considera "fino aroma". Estos criterios, detallados a continuación, aseguran una clasificación precisa:

Tiempo de maduración: Un tiempo de maduración adecuado garantiza un perfil de sabor y aroma óptimo en el cacao. La identificación precisa de la coloración (rojo, amarillo, verde) es esencial para clasificar el cacao como "fino aroma", asegurando un producto de alta calidad que satisfaga las preferencias internacionales.

- a) Días de fermentación: Los días de fermentación son críticos para el desarrollo de los sabores característicos del cacao. Este proceso esencial contribuye a la complejidad y profundidad del sabor.
- b) Días de secado: El proceso de secado, siendo delicado y dependiendo del tipo de semilla, afecta directamente la calidad del cacao. Un secado inadecuado puede

resultar en pérdida de textura, color y sabor, lo que comprometería la aceptación en los mercados internacionales.

- c) Acidez: La acidez, medida por el pH, contribuye significativamente al perfil de sabor. Un equilibrio cuidadoso de acidez es crucial para producir cacao con un perfil de sabor balanceado y apreciado por los consumidores internacionales.
- d) Peso: El peso del grano de cacao no solo es un factor económico que determina el precio, sino que también puede indicar la presencia de frutos más desarrollados y, por ende, de mayor calidad.
- e) Pesticida: El manejo responsable de pesticidas es vital para cumplir con los estándares internacionales y garantizar la seguridad del producto. Prácticas sostenibles y libres de pesticidas en la medida de lo posible contribuyen a la aceptación en mercados internacionales cada vez más conscientes de la sostenibilidad.
- f) Métodos de cultivo: Los métodos de cultivo sostenibles afectan directamente la calidad del cacao. Prácticas respetuosas con el medio ambiente no solo mantienen la integridad del producto, sino que también responden a la creciente demanda internacional de productos sostenibles.
- g) Gestor de la planta: La supervisión del gestor de la planta es esencial para garantizar que todos los procesos se realicen de manera eficiente y cumpliendo con estándares de calidad. Esto es crítico para mantener la reputación y competitividad en los mercados internacionales.

En otras palabras. Este proyecto busca la necesidad inmediata de garantizar la calidad del cacao exportado y contribuye a la construcción de una base sólida para el crecimiento continuo y

sostenible de la industria cacaotera, alineándose con estándares internacionales y expectativas de nuestros futuros clientes.

La comprensión detallada de los factores que impulsan el éxito del sector cacaotero no solo beneficiará a los actores directos en la industria, sino que también contribuirá al desarrollo sostenible a nivel nacional. Al abordar las cuestiones fundamentales identificadas, este proyecto se pretende ser un impulsor para que la industria del cacao siga creciendo de manera positiva.

Alcance

La aplicación de algoritmos de machine learning para clasificar la calidad del cacao fino aroma podría ser relevante para un amplio público que incluye a agricultores, exportadores, investigadores, amantes del chocolate, profesionales de la industria y académicos interesados en la calidad de los productos agrícolas y la tecnología de machine learning.

- a) Agricultores y productores de cacao: Los agricultores que cultivan cacao fino aroma podrían estar interesados en aprender cómo los algoritmos de machine learning pueden ayudar a mejorar la calidad de su producto y, en última instancia, aumentar sus ganancias.
- b) Empresas exportadoras de cacao: Las empresas que exportan cacao fino aroma a nivel internacional podrían encontrar valiosa información sobre cómo utilizar algoritmos de machine learning para optimizar la clasificación y selección de los granos de cacao antes de la exportación.
- c) Investigadores y científicos: Los expertos en el campo de la ciencia de los alimentos y la agricultura podrían estar interesados en conocer las últimas investigaciones y avances en la aplicación de machine learning en la industria del cacao.
- d) Amantes del chocolate: Los consumidores apasionados por el chocolate de alta calidad pueden encontrar interesante entender cómo se evalúa y clasifica el cacao fino aroma, ya que esto puede influir en la calidad del chocolate que compran.
- e) Profesionales de la industria del chocolate: Expertos en cata de chocolate y profesionales de la industria del chocolate podrían estar interesados en conocer las técnicas y tecnologías utilizadas para evaluar y clasificar el cacao fino aroma, ya que esto influye en la calidad de sus productos.

- f) Estudiantes y académicos: Estudiantes de ciencias agrícolas, ciencia de datos, inteligencia artificial y disciplinas relacionadas pueden encontrar inspiración y oportunidades de investigación en este tema, brindando conocimientos valiosos y una comprensión más profunda de cómo se aplican estos conceptos en un contexto real.

En definitiva, el presente trabajo investigativo contribuye al conocimiento al introducir enfoques y métodos innovadores en el ámbito de la clasificación de la calidad del cacao fino aroma, con el propósito de aportar una base sólida para la ampliación del entendimiento de los lectores interesados en el tema. En pocas palabras, comprender el uso y aplicaciones prácticas en la industria alimentaria y agrícola de algoritmos de aprendizaje automático. Asimismo, la calidad del cacao fino es un tema relevante global económica y socialmente por su impacto significativo en los productores y exportadores de cacao, pero sobre todo relacionada con la sostenibilidad agrícola y la preservación del medio ambiente. En pocas palabras, proporciona una perspectiva académica como práctica porque ofrece conocimientos valiosos, ideas innovadoras que contribuyen al avance en la industria alimentaria y agrícola.

Objetivos

Objetivo General

Identificar, implementar y validar un modelo de Machine Learning para la clasificación eficiente de la calidad del cacao fino aroma, con el fin de mejorar el proceso de exportación para la empresa Cocholato S.A.

Objetivos específicos

- a) Realizar un análisis detallado de los conceptos relacionados con el tema del proyecto investigativo mediante una revisión literaria amplia.
- b) Analizar diversos modelos estadísticos basados en la inteligencia artificial y su aplicabilidad para la clasificación de la calidad del cacao.
- c) Evaluar la eficacia del algoritmo del machine learning en la clasificación de la calidad del cacao fino aroma.

Preguntas de investigación

- a) ¿Cuáles son los criterios claves para clasificar la calidad del cacao dentro de la empresa Cocholato S.A.?
- b) ¿Cómo afectaría la implementación de machine learning a la eficiencia del proceso de clasificación dentro de la empresa Cocholato S.A.?
- c) ¿Cuál sería el impacto potencial en la competitividad de la exportadora Cocholato S.A. al implementar este método de clasificación de calidad de cacao basado en machine learning?

Capítulo II

Marco Teórico

La Inteligencia Artificial (IA) ha captado la atención desde sus inicios en la mitad del siglo XX hasta la actualidad, emergiendo como un tema complejo y extenso que resulta desafiante de estudiar y analizar debido a su arraigado en las ciencias exactas. Este campo de conocimiento refleja la constante búsqueda humana de herramientas que puedan imitar y, en algunos casos, superar procesos cognitivos, reduciendo al mismo tiempo la necesidad de interacción humana. Según la definición proporcionada por Munárriz (1994) “La inteligencia artificial es una ciencia que tiene como objetivo el diseño y construcción de máquinas capaces de imitar el comportamiento inteligente de las personas” (p. 19)

En la comprensión de la utilidad de la implementación del Machine Learning (ML) en las empresas, es esencial considerar las teorías que sirven como base para la investigación. Este proyecto investigativo profundiza en la relevancia de la IA en la actualidad, destacando su evolución y la complejidad a su aplicación en diverso ámbito (Corvalán, 2018).

Introducción al Aprendizaje Estadístico

La Ciencia de Datos es el conjunto de conocimientos y herramientas utilizados en las distintas etapas del análisis de datos. Es decir, un campo multidisciplinar, con importantes aportaciones estadísticas e informáticas, dentro del que se incluirían disciplinas como minería de datos, aprendizaje automático, aprendizaje profundo, modelado predictivo y también el aprendizaje estadístico. (IBM, 2021).

El aprendizaje estadístico ha experimentado un crecimiento significativo gracias a la convergencia de la disponibilidad de datos, la potencia computacional y los avances en algoritmos

que permiten el descubrimiento de patrones complejos en datos de alta dimensionalidad. Su objetivo fundamental es capacitar a las máquinas para realizar tareas complejas basadas en datos, lo que ha llevado a su aplicación en una amplia gama de campos para la automatización y la toma de decisiones fundamentadas (Munárriz, 1994).

Inteligencia Artificial

La Inteligencia Artificial (IA) se define como “La habilidad de los ordenadores para realizar actividades que normalmente requieren inteligencia humana”, según la descripción proporcionada por Rouhiainen en el año 2018. Es decir que, la IA ha revolucionado la forma en las maquinas puedan hacer tareas que anteriormente requerían de inteligencia humana. Este campo de estudio se basa en el desarrollo y aplicación de algoritmos que permiten a las máquinas aprender de los datos y aplicar ese conocimiento en la toma de decisiones, imitando procesos cognitivos humanos. A diferencia de los seres humanos, las máquinas basadas en la IA pueden operar sin la necesidad de descanso, lo que convierte en herramientas altamente eficientes para analizar grandes volúmenes de información de manera continua. Este atributo, combinado con una proporción de errores significativamente menor en comparación con las tareas realizadas por humanos, destaca la capacidad única de la IA para mejorar la precisión y la eficacia en diversas áreas de aplicación (Rouhiainen, 2018).

La Inteligencia Artificial (IA) ha tenido un impacto significativo en ámbito empresarial y en nuestra vida cotidiana. Según Russel y Norvig (2010), resalta la importancia de los agentes inteligentes. Estos agentes son sistemas diseñados para percibir su entorno mediante sensores, actuar en respuesta a este entorno a través de actuadores y tomar decisiones con el fin de alcanzar objetivos específicos. La interacción dinámica de estos agentes con su entorno, así como su capacidad para adaptarse y aprender, son elementos cruciales en la conceptualización de la

inteligencia artificial. Además, el libro explora una amplia variedad de enfoques y técnicas en el campo de la IA. Esto incluye aspectos como la resolución de problemas, el aprendizaje automático o conocido como Machine Learning, la representación del conocimiento, el razonamiento, el aprendizaje profundo o también conocido como Deep Learning, las redes neuronales y otros conceptos fundamentales que contribuyen a la comprensión de la Inteligencia Artificial (p- 2-33).

Machine Learning

El aprendizaje automático es una rama en evolución de los algoritmos computacionales que están diseñados para emular la inteligencia humana aprendiendo del entorno circundante. Las técnicas basadas en el aprendizaje automático se han aplicado con éxito en diversos campos que van desde el reconocimiento de patrones, la biología computacional hasta las aplicaciones biomédicas y médicas. El grado de complejidad de estos procesos puede variar e implicar varias etapas de sofisticadas interacciones hombre-máquina y toma de decisiones, lo que naturalmente invitaría al uso de algoritmos de aprendizaje automático para optimizar y automatizar estos procesos. La capacidad de los algoritmos de aprendizaje automático para aprender del contexto actual y generalizar en tareas no vistas permitiría mejorar tanto la seguridad como la eficacia de la práctica de la radioterapia, lo que redundaría en mejores resultados (Chamorro, Laza, Noriega, Rojano, & Heredia, 2021)

Aprendizaje Supervisado

El aprendizaje supervisado, según los principios establecidos en el campo del aprendizaje automático, se caracteriza por entrenar un modelo utilizando un conjunto de datos que contiene pares de entrada y salida conocidos. Cada instancia del conjunto de datos de entrenamiento consiste en datos de entrada acompañados de la salida correspondiente, también conocida como etiqueta. Durante el proceso de entrenamiento, el modelo ajusta sus parámetros internos a medida que se

expone a más ejemplos de entrenamiento, con el objetivo de aprender la relación subyacente entre las entradas y las salidas. Este enfoque se aplica a dos tipos principales de problemas: regresión, donde la salida es un valor continuo, y clasificación, donde la salida es una etiqueta o categoría discreta (James, Witten, Hastie, & Tibshirani, 2023).

El aprendizaje supervisado es esencial en diversas aplicaciones, desde predecir precios en problemas de regresión hasta clasificar correos electrónicos en problemas de clasificación, ofreciendo así una base fundamental para la toma de decisiones automatizada en una amplia variedad de campos (Murphy, 2012).

Modelo de clasificación

La clasificación emplea un algoritmo para destinar con precisión los datos de prueba en categorías determinadas. Identifica entidades específicas dentro del conjunto de datos e intenta sacar algunas conclusiones sobre cómo esas entidades deben etiquetarse o definirse. Por ejemplo, existen algoritmos de clasificación lineal, también hay máquinas de vectores de soporte (SVM), los árboles de decisión, los bosques aleatorios y k vecinos más cercanos (IBM, 2023).

Este tipo de modelo utiliza ecuaciones matemáticas asociadas a:

Regresión logística: Es un modelo comúnmente utilizado para problemas de clasificación binaria (0,1). La función logística se utiliza para transformar una combinación lineal de las características a una probabilidad (Jurafsky & H. Martin, 2008).

$$P\left(Y = \frac{1}{X}\right) = \frac{1}{1 + e^{-(WX+b)}}$$

Donde,

$P(Y=1/X)$ es la probabilidad de que la clase sea 1 dadas las características de X.

W son los pesos del modelo.

B es el sesgo

Máquinas de soporte vectorial (SVM): Se busca encontrar el hiperplano óptimo que maximiza el margen entre las dos clases.

$$y(x) = \text{sign}(x \cdot w + b)$$

Donde,

W son los pesos del modelo.

X son las características de entrada.

B es el sesgo.

Aprendizaje No Supervisado

El aprendizaje no supervisado es un enfoque de aprendizaje automático donde el algoritmo se entrena en un conjunto de datos sin etiquetas. A diferencia del aprendizaje supervisado, donde el modelo se entrena con datos etiquetados (entrada y salida esperada), en el aprendizaje no supervisado, el algoritmo debe descubrir patrones o estructuras en los datos sin tener la guía explícita de etiquetas (Russell & Norvig, 2004).

En "Inteligencia Artificial: Un Enfoque Moderno" de Russell y Norvig, se aborda el aprendizaje no supervisado como una forma de explorar la estructura oculta o la organización subyacente en

los datos. Algunos de los enfoques comunes dentro del aprendizaje no supervisado incluyen (Russell & Norvig, 2004):

- a) Agrupamiento (Clustering): El agrupamiento implica dividir un conjunto de datos en grupos o clústeres basados en similitudes entre las muestras. Los algoritmos de agrupamiento intentan encontrar estructuras intrínsecas en los datos sin conocer las etiquetas de antemano (IBM, 2023).
- b) Reducción de Dimensionalidad: Este enfoque busca reducir la complejidad de los datos conservando características importantes. Métodos como Análisis de Componentes Principales (PCA) son ejemplos de técnicas de reducción de dimensionalidad en el aprendizaje no supervisado (IBM, 2023).
- c) Asociación: Identifica patrones de asociación o relaciones entre variables en grandes conjuntos de datos. El descubrimiento de reglas de asociación es un ejemplo común de este tipo de técnicas (IBM, 2023).

La aplicación práctica del aprendizaje no supervisado puede encontrarse en varios contextos:

- a) Segmentación de Mercado: Identificar grupos de clientes con comportamientos y preferencias similares sin tener información previa sobre las categorías (Russell & Norvig, 2004).
- b) Recomendación de Contenido: Descubrir patrones de comportamiento en usuarios para ofrecer recomendaciones personalizadas sin utilizar etiquetas explícitas (Russell & Norvig, 2004).
- c) Detección de Anomalías: Identificar patrones anómalos o inusuales en datos, como fraudes en transacciones financieras (Russell & Norvig, 2004).

- d) **Análisis de Imágenes y Sonido:** En el procesamiento de imágenes y señales de audio, el aprendizaje no supervisado puede ayudar a encontrar patrones y características relevantes sin tener etiquetas específicas (Russell & Norvig, 2004).

Aprendizaje por Refuerzo

El aprendizaje por refuerzo es un enfoque del aprendizaje automático en el cual un agente interactúa con su entorno y aprende a través de la retroalimentación positiva o negativa que recibe en función de las acciones que realiza. El agente tiene como objetivo maximizar una señal de recompensa a lo largo del tiempo, aprendiendo así a tomar decisiones que conduzcan a resultados más favorables (Russell & Norvig, 2004).

Algunos conceptos clave relacionados con el aprendizaje por refuerzo incluyen:

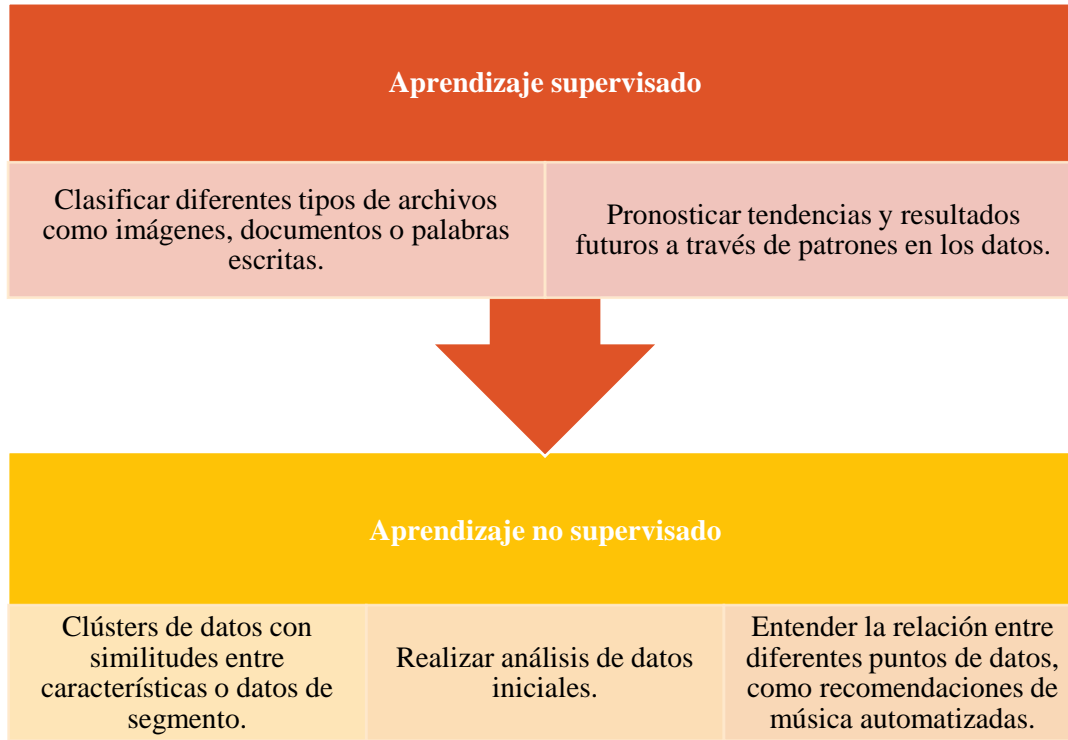
- a) **Agente:** La entidad que realiza acciones en un entorno para lograr sus objetivos.
- b) **Entorno:** El contexto o el mundo en el que el agente opera y realiza acciones.
- c) **Estado:** La representación de la situación actual del agente en el entorno.
- d) **Acción:** La decisión o la operación que el agente elige realizar en un estado dado.
- e) **Recompensa:** La señal que el agente recibe del entorno como consecuencia de sus acciones.

La recompensa puede ser positiva, negativa o neutra.

Diferencias entre aprendizaje supervisado y no supervisado

Figura 1

Aprendizaje supervisado vs No supervisado



Nota. *Explicación de las diferencias entre el aprendizaje supervisado y no supervisado. Fuente: International Business Machines (IBM)

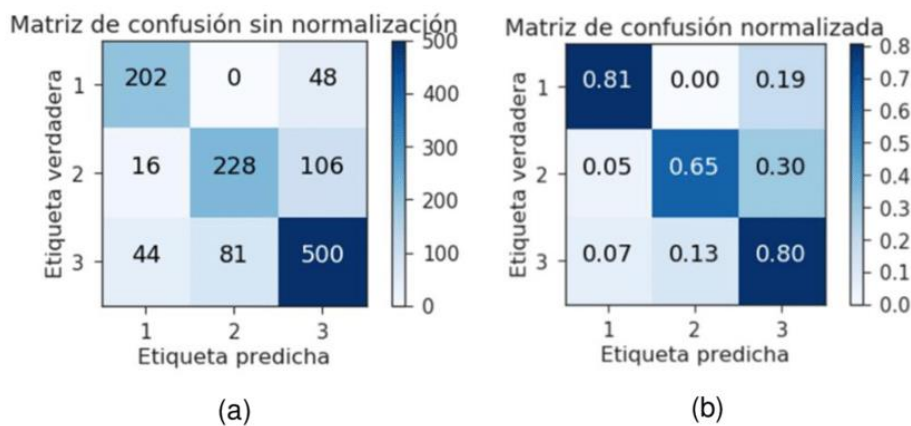
La principal diferencia entre el aprendizaje supervisado y el no supervisado es la necesidad de datos etiquetados, dicho de otra manera, aprendizaje supervisado se basa en datos de entrada y salida etiquetados, mientras que el aprendizaje no supervisado procesa datos sin etiquetar o sin procesar. Por otra parte, el aprendizaje automático supervisado generalmente se usa para clasificar datos o hacer predicciones, mientras que el aprendizaje no supervisado generalmente se usa para comprender las relaciones dentro de los conjuntos de datos. (Universidad Europea, 2022)

Técnicas principales de clasificación

Matriz de confusión. La matriz de confusión es una herramienta que se utiliza en el campo de la clasificación en el aprendizaje automático para evaluar el rendimiento de un modelo al comparar sus predicciones con las verdaderas etiquetas de un conjunto de datos. Proporciona una visión detallada de cómo el modelo está realizando la clasificación en términos de verdaderos positivos, verdaderos negativos, falsos positivos y falsos negativos (Ariza-López, Rodríguez-Avi, & Alba-Fernández, 2018).

Figura 2

Matriz de confusión



Nota. Se visualiza una matriz de confusión sin normalización y con normalización. Fuente:

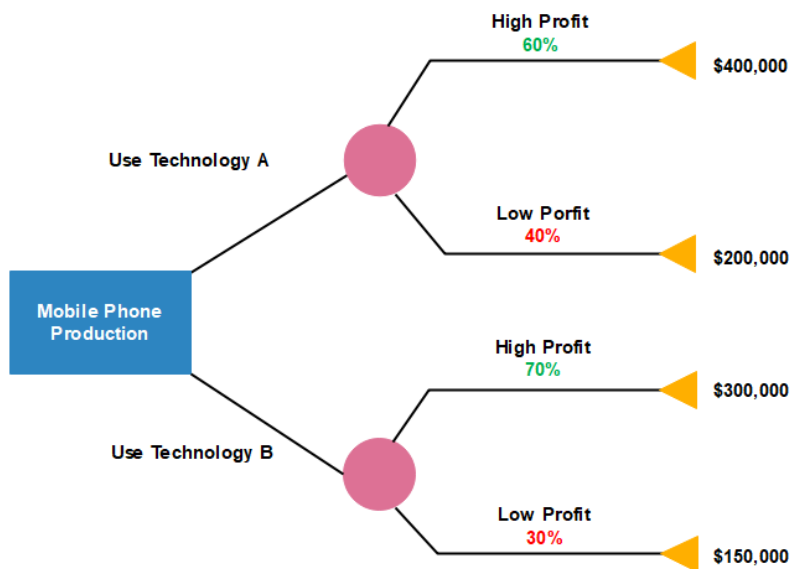
ResearchGate

Árbol de decisión. Los árboles de decisión son modelos de aprendizaje supervisado que se utilizan tanto para clasificación como para regresión. En este enfoque, el conjunto de datos se divide recursivamente en subconjuntos más pequeños, y en cada nodo del árbol se toma una decisión basada en una característica específica. Esta división se realiza de manera que maximice la homogeneidad de las clases en cada rama (Munárriz, 1994)

- a) Nodos de decisión: Representado por un cuadro. Indica un punto crítico donde se debe tomar una decisión durante el proceso.
- b) Nodo de probabilidad: Representado por un círculo. Señala la presencia de un evento aleatorio en ese punto específico.
- c) Rama: Se refiere a los caminos que conducen a la toma de decisiones, conectado nodos a lo largo del proceso.
- d) Nodo final: Estos son los resultados finales, marcando el final de una serie de decisiones y caminos tomados.

Figura 3

Árbol de decisión



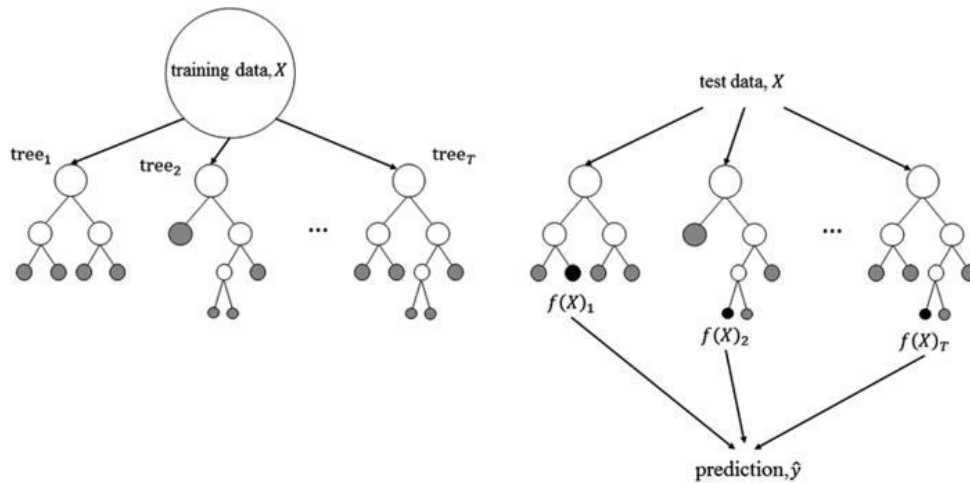
Nota. Recuperado de EDraw (2023)

Bosque Aleatorio. Un bosque aleatorio es un modelo de aprendizaje supervisado que se utiliza para tareas de clasificación, regresión y otras aplicaciones predictivas. Se basa en la construcción de múltiples árboles de decisión durante el proceso de entrenamiento y se combina

para obtener una predicción más robusta y generalizada (James, Witten, Hastie, & Tibshirani, 2023).

Figura 4

Bosque Aleatorio



Nota. Random Forest Diagram. Daniel James Mennit (2014)

Un bosque aleatorio consta de múltiples árboles de decisión que trabajan de manera conjunta para realizar predicciones. Aunque visualizar todos los árboles individualmente en un bosque puede ser complejo (IBM, 2023).

1. Inicio

El bosque aleatorio comienza con un conjunto de datos de entrenamiento.

2. Construcción de árbol

Se construyen árboles de decisión utilizando subconjuntos aleatorios de datos y características.

Cada árbol se desarrolla de manera independiente.

3. Proceso de entrenamiento

Cada árbol sigue el proceso de entrenamiento de un árbol de decisión estándar, dividiendo el conjunto de datos en nodos basados en características específicas.

La aleatorización en la construcción implica seleccionar subconjuntos aleatorios de datos y considerar solo un conjunto aleatorio de características en cada división de nodo.

4. Conjunto de arboles

El bosque aleatorio final consiste en un conjunto de estos árboles independientes y aleatorizados.

5. Predicción

Para realizar predicciones, se pasa una nuestra instancia de cada árbol del bosque.

En clasificación, se puede realizar una votación para determinar la clase más frecuente entre los árboles.

En regresión, se promedian las predicciones de todos los árboles

Así mismo, la técnica del bosque aleatorio es una técnica de clasificación más acertada que existe el día de hoy. Breiman (2001) ha afirmado lo siguiente “ El algoritmo Random Forest es una técnica de aprendizaje supervisado que genera múltiples árboles de decisión sobre un conjunto de datos de entrenamiento, los resultados obtenidos se combinan a fin de obtener un modelo único más robusto en comparación con los resultados de cada árbol por separado" (p, 2)., dicho de otra manera, es un método preciso para seleccionar las características más importantes para el clasificador, con relación al cacao de fino aroma.

Las ventajas de los bosques aleatorios incluyen su capacidad para manejar conjuntos de datos grandes, la resistencia al sobreajuste y la capacidad de manejar características faltantes y datos desequilibrados. Además, pueden proporcionar una estimación de la importancia de las características, lo que puede ser útil para seleccionar características relevantes en el proceso de modelado. En resumen, los bosques aleatorios son una técnica de conjunto que combina múltiples árboles de decisión para realizar predicciones más precisas y robustas. La aleatorización en la construcción de árboles y la combinación de múltiples árboles son las características clave que los hacen efectivos (Medina Merino & Ñique Chacón, 2017).

La correlación entre árboles se da cuando el árbol es un modelo adecuado para describir la relación entre los predictores y la respuesta, y también cuando uno de los predictores es muy fuerte, es decir, es especialmente relevante, con lo cual casi siempre va a estar en el primer corte. Esta correlación entre árboles se va a traducir en una correlación entre sus predicciones (más formalmente, entre los predictores) (Medina Merino & Ñique Chacón, 2017).

Promediar variables altamente correlacionadas produce una reducción de la varianza mucho menor que si promediamos variables incorrelacionadas. La solución pasa por añadir aleatoriedad al proceso de construcción de los árboles, para que estos dejen de estar correlacionados. Hubo varios intentos, entre los que destaca Dietterich (2000) al proponer la idea de introducir aleatoriedad en la selección de las variables de cada corte. Breiman (2001) propuso un algoritmo unificado al que llamó bosques aleatorios. En la construcción de cada uno de los árboles que finalmente constituirán el bosque, se van haciendo cortes binarios, y para cada corte hay que seleccionar una variable predictora.

Funcionamiento de los bosques aleatorios. Los algoritmos de bosque aleatorio tienen tres hiperparámetros principales, que deben establecerse antes del entrenamiento. Estos son el tamaño

de nodo, el número de árboles y el número de características muestreadas. A partir de aquí, se puede utilizar el clasificador de bosque aleatorio para resolver problemas de regresión o clasificación.

El algoritmo de bosque aleatorio se compone de una colección de árboles de decisiones, y cada árbol del conjunto se compone de una muestra de datos extraída de un conjunto de entrenamiento con sustitución, que se denomina muestra de programa de arranque. De esa muestra de entrenamiento, un tercio se establece aparte como datos de prueba, lo que se conoce como la muestra OOB (Out-Of-Bag), a la que volveremos más adelante. A continuación, se inyecta otra instancia de aleatoriedad a través de la agregación autoencodificadas de características, lo que añade más diversidad al conjunto de datos y reduce la correlación entre los árboles de decisiones. Dependiendo del tipo de problema, la determinación de la predicción variará. Para una tarea de regresión, los árboles de decisiones individuales se promediarán y, para una tarea de clasificación, un voto mayoritario (por ejemplo, la variable categórica más frecuente) dará como resultado la clase pronosticada. Por último, se utiliza la muestra OOB para la validación cruzada, lo que finaliza la predicción (Medina Merino & Ñique Chacón, 2017).

Marco Conceptual

Cacao

El cacao es fruto del árbol teobroma cacao y se cultiva principalmente por sus semillas, las cuales son la materia prima esencial para la elaboración de chocolate. Estas semillas contienen compuestos como la teobromina, que contribuyen al distintivo sabor del chocolate. Más allá de ser un placer culinario, el cacao ha sido apreciado a lo largo de la historia por diversas culturas debido a sus propiedades estimulantes. Su cultivo desempeña un papel crucial en las economías de regiones tropicales, sirviendo como fuente de ingresos para los agricultores (CEFA, 2023).

Compra de cacao

Proceso de adquisición de los granos de cacao. Este proceso involucra transacciones comerciales entre productores de cacao y compradores, que pueden ser empresas chocolateras u otros intermediarios. La compra de cacao juega un papel crucial en la cadena de suministro. Las negociaciones de precios, calidad y condiciones comerciales son elementos clave en estas transacciones (International Cocoa Organization, 2022).

Cadena de suministro

Es una función estratégica y logística que involucra todas las operaciones que son indispensables para que una mercancía logre llegar al cliente final en óptimas condiciones. Esto es, desde la búsqueda de materias primas, su posterior transformación y hasta la fabricación, transporte y entrega al consumidor final. (Roldan, 2020)

Rentabilidad

Hace referencia a los beneficios que se han obtenido o se pueden obtener de una inversión. La rentabilidad económica (RE) es diferente de la rentabilidad financiera (RF). La rentabilidad

económica tiene en cuenta todos los activos utilizados para generar rentabilidad. En contraste, la rentabilidad financiera sólo tiene en cuenta la cantidad de recursos propios utilizados. (Arias, 2020)

Perfil del cliente

El perfil de cliente es un conjunto de información demográfica, psicográfica y de comportamiento que caracteriza al comprador ideal de una marca. Crear el perfil de cliente sirve para facilitar la atracción de consumidores “ideales” según las características de cada negocio, de esa forma, será más fácil para el prospecto de negocio avanzar a través del embudo de conversión. (Zendesk, 2022)

Servicio

Un servicio, en el ámbito económico, es la acción o conjunto de actividades destinadas a satisfacer una determinada necesidad de los clientes, brindando un producto inmaterial y personalizado. Por otro lado, al ser los servicios principalmente intangibles o no materiales (más adelante ahondaremos al respecto), un consumidor no los puede poseer. Un claro ejemplo de este tipo serían los servicios de atención al cliente, habituales en grandes marcas. (Galán, 2020)

Exportación

Es básicamente todo bien y/o servicio legítimo que el país productor o emisor (el exportador) envíe como mercancía a un tercero (importador), para su compra o utilización. (Oca, 2020)

El mercado mundial del cacao

El comercio mundial de cacao en grano, que constituye la base de la economía nacional de la mayoría de los países subdesarrollados, ha experimentado cambios significativos de declive que han afectado las relaciones de intercambio de los países productores de materias primas, debido a una combinación de factores, como la volatilidad de los precios, la concentración de la producción, los problemas de sostenibilidad, la demanda de chocolate y la influencia de organizaciones y

acuerdos comerciales. En 1980, por ejemplo, tales exportaciones de los países de América Latina y el Caribe representaban 82% del valor total de las exportaciones FOB de bienes, mientras que para el año 2001 esa cuota había descendido a 41% (CEPAL, 2003).

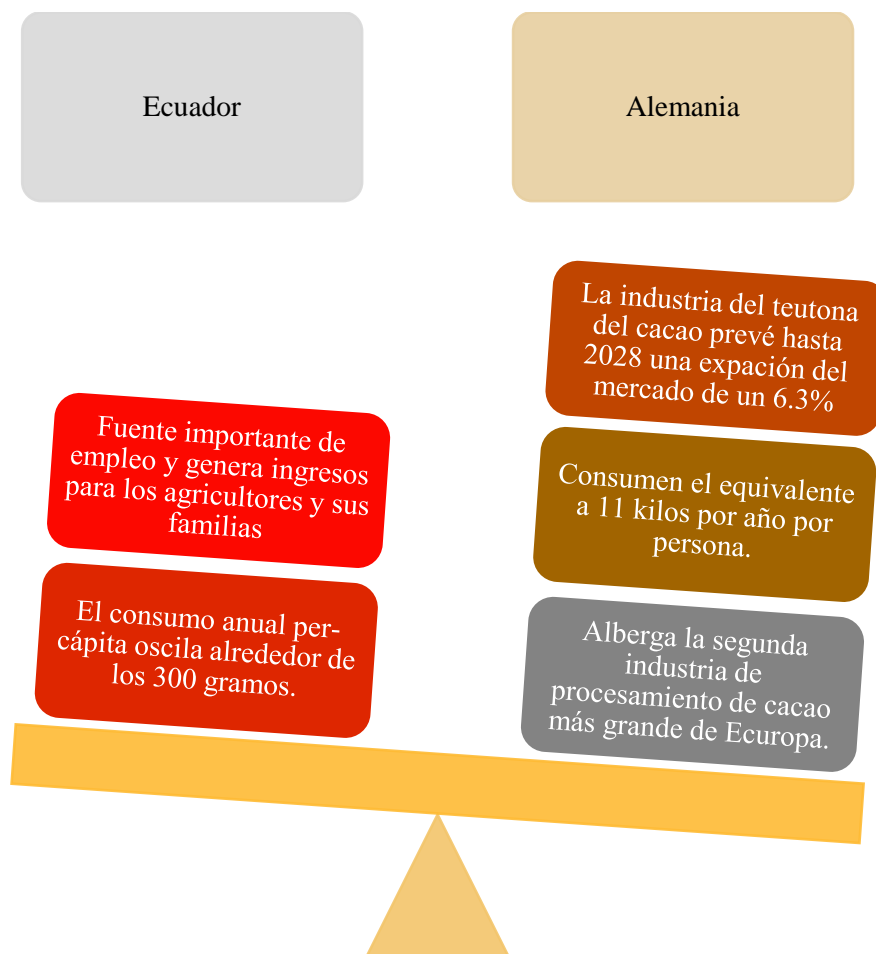
Bajo este escenario el presente artículo tiene como objetivo caracterizar el comercio mundial de cacao en grano planteando inicialmente la descripción y evolución de las principales variables que afectan el comercio mundial, así como los cambios en su funcionamiento, debido a que el cacao es considerado bien de lujo, no es indispensable, es decir no es considerado bien de primera necesidad en la alimentación. Destacando principalmente que la calidad del grano ha sido otro factor importante en la evolución del comercio mundial, debido a los cambios en el tipo de grano ofertado y en la fijación de precios.

En los últimos años se ha evidenciado una serie de cambios de gran importancia en la economía mundial, de productos alimentarios, entre estos cambios se pueden resaltar el cacao en grano, con un enfoque creciente en la estabilidad de precios, la sostenibilidad, la inversión en la cadena de valor, la diversificación de productos y mercados, la adaptación al cambio climático, y la importancia del comercio justo y la trazabilidad. Estos cambios reflejan los desafíos y las oportunidades en esta industria en constante evolución.

El consumo mundial anual de chocolate es de 7,5 millones de toneladas. Los principales países consumidores de chocolate, en volumen total, son Estados Unidos, Alemania, Reino Unido, Francia y la Federación de Rusia, con un consumo per cápita es cerca de un kilogramo por persona (María Liliana Quintero R; Katty Marisabel Díaz Morales, 2004). Por otra parte, en Ecuador el consumo per cápita de chocolate es de 800 gramos por persona, un número para nada despreciable teniendo en cuenta la población que posee el país.

Figura 5

Mercado actual del mercado



Nota. Mercado de cacao entre Ecuador y Alemania

Exportación de cacao ecuatoriano

A finales del siglo XIX Ecuador se convirtió en el mayor exportador de cacao. La exportación de cacao ecuatoriano mejora la balanza comercial del país, debido a que genera empleo y bienestar a la población. Además, es un ingrediente clave de tradición gastronómica en la creación de chocolate de alta gama por sus características fundamentales en sabor y aroma.

El cacao ecuatoriano se exporta en diferentes presentaciones como producto vegetal y también como subproductos de origen vegetal, es decir, granos de cacao, semielaborados y/o producto terminado:

- a) Granos de cacao: La forma más básica y cruda del producto. Estos granos son extraídos de las vainas de cacao y son la materia prima esencial para la producción de chocolate.
- b) Semielaborados: Ecuador exporta también cacao en forma de productos semielaborados, como la pasta de cacao, tabletas o barras de cacao, jugo de cacao, nibs de cacao, polvo de cacao, manteca de cacao y licor de cacao. Estos productos son el resultado de procesos de transformación que pueden incluir la fermentación, secado, tostado y molienda de los granos de cacao
- c) Producto terminado: Además de los productos semielaborados, Ecuador exporta chocolate y otros productos de cacao ya terminados. Estos productos pueden incluir tabletas de chocolate, bombones, cacao en polvo, entre otros.

Los productos semielaborados derivados del cacao son todos aquellos que se obtienen a partir del procesamiento de esta semilla y determinados procesos de fermentación y otro de secado. A diferencia del cacao en grano, para una buena calidad en los diferentes subproductos, lo esencial es, maquinaria para chocolatería, utensilios y técnicas profesionales, que permitan exactitud en los pasos y tiempos de cada proceso.

Se certificaron 331.028,57 toneladas de cacao en grano para la exportación, por un valor FOB aproximado de 758 millones de dólares; esto significa un incremento del 2% con relación a lo exportado en 2020, donde se registró una exportación de 325.208,04 toneladas de cacao. (ANECACAO, 2022). En cuanto a la calidad referida en la Norma Técnica Ecuatoriana NTE-INEN 176, se registra que el 81% del cacao certificado corresponde a cacao fino de aroma con

266.789,23 toneladas exportadas y un 19% a cacao CCN-51 con 64.239,34 toneladas exportadas. Este importante producto vegetal ha sido destinado hacia 40 diferentes destinos mundiales, de los cuales los más destacados son: Estados Unidos, Indonesia, Malasia, Países Bajos, México, Bélgica Italia y Alemania. (ANECACAO, 2022). Por otra parte, el sector de exportación de cacao y de semielaborados productos sumaron USD 940 millones marcando un récord de ventas por segundo año consecutivo.

Descripción y principales características del cacao

El cacao (*Theobroma cacao* L.) es una planta originaria de América tropical, perteneciente a la familia Sterculiaceae. De acuerdo con la clasificación botánica, el cacao pertenece al género *Theobroma*, especie de cacao, una planta valiosa que ha desempeñado un papel fundamental en la industria del chocolate y en la cultura de muchas sociedades. Además de su importancia económica, el cacao es apreciado por su sabor y sus beneficios para la salud.

El cacao es un árbol perenne que puede alcanzar una altura de 4 a 8 metros. Sus hojas son grandes y perennes, de forma elíptica y de color verde oscuro y las condiciones más propicias para el cultivo se encuentran confinadas a las áreas de los bosques húmedos tropicales, con temperaturas entre 20 y 30 grados Celsius y una precipitación anual de al menos 1.500 mm sin vientos fuertes ni constantes porque pueden romper ramas y dañar la planta. El cacao produce flores pequeñas y hermafroditas en el tronco y las ramas principales. Estas flores son de color blanco o amarillo pálido y contienen tanto órganos masculinos como femeninos. Los insectos, como los mosquitos y las abejas, son responsables de la polinización de las flores de cacao, lo que es esencial para el desarrollo de los frutos.

Clasificación botánica del cacao

La clasificación botánica es importante en la agricultura y la industria del cacao para comprender las diferentes características de las plantas y seleccionar las variedades adecuadas para

su cultivo y procesamiento. Cada variedad de cacao puede tener características distintivas que influyen en el sabor y la calidad del chocolate producido a partir de sus granos.

- a) Cacao Criollo: El árbol que da este fruto es frágil, de poco vigor, sensible y de bajo rendimiento, lo cual explica su nivel de producción tan rara y escasa (1% de producción mundial), destacándose la alta calidad de sus semillas. Sin embargo, se destaca por la alta calidad de sus semillas, su fruto es grande y de color verde o rojo. Los granos son grandes y de color claro. Es un cacao excepcional, aromático y poco amargo, con un olor de cacao dulce unido a un aroma delicado característico. Se cultiva en Méjico, Nicaragua, Guatemala, Colombia, Venezuela, Madagascar, Sri Lanka, Indonesia (Java), Islas Samoa., en el Caribe y Nueva Guinea Papua.
- b) Cacao Forastero: El árbol que lo produce es muy resistente a las enfermedades, de mayor tolerancia a las que el cacao Criollo. El fruto es de color verde y amarillo, los granos son aplanados de un color violeta, tiene un sabor fuerte y su aroma es amargo y ácido. Representa aproximadamente un 95% de la producción mundial, proveniente de los países de África Occidental y Brasil. Es el más cultivado (80% de producción mundial) y básicamente se planta en Perú, Ecuador, Colombia, Brasil y África en general.
- c) Cacao Trinitario: Este tipo de cacao es un híbrido del cacao forastero y el criollo, es decir, de inferior calidad, pero es más resistente y productivo que en ocasiones reemplaza las plantaciones de Criollo. En cuanto a la producción a nivel mundial representa entre un 10% y un 15%. Se empezó a cultivar en la Isla de Trinidad (Caribe) cuando un ciclón arrasó las plantaciones de Criollo, pero actualmente es producido en Granada, Jamaica, Trinidad y Tobago, Colombia, Venezuela y América Central.

Clasificación comercial del cacao

La clasificación comercial del cacao se basa en diferentes factores, como el tamaño, el contenido de grasa y el perfil de sabor de los granos de cacao. Esta clasificación es fundamental para la comercialización y la industria del chocolate, que es esencial para los productores, fabricantes de chocolate y consumidores, ya que ayuda a determinar la calidad y las características de los granos de cacao utilizados en la producción de chocolate y otros productos a base de cacao.

- a) Cacao ordinario: Granos producidos por los cacaos tipo forastero, un ejemplo es Nacional de Ecuador, considerado fino, pero que proviene de árboles tipo Forastero, éstos son utilizados en la fabricación de manteca de cacao y de productos que tengan una elevada proporción de chocolate. Este cacao se utiliza en la producción de chocolate masiva, productos de cacao grandes y aplicaciones industriales y carece de las notas florales, afrutadas o especiadas, es decir, no se distingue por un sabor distintivo.
- b) Cacao fino o de aroma: En términos generales, los granos de cacaos criollos y trinitarios corresponden a lo que en el mercado mundial se conoce como cacao fino o de aroma. Los granos correspondientes a esta categoría dan características específicas de aroma o color en chocolates finos de revestimientos o capas de cobertura. También se usan (aunque cada vez menos) para producir cacao en polvo que se emplea como aroma en algunas recetas y en la preparación de algunos alimentos y bebidas. El cacao fino o de aroma se define como cacao libre de defectos de aroma que al mismo tiempo ofrece un perfil aromático complejo que refleja la pericia del productor y del “terroir”, o se asocia con el entorno específico en el que se cultiva, se fermenta y se seca el cacao. El cacao fino o de aroma que reúne estos criterios básicos también puede ofrecer una notable diversidad genética, además de

un patrimonio histórico y cultural. Se define como cacao que apenas tiene defectos aromáticos y que ofrece unas características aromáticas o de sabor valoradas tradicionalmente en las mezclas. El cacao fino o de aroma que reúne estos criterios básicos de calidad también puede ofrecer una notable diversidad genética, además de un patrimonio histórico y cultural.

Comercio internacional del cacao

El cacao es uno de los productos agroalimentarios de mayor penetración en el mercado mundial. El comercio internacional del cacao implica la exportación e importación de granos de cacao y productos relacionados en todo el mundo. En el caso particular del cacao en grano, se pueden distinguir diferentes actores y procesos de comercialización (producción, transformación, comercialización y consumo final); intervienen así productores, intermediarios (brokers), casas de bolsa, empresas consumidoras (transformadoras y comercializadoras de productos con base de cacao) y organizaciones

Principales productores de cacao

El proceso de comercialización del cacao es largo, complejo y requiere coordinación y colaboración a lo largo de toda la cadena de suministro, involucrando a numerosos actores, desde los agricultores que cultivan el cacao hasta los fabricantes de productos de chocolate y los minoristas que venden esos productos. La cadena de suministro incluye actividades como la cosecha, el procesamiento, la exportación, el transporte, la fabricación y la distribución. Cada etapa de esta cadena puede involucrar a múltiples intermediarios y empresas, lo que puede hacer que sea difícil rastrear el origen del cacao en un producto final. Sin embargo, la calidad, la sostenibilidad y la ética en la producción y comercialización del cacao son aspectos cada vez más importantes a medida que los consumidores buscan productos de cacao de alta calidad y responsables.

Estándares de calidad del cacao

Los estándares de calidad del cacao son criterios y parámetros utilizados para evaluar la calidad del cacao en términos de su sabor, aroma, contenido de humedad, tamaño de los granos, entre otros factores. Estos estándares varían según las organizaciones y las regulaciones de la industria del cacao y el chocolate. Organismos como la Organización Internacional del Cacao (ICCO) y diversas certificaciones (Fair Trade, UTZ, Rainforest Alliance) establecen requisitos específicos para asegurar la calidad y la sostenibilidad en la producción de cacao

Según datos de la Organización Internacional del Cacao, el nuevo estándar ISO para cacao podrá ser una opción de mejora para la producción de cacao en el Ecuador. La ISO 2451, especifica los requisitos, clasificación, muestreo, métodos de prueba, empaque y marcado de los granos de cacao, con recomendaciones adicionales relacionadas con el almacenamiento y la desinfección. Esta norma cubre todo, desde tamaño y color hasta contenido de humedad, preparación y clasificación de los granos. (Servicio de acreditación ecuatoriano, 2018).

Tabla 1

Características del grano fermentado

Características del grano seco	Grano bien fermentado	Grano que le faltó fermentado	Grano sin fermentar
Forma	Hinchado	Algo Aplanado o pacho	Aplanado o panchito
Color del grano por fuera	Café oscuro	Amarillo claro, amarillo rojizo	Blanquecino o rojizo
Cascarilla	Se desprende fácilmente al tocarla con los dedos	Difícil arrancarla con las uñas	No se desprende. Está pegada al grano
Consistencia del grano	Fácil de quebrar y desbaratar con los dedos	Se desbarata con los dedos	Es duro como de hule, solo se puede partir con navaja

El grano por dentro	Está todo quebradito	No representa quebradura	Muy duro y sólido
Color del grano por dentro	Color chocolate o café claro	Entre cenizo y morado	Violeta
Olor	A chocolate aromático	A vinagre desagradable	Sin olor o con olor a moho
Sabor o gusto	Amargo agradable	Amargo	Muy amargo

Nota. Fuente: Cacaomovil (2024).

Calidad del grano fresco en baba

Los centros de acopio y beneficiado compran grano fresco de cacao. Este grano se conoce como cacao en baba puesto que el grano en la mazorca madura está rodeado de mucilago blanco y húmedo. Es necesario evaluar la calidad de este grano, ya que conlleva a obtener un grano seco de calidad. (Aguilar, 2016).

- a) Debe provenir de mazorcas maduras, sanas y sin daños. Esto solo se puede evaluar en área de partido de la mazorca el cual generalmente es en la parcela.
- b) El grano en baba debe ser blanco, húmedo y brillante. El cacao amarillento es indicativo de tener mucho tiempo de extraído. Si la baba está seca o poco jugosa y opaca es indicativo de provenir de mazorca inmadura, sobre madura o enferma.
- c) No debe haber granos negros, café o donde trasluce el color café del grano

Figura 6

Calidad del cacao



Nota. El grano en su estado de mucílago debe presentar un aspecto blanco, encontrarse en una condición húmeda y exhibir un brillo característico.

Marco Legal

En el ámbito legal se han generado varios planes alineados a una normativa legal entre los cuales se destacan: El Plan Nacional de Telecomunicaciones y Tecnologías de Información del Ecuador 2016-2021. El Convenio Marco de Cooperación Interinstitucional firmado con los diferentes entes como el Ministerio de Telecomunicaciones y de La Sociedad de la Información, la Secretaría de Educación Superior, Ciencia, Tecnología e Innovación.

Los Planes Nacionales para el Desarrollo (2013-2017). La Ley Orgánica de Telecomunicaciones. El Código Orgánico de La Economía Social de los Conocimientos, Creatividad e Innovación. Así como los artículos que contemplan el desarrollo de la investigación científica y de la innovación tecnológica, de las TIC y diversas áreas del conocimiento que están en la Constitución de la República del Ecuador de los cuales se puede destacar los (Art. 281, Art. 385, Art. 387, Art. 423, Art. 388, Art. 313) que permiten alinearse y ser partícipes de la adopción de estas tecnologías emergentes para el cambio y adopción de una cultura de transformación digital para un mejor desarrollo a nivel empresarial en el ámbito local con proyección internacional, lo que se convierte en un impacto positivo.

Capítulo III

Metodología

Para este proyecto investigativo se empleó el algoritmo de Random Forest o también conocido como Bosque Aleatorio para la clasificación. Ya que con la combinación de múltiples arboles de decisión, cada uno entrenado en una muestra aleatoria del conjunto de datos. Esto reduce la sensibilidad a que pequeñas variaciones en los datos de entrenamiento, lo que resulta ser un modelo más robusto y menos propensos al sobreajuste. Además, se utiliza la matriz de confusión, con el objetivo de evaluar el rendimiento del modelo.

Pasos para desarrollar

Para la construcción del modelo, optaremos por utilizar RStudio, un software que ofrece una interfaz intuitiva para facilitar la manipulación del lenguaje de programación R. Esta herramienta proporciona mayor comodidad y fluidez en comparación con el uso directo del lenguaje R. Para proceder, es fundamental contar con una base de datos preparada, siendo recomendable tener un mínimo de 50 observaciones por cada clase o categoría. A continuación, se presenta un breve ejemplo de la base de datos que será objeto de estudio en este proyecto:

Tabla 2*Base de datos tesis*

HectareasNa	HectareasCCN51	EdadPlantacion	lotes	Diasfermentacion	DiasSecado	Pesticida	MCultivo	Genetica	Acidez
0	15	6	3	4	30	1	2	1	3
2	3	23	1	3	75	3	4	2	3
18	2	14	4	4	45	2	1	1	3,5
4	20	16	3	5	75	1	4	3	3
10	13	4	2	6	67	1	2	1	3
7	3	23	6	7	34	2	3	3	3,2
9	4,5	40	7	8	41	3	1	2	3,3
4	4	21	10	9	75	1	4	3	3
2	3	34	12	9	70	3	3	2	3,2
6	2	23	2	9	15	2	2	1	3,5
7,5	2	12	1	8	40	2	2	3	3,5
12	5	3	4	7	50	2	3	1	3,5
4	12	4	4	10	23	3	3	2	3,4
11	2,5	9	5	8	56	1	3	2	3,3
5	2	10	3	6	34	3	2	1	3,2
1	3	14	8	4	75	2	1	3	3
1	2	15	10	6	63	2	3	1	3
10	10	1	12	5	12	3	2	2	3,5
4	3	40	2	4	43	1	2	2	3
3	13	38	4	10	30	3	3	1	3
2	3	35	5	9	61	2	2	3	3,2

Nota. Visualización de los primeros 21 datos que se trabajará a lo largo del proyecto.

Para iniciar el proceso, será necesario acceder a la base de datos mediante la función (*read.csv*) o (*read.csv2*), dependiendo del formato de base de datos que estemos utilizando. Es esencial seleccionar la función adecuada según el tipo de base de datos que estemos incorporando, asegurando así una integración óptima con nuestro modelo.

```
cacao <- read.csv2(DATOS/CACA0MyA.csv)
```

Una vez que hayamos preparado la base de datos, podemos iniciar la codificación de las funciones esenciales para la generación del algoritmo Random Forest o Bosque Aleatorio. En esta fase, es crucial descargar una serie de paquete. Para utilizar los paquetes en RStudio, se lleva a cabo la instalación de los paquetes a través de la función (*install.packages*) y luego se hace referencia a las bibliotecas mediante la función (*library*), como se observa a continuación:

```
#INSTALACIÓN DE PAQUETES
install.packages("caret")
install.packages("gridExtra")
install.packages("ggplot2")
install.packages("randomForest")
install.packages("tidyverse")
install.packages("factoextra")
install.packages("cluster")
install.packages("Nbclust")
install.packages("GGally")
library(GGally)
library(caret)
library(gridExtra)
library(ggplot2)
library(randomForest)
library(tidyverse)
library(cluster)
library(factoextra)
library(NbClust)
```

Los paquetes por instalar dependerán en el algoritmo y base de datos a utilizar. A continuación, se explicará detalladamente los paquetes instalados para este proyecto.

Caret. El paquete *caret* (abreviatura de Classification And Regression Training) contiene funciones para agilizar el proceso de entrenamiento de modelos para problemas complejos de regresión y clasificación. El paquete utiliza una serie de paquetes de R, pero intenta no cargarlos todos al iniciar el paquete (eliminando las dependencias formales de los paquetes, el tiempo de inicio del paquete puede reducirse considerablemente). *Caret* carga los paquetes según sea necesario y asume que están instalados. Si falta un paquete de modelado, aparece un aviso para instalarlo (Berrendero, RPubS, 2017).

GridExtra. El paquete *grid* proporciona funciones de bajo nivel para crear objetos gráficos (*grobs*) y colocarlos en una página en ventanas específicas. El paquete *gtable* introdujo un esquema de diseño de más alto nivel, posiblemente más susceptible de interacción con el usuario. Con el par de funciones *arrangeGrob/grid.arrange ()*, *gridExtra* se basa en *gtable* para organizar múltiples *grobs* en una página. La función *grid.arrange ()* dibuja en el dispositivo; para disposiciones más complejas, puede que queramos almacenar la *gtable* y combinarla con otros objetos, por ejemplo, formando disposiciones anidadas. Para ello, utilice *arrangeGrob ()* (Auguie, 2017).

Ggplot 2. Es un paquete para graficar, que facilita crear gráficas complejas a partir de datos en un data frame. Incluye varias funciones para especificar que variables graficar, como éstas son expuestas, y varias otras características visuales. El paquete *ggplot2* permite personalizar los gráficos con temas. Es posible personalizar cada detalle de un gráfico, como los colores, tipos de líneas, fuentes o alineación, entre otros, usando los componentes de la función *theme*. Además, la librería proporciona otras funciones que permiten añadir títulos, subtítulos, líneas, flechas o textos (Ballari, 2018).

Algunas capas tienen una orientación específica. *geom_bar ()* por ejemplo, tendrá las barras a lo largo de un eje, *geom_line ()* ordenará la entrada por un eje, etc. El enfoque original para

utilizar estos geoms en la otra orientación era añadir `coord_flip ()` al trazado para cambiar la posición de los ejes x e y. A partir de `ggplot2 v3.3` todos los geoms funcionarán de forma nativa en ambas orientaciones sin `coord_flip ()`. El mecanismo es que la capa intentará adivinar la orientación a partir de los datos mapeados, o tomar la dirección del usuario usando el parámetro `orientation`. Para replicar esta funcionalidad en nuevas stats y geoms hay que dar unos pasos. Veremos la capa `boxplot` como ejemplo en lugar de crear una nueva desde cero (Ballari, 2018).

Algunas capas no tendrán una interpretación clara de sus datos en términos de orientación. Un ejemplo clásico es `geom_line ()` que sólo por convención corre a lo largo del eje x. No hay nada en los datos que lo indique. Para estos geoms el usuario debe indicar una orientación invertida estableciendo `orientation = "y"`. El stat o geom llamará entonces `has_flipped_aes ()` con `ambiguous = TRUE` para cancelar cualquier suposición basada en el formato de los datos. Como ejemplo podemos ver el método `setup_params ()` de `GeomLine ()` (Ballari, 2018)

randomForest. Implementa el algoritmo de Bosques Aleatorios de Breiman (basado en el código original en Fortran de breiman y Cutler) para clasificación y regresión. También puede utilizarse en modo no supervisado para evaluar proximidades entre puntos de datos (Abdala, 2020).

Tidyverse. Es un conjunto de paquetes que funcionan de manera armoniosa debido a que comparten representaciones de datos comunes y un diseño de 'API'. Este paquete está diseñado para facilitar la instalación y carga de varios paquetes del 'tidyverse' en un solo paso (Araneda, 2021).

Clúster. El agrupamiento es un método para encontrar subgrupos de observaciones dentro de un conjunto de datos. Cuando realizamos agrupamiento, necesitamos que las observaciones en el mismo grupo tengan patrones similares y que las observaciones en diferentes grupos sean diferentes entre sí (Romero, Ramirez, Rojas, & Darghan, 2020).

Si no hay una variable de respuesta, es adecuado para un método no supervisado, lo que implica que busca encontrar relaciones entre las n observaciones sin estar entrenado por una variable de respuesta. El agrupamiento nos permite identificar grupos homogéneos y categorizarlos en el conjunto de datos (Romero, Ramirez, Rojas, & Darghan, 2020).

factoExtra. El paquete factoextra en R simplifica la extracción y representación visual de los resultados de análisis exploratorio de datos multivariados, abordando distintas técnicas, tales como (R-Project, 2020):

NbClust. El paquete ofrece 30 índices para determinar el número de grupos y sugiere al usuario el mejor esquema de agrupamiento a partir de los diferentes resultados obtenidos al variar todas las combinaciones de número de grupos, medidas de distancia y métodos de agrupamiento.

Notas sobre el argumento "Distancia" Las siguientes medidas de distancia se aplican a dos vectores, x e y . Se utilizan cuando los datos son vectores de 2 dimensiones que surgen al medir 2 características en cada uno de los n objetos o individuos (RDocumentation, 2014).

SKcacao < -scale(cacao)

Después de instalar la serie de paquetes a utilizar, se procederá a llevar a cabo la normalización de los datos mediante la escalación. Este proceso tiene como objetivo garantizar que todas las variables consideradas tengan igual relevancia y previene errores durante la ejecución del programa. La escalación se llevará a cabo mediante el empleo de la función "Scales".

Scale: Las escalas gráficas asignan datos a estéticas y ofrecen métodos para determinar automáticamente los intervalos y etiquetas de ejes y leyendas. En este caso lo hicimos con la base de datos llamado cacao.

Luego de normalizar los datos, el siguiente paso consistirá en determinar el número ideal de clústeres para llevar a cabo la división mediante el método de k-means.

```

#ESTIMAR EL NUMERO DE CLUSTERS
NumClust <- NbClust(SKcacao, distante = "euclidean",
                    min.nc = 2, max.nc = 10, method = "kmeans",
                    index = "allong")
fviz_nbclust(numClust)

```

En este punto, se emplea el paquete NbClust () para determinar el número óptimo de clústeres. Se utiliza la distancia euclidiana con el propósito de minimizar la distancia media entre los puntos y los centroides. Se establece un rango de evaluación de clústeres desde un mínimo de 2 hasta un máximo de 10. Se opta por el método kmeans y el índice de mediación será constante en todo momento. El paquete trabaja de la siguiente manera:

Data: Es la base de datos con el que trabajara el paquete. En este caso SKcacao, debido a que se realizó el escalamiento de los datos.

Distance: Es la métrica de distancia que se empleará para calcular la matriz de disimilitud; en este caso, se utiliza la métrica “euclidiana”.

Euclidiana: Distancia cuadrática habitual entre los dos vectores (norma 2).

Min.nc: Representa el número mínimo de clústeres y se encuentra en el rango de 1 al total de objetos -1.

Max.nc: Representa el número máximo de clústeres, y su valor va desde 2 hasta el total de objeto -1.

Method: Indica el método de análisis de conglomerados a emplear, siendo “kmeans” la elección es este caso.

Index: Hace referencia al índice que se calculará, siendo en este caso “allong” para los clústeres.

Allong: Se utiliza para calcular múltiples índices de validación del agrupamiento, lo que nos proporcionará una visión más completa de la calidad de los clústeres generados por el algoritmo K-means.

Luego de la ejecución del programa, se importarán dos conjuntos de datos en RStudio.

CALCULAMOS LOS DOS CLÚSTER

```
k6 <- kmeans(SKcacao, centers = 6, nstart = 25)
k6
str(k6)
```

Como siguiente paso, calculamos los clústeres, para ello se utiliza el paquete (kmeans) para realizar el algoritmo de K-means en los datos escalados anteriormente. Se especifica que se desean encontrar 6 centro de clústeres y se realiza el proceso de inicio múltiple para mejorar la estabilidad del algoritmo.

Centers: Se requiere especificar el número de clústeres, denotado como “k”, o proporcionar un conjunto inicial de clústeres distintos. En este caso de especificar un número, se elige de manera aleatoria un conjunto de filas (distintas) de x como los centros iniciales.

Nstart: Si, center es un número, ¿Cuántos conjuntos aleatorios deben seleccionarse?

#PLOTAR LOS CLÚSTER

```
fviz_cluster(k6, data = SKcacao)
fviz_cluster(k6, data = SKcacao, ellipse.type = euclid,
              repel = TRUE, start.plot = TRUE) # ellipse.type = "t", "norm", "euclid"
fviz_cluster(k6, data = SKcacao, ellipse.type = "norm")
fviz_cluster(k6, data = SKcacao, ellipse.type = norm,
              palette = Set2, ggtheme = theme_minimal())
```

Luego plotamos los clústeres, para visualizar los resultados del análisis de clúster utilizando la codificación *fviz_clúster ()*. Asimismo, plotamos los clústeres, pero esta vez con elipse configurada en “euclid” para visualizar elipses de confianza y con “repel=TRUE” para evitar

superposiciones de etiquetas. La opción “star. plot=TRUE” agrega un gráfico de estrella para mostrar la varianza explicada.

Data: La información empleada para la formación de los clústeres. Es necesario especificar esto solo cuando el objeto pertenece a la categoría kmeans o dbscan.

Ellipse.type: Se refiere a un carácter que especifica el tipo de marco. Los valores posibles son “convex”, “confidence” o tipos admitidos por stat_ellipse, incluyendo uno de (“t”, “norm”, “euclid”).

Repel: Un valor booleano que indica si se debe utilizar ggrepel para evitar la superposición de etiquetas de texto o no.

Star.plot: Controla la visualización de estrellas en el gráfico de clústeres.

Palette: Se refiere a la paleta de colores que se utilizará en el gráfico.

Ggtheme: Función que trabaja de la mano con el paquete ggplot2.

En conjunto, la codificación escrita proporciona diversas visualizaciones de los clústeres resultantes del análisis de K-medias, ofreciendo diferentes perspectivas sobre la distribución y la forma de los grupos identificado en los datos escalados.

#CODIFICO LA VARIABLE DE RESPUESTA

```
cacao$K_Means <- factor(cacao$K_Means,
                      levels = c(1,2,3,4,5,6),
                      labels = c("Exportacion 1", "Nacional",
                                "Exportación 3", "Exportación 4",
                                "Exportacion 5", "Exportacion 6"))
cacao$Pesticida <- factor(cacao$Pesticida,
                        levels = c(1,2,3),
                        labels = c("Fungicida", "Herbicida", "Insecticida"))
cacao$MCultivo <- factor(cacao$MCultivo,
                       levels = c(1,2,3,4),
                       labels = c("Agroforesteria", "EnPlanta",
                                   "Monocultivo", "Poda"))
cacao$Genetica <- factor(cacao$Genetica,
                       levels = c(1,2,3),
                       labels = c("criollo", "forastero", "trinitario"))
```

Como siguiente paso, codificamos nuestras variables categóricas, utilizando la función (factor). Se especifican los niveles y las etiquetas para cada variable.

Levels: Este parámetro es un vector opcional que contiene los valores únicos que podría haber tomado la variable categórica “x”. Estos valores se proporcionan como cadenas de caracteres.

Labels: Otro parámetro que puede ser un vector de etiquetas para los niveles o puede ser una cadena de caracteres de longitud 1.

```
#PLANTAMOS SEMILLA  
set.seed(123)
```

Después de la codificación de las variables, establecemos una semilla, que marca el inicio de la implementación de procedimientos asociados a árboles. Aunque existan diversas opciones de semillas, en este contexto de clasificación, la selección más apropiada es la semilla (123).

```
#ENTRENAMOS LOS DATOS
```

```
dataentrenamiento <- createDataPartition (cacao$K_Means, p = 0.80, list  
= FALSE)
```

Luego preparamos los datos de entrenamiento.

createDatapartition: Crea particiones de prueba/entrenamiento en un conjunto de datos. Es útil para dividir los datos en conjunto de entrenamiento y prueba de manera estratificada, asegurando una distribución equitativa de las clases entre los conjuntos.

p: Este parámetro representa el porcentaje de datos que se destinará al conjunto de entrenamiento. Determina la proporción de datos que se utilizará para entrenar un modelo en comparación con el conjunto total de datos.

list: Es un parámetro lógico que determina el formato de los resultados. En este caso se utiliza “FALSE”, los resultados se presentarán en una matriz donde el número de filas es igual a “floor(p*length(y))” y el número de columnas es el número de veces que se realiza la partición.

#ARMAMOS EL BOSQUE ALEATORIO

```
RandomTreeModel <- randomForest (x = cacao {dataentrenamiento, 1:10},  
                                  y = cacao {dataentrenamiento, 11},  
                                  ntree = 3000, keep. forest = TRUE
```

```
RandomTreeModel$importance
```

#REALIZO LA PREDICCIÓN

```
Prediccion <- predict(RandomTreeModel, cacao)
```

```
Prediccion
```

Después, construimos el bosque aleatorio con la función del paquete (randomForest), especificando las filas correspondientes al conjunto de entrenamiento (x=cacao [dataentrenamiento, 1:10]), luego se selecciona la columna 11 de la base de datos como la variable respuesta. Además, se especifica el número de árboles en el bosque aleatorios (3000). Realizamos la predicción.

X: Representa un marco de datos o una matriz que contiene las variables predictoras para ajustar el modelo.

Y: Es un vector de respuesta.

Xtest: Representa un marco de datos o una matriz similar a “x” que contiene las variables predictoras para el conjunto de prueba.

Ytest: Es el vector de respuesta correspondiente al conjunto de prueba.

Ntree: Representa el número de árboles que se deben cultivar en el bosque aleatorio.

Keep.forest: Si esta función se establece en “FALSE”, el bosque no se retendrá en el objeto de salida.

Predict: Es una función que se utiliza para realizar predicciones a partir de los resultados de diversos modelos ajustados. La función invoca métodos particulares que dependen de la clase del primer argumento proporcionado.

```
#ARMO LA MATRIZ DE CONFUSIÓN
```

```
Matriz <- table(cacao$K_Means, Prediccion, dnn = c("Actual", "Predicho"))
```

```
Matriz
```

```
#VALORES CLASIFICADOS
```

```
cacao$Prediccion <- predict(RandomTreeModel, cacao
```

```
#PROBABILIDAD DE ÉXITO
```

```
cacao$RandProb <- predict(RandomTreeModel, cacao, type = "prob")
```

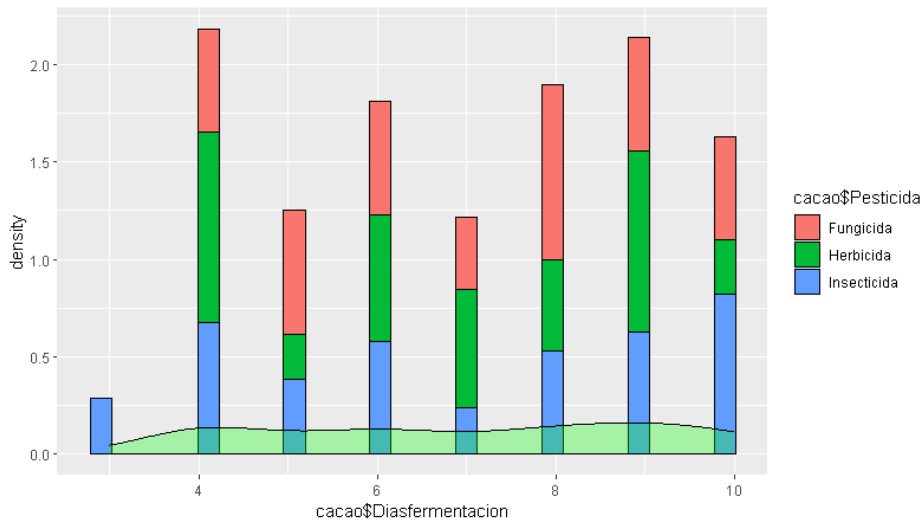
Finalmente, se evalúa el rendimiento del modelo de bosque aleatorio mediante la construcción de una matriz de confusión y realiza predicciones en el conjunto de datos completo, incluyendo la asignación de probabilidades de éxito a cada observación.

Resultados

Antes de introducirnos en los resultados detallados de los modelos que se realizó anteriormente en RStudio, es esencial realizar un análisis descriptivo que nos permita entender a fondo las características fundamentales las variables claves. A continuación, presentaremos una serie de gráficos descriptivos que proporcionaran una visión visual de las tendencias y patrones observado en nuestros datos:

Figura 7

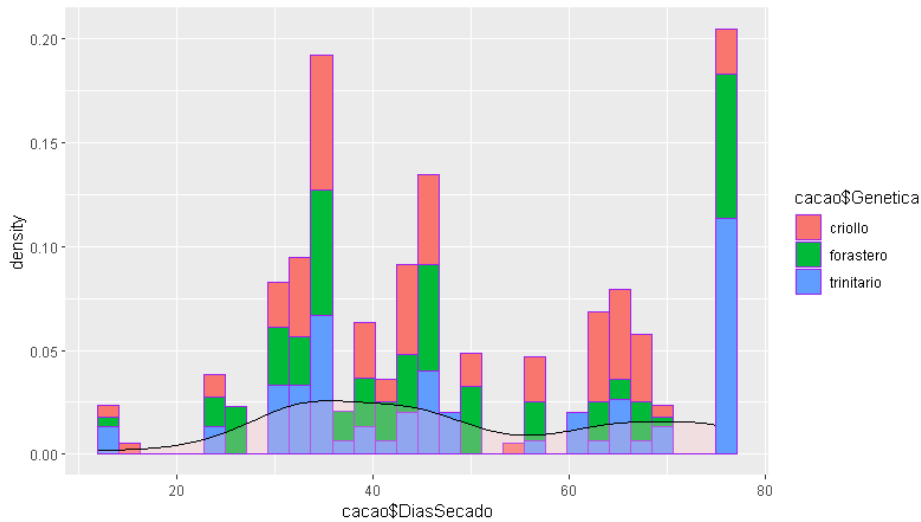
Gráfico de normalidad pesticida - días de fermentación



En la figura 7, Un análisis descriptivo de variables de pesticidas para el cacao implica examinar y describir las características clave de los datos relacionados con el uso de pesticidas en la producción de cacao, identificando y enumerando los diferentes tipos de pesticidas utilizados en la producción de cacao.

Figura 8

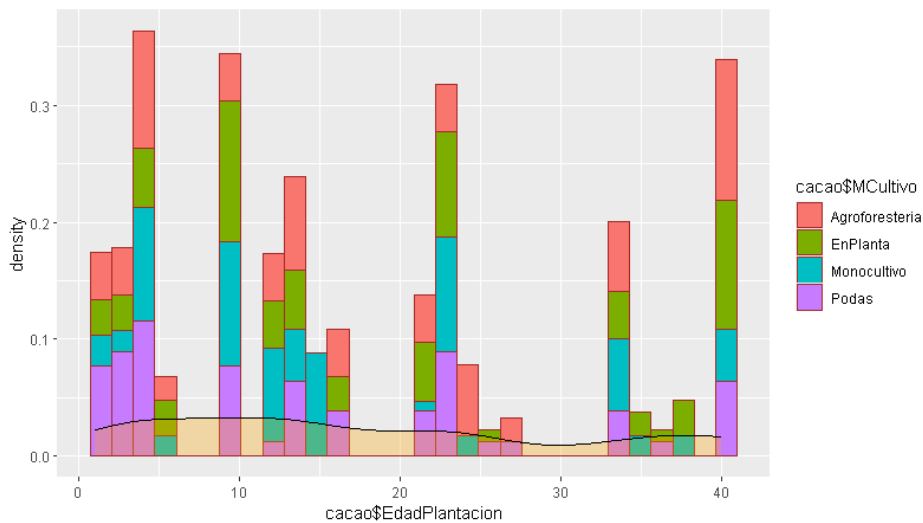
Gráfico de normalidad Genética - días de secado



En la figura 8, El gráfico de normalidad se utiliza para evaluar si la variable relacionada con la genética del cacao sigue una distribución normal. Puede representar, por ejemplo, la diversidad genética entre diferentes variedades o cepas de cacao, proporcionando información valiosa para programas de reproducción y mejora genética.

Figura 9

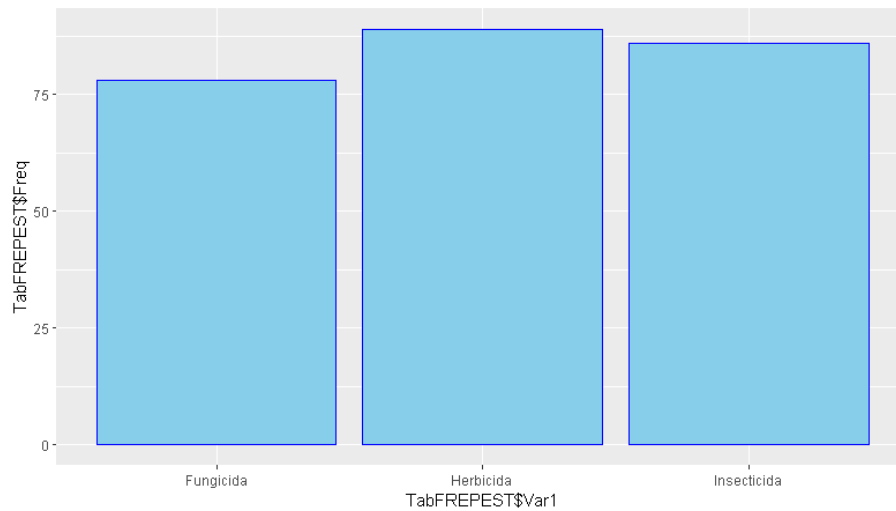
Gráfico de normalidad MCultivo - Edad de plantación



En la figura 9, El gráfico de normalidad aplicado al método de cultivo del cacao proporciona una herramienta visual para evaluar la distribución de una variable específica y entender la variabilidad asociada al método de cultivo, esto puede ser útil para optimizar prácticas agrícolas, identificar áreas de mejora y tomar decisiones informadas para mejorar la producción de cacao.

Figura 10

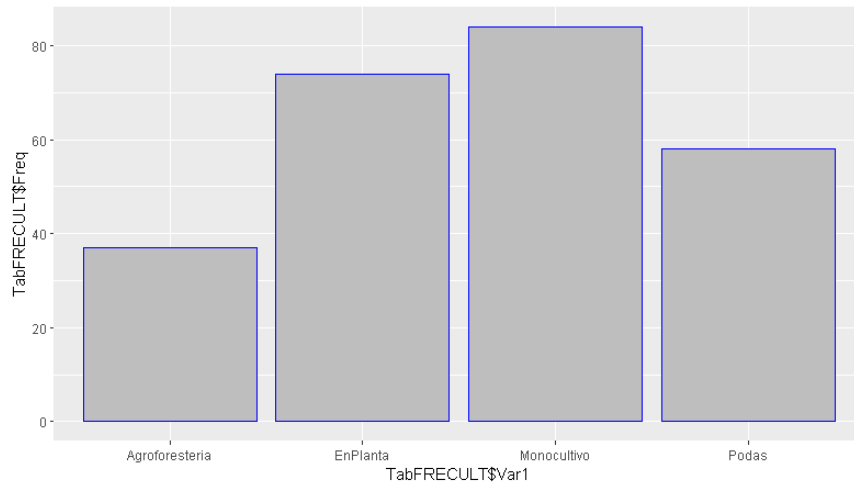
Gráfico de frecuencia Pesticida



En la figura 10, Para cada tipo de pesticida (insecticida, herbicida, fungicida), el gráfico de frecuencia mostrará con qué frecuencia se aplican estos productos en la muestra de datos, ayudando a identificar cuál de los pesticidas se utiliza con mayor o menor frecuencia.

Figura 11

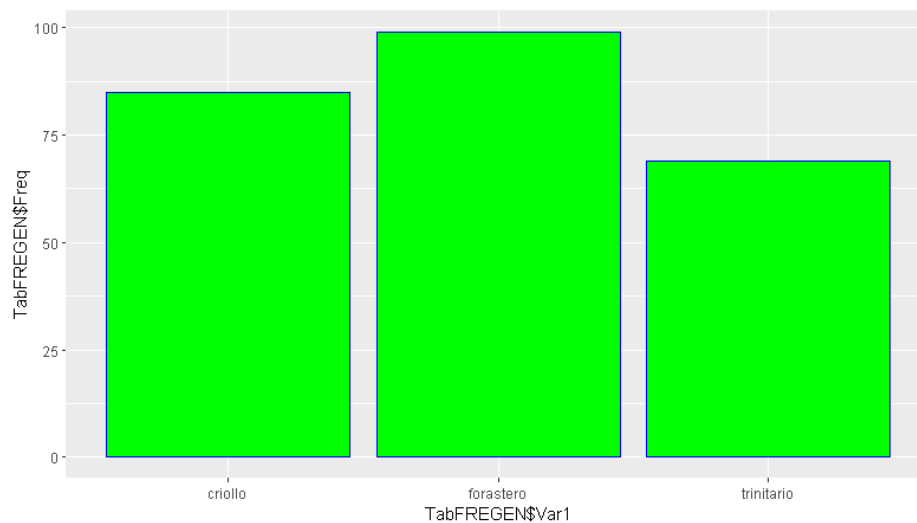
Gráfico de frecuencia cultivo



En la figura 11, El gráfico de frecuencia del método de cultivo del cacao proporciona una representación visual de la distribución de prácticas agrícolas en la muestra de datos, observando cuál es el método de cultivo más frecuentemente utilizado para identificar prácticas agrícolas dominantes en la producción de cacao, comprendiendo la diversidad de enfoques en la producción de cacao.

Figura 12

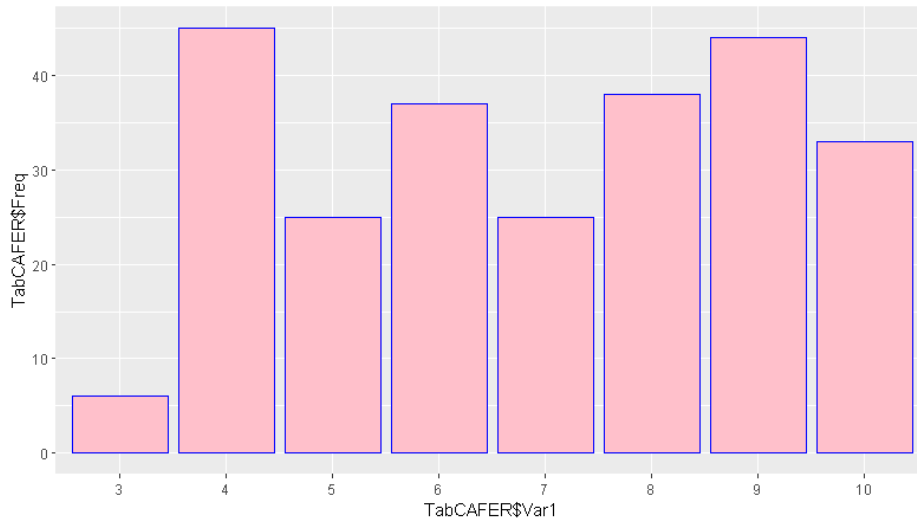
Gráfico de frecuencia Genética



En la figura 12, gráfico de frecuencia de la genética del cacao, mostrará cuántas veces se registra cada variedad genética (criollo, forastero, trinitario) en el conjunto de datos, observando la frecuencia relativa de cada tipo de genética.

Figura 13

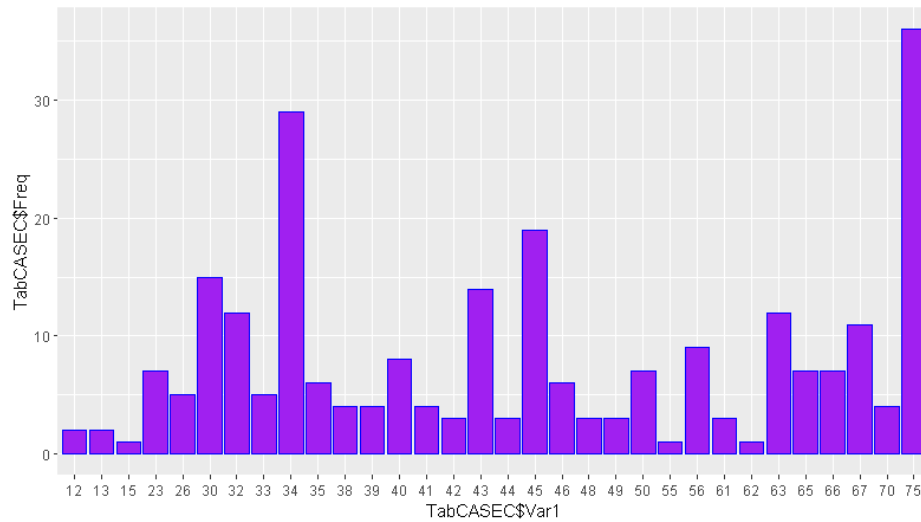
Gráfico de frecuencia Días de fermentación



En la figura 13, El gráfico de frecuencia de los días de fermentación del cacao proporciona una representación visual de la distribución de las duraciones de fermentación en la muestra de datos, la recomendación es fermentar de 3 a 10 días la masa de cacao y se evalúa por el olor y el color de los cambios externos e internos en los granos, mediciones de temperatura y pH.

Figura 14

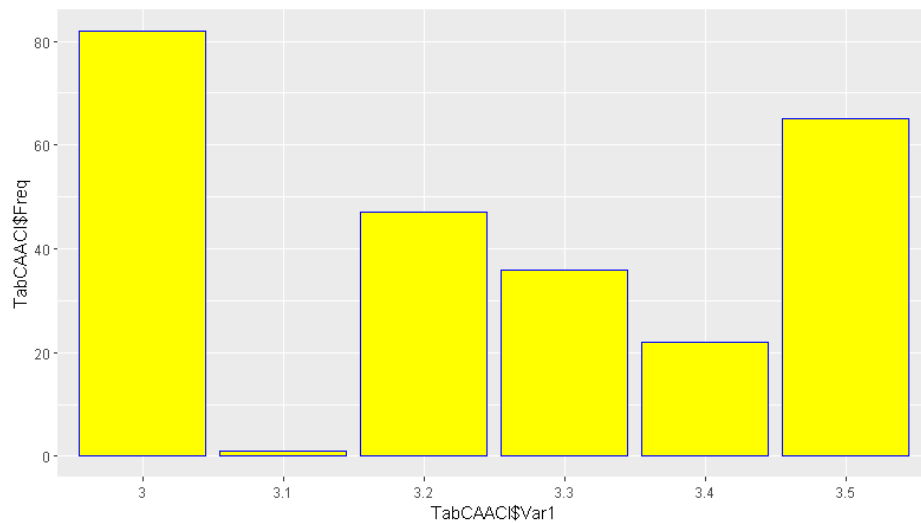
Gráfico de frecuencia Días de secado



En la figura 14, demuestra que se lo realiza por un tiempo máximo de 75 días, y así se obtiene el grano con un mejor aroma, identificando si hay correlaciones entre la duración de secado y la calidad del cacao, es decir, la duración y las condiciones de secado pueden afectar la calidad del grano y características.

Figura 15

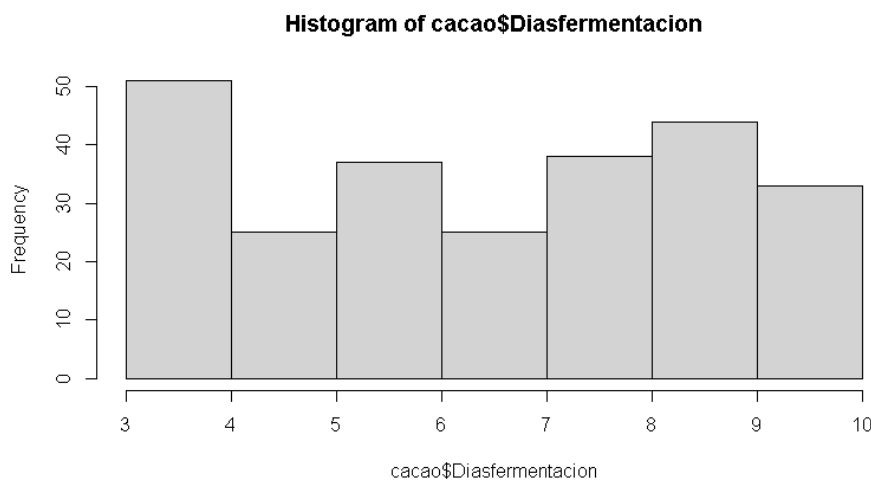
Gráfico de frecuencia Acidez



En la figura 15, El grafico de frecuencia de la acidez analiza correlaciones entre los niveles de acidez y la calidad del cacao, algunos productores pueden preferir niveles específicos de acidez, con pH entre 3.0 y 3.5, debido al contenido de ácido cítrico que contribuyan a ciertas características de sabor.

Figura 16

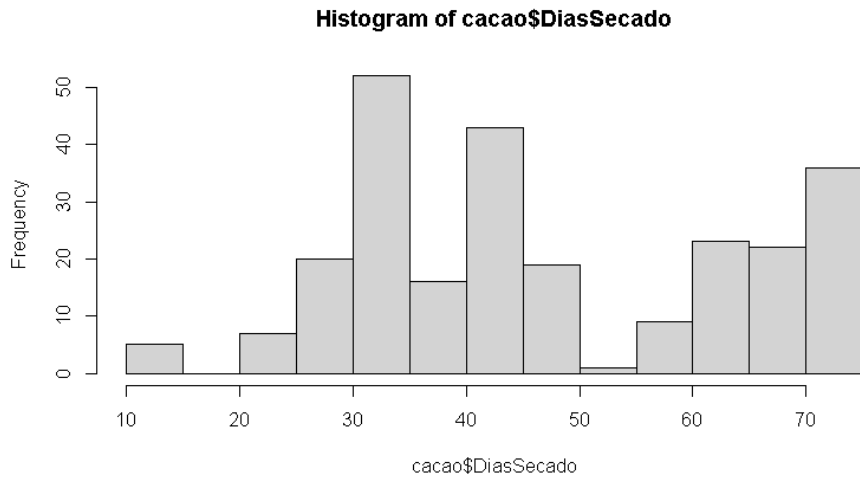
Histograma días de fermentación



En la figura 16, Representación gráfica que proporciona información sobre la distribución de la duración de la fermentación en un conjunto de datos, muestra cómo se distribuyen los diferentes períodos de fermentación en el conjunto de datos, la amplitud de las barras en el histograma representa la frecuencia relativa de cada período de fermentación, de 3 a 10 días. Esta herramienta gráfica es útil para entender patrones, variabilidad y tendencias relacionadas con el proceso de fermentación del cacao, lo que puede tener un impacto significativo en la calidad del producto final.

Figura 17

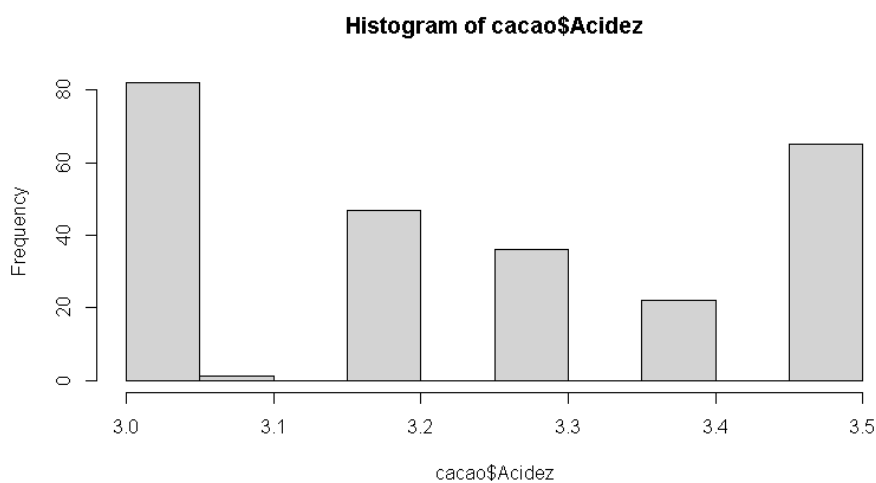
Histograma Días secado



En la figura 17, La amplitud de las barras en el histograma representa la frecuencia relativa de cada período de secado, se lo realiza por un tiempo máximo de 5 días, y así se obtiene el grano con un mejor aroma, esto es útil para comprender la variabilidad en los procesos de secado del cacao.

Figura 18

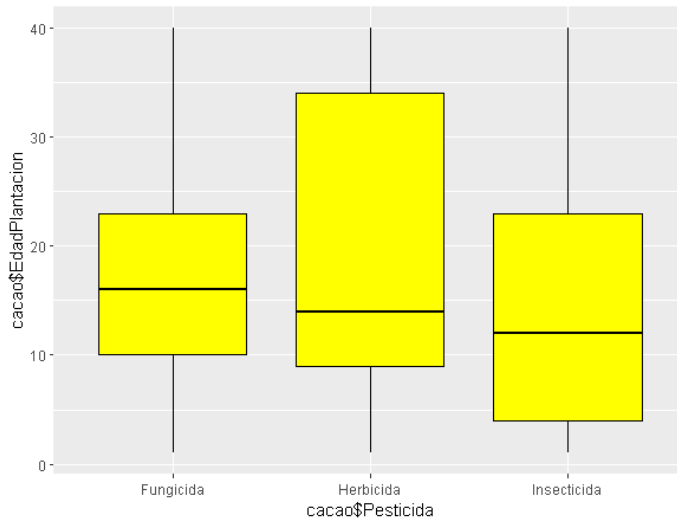
Histograma Acidez



En la figura 18, Un histograma de acidez del cacao analiza la variabilidad en los niveles de acidez y la frecuencia con la que ocurren ciertos rangos de valores, la acidez es un componente importante del sabor y aroma del chocolate.

Figura 19

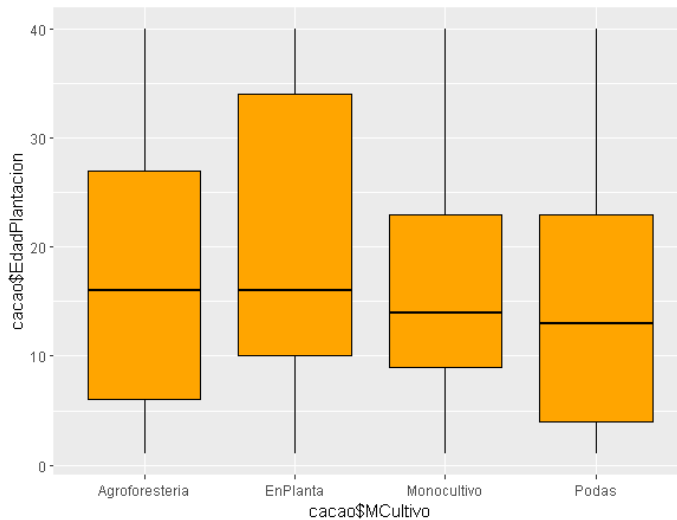
Gráfico evolutivo Edad de plantación vs Pesticida



En la figura 19, El gráfico evolutivo permite observar cómo ha evolucionado la edad de plantación del cacao a lo largo del tiempo en relación con el uso de pesticidas. Además, analizar si hay una correlación visual entre la edad de plantación de los cacaotales y la aplicación de pesticidas. Por ejemplo, observar si a medida que las plantaciones envejecen, hay un cambio en las prácticas de manejo de plagas, si hay una disminución en el uso de pesticidas a medida que las plantaciones envejecen, esto podría indicar la adopción de prácticas agrícolas más sostenibles o métodos alternativos de manejo de plagas.

Figura 20

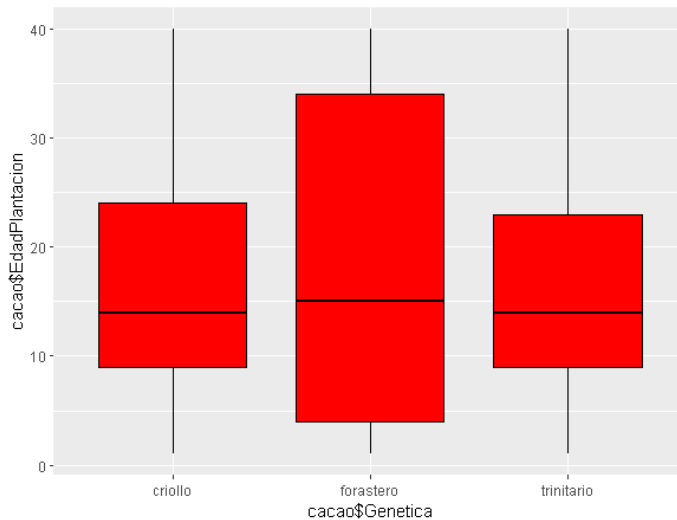
Gráfico evolutivo Edad de plantación vs Cultivo



En la figura 19, El gráfico evolutivo permite observar cómo ha cambiado la relación entre la edad de plantación del cacao y los métodos de cultivo a lo largo del tiempo, analizar si hay una correlación visual entre la edad de las plantaciones de cacao y los métodos de cultivo utilizados. Por ejemplo, observar si ciertos métodos de cultivo son más comunes en plantaciones más antiguas o recientes, si hay un aumento en la adopción de métodos de cultivo sostenibles a medida que las plantaciones envejecen, esto podría indicar una mayor conciencia ambiental y prácticas agrícolas más respetuosas con el medio ambiente. Además, evaluar si hay correlaciones entre los métodos de cultivo y factores ambientales o sostenibles, ya que esto puede tener implicaciones para la sostenibilidad y la eficiencia agrícola.

Figura 21

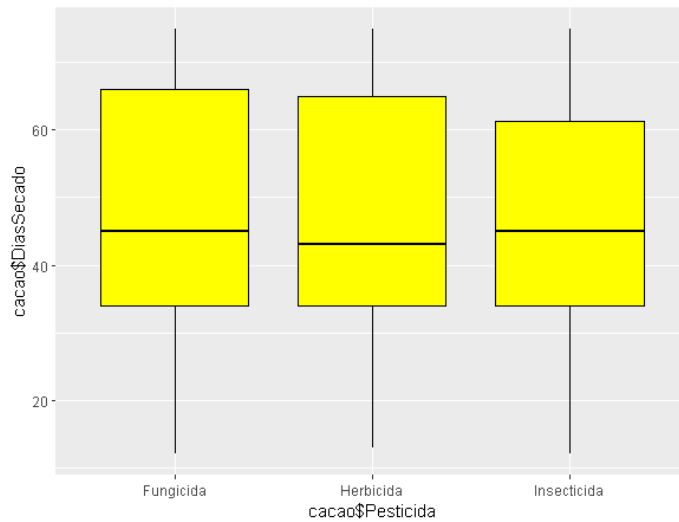
Gráfico evolutivo Edad de plantación vs genética



En la figura 21, Analizar si hay una correlación visual entre la edad de las plantaciones de cacao y las variedades genéticas presentes. Por ejemplo, observar si ciertas variedades genéticas son más comunes en plantaciones más antiguas o recientes. Observar cómo la edad de las plantaciones se relaciona con la genética del cacao proporciona información sobre cómo el envejecimiento de los cacaotales podría influir en la diversidad genética y la selección de variedades. En resumen, la interpretación de este gráfico es valiosa para comprender la evolución de la diversidad genética en las plantaciones de cacao y guiar decisiones relacionadas con la gestión de cultivos.

Figura 22

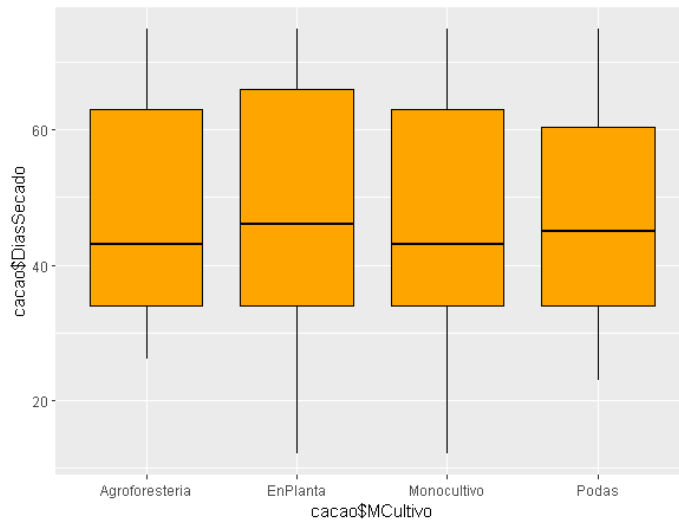
Gráfico evolutivo Días de secado vs Pesticida



En la figura 22, El gráfico evolutivo permite observar cómo ha cambiado la relación entre los días de secado del cacao y el uso de pesticidas a lo largo del tiempo, analizar si hay una correlación visual entre la duración de los días de secado y la cantidad o frecuencia de aplicación de pesticidas. Por ejemplo, observar si hay cambios en el uso de pesticidas en función de los diferentes períodos de secado.

Figura 23

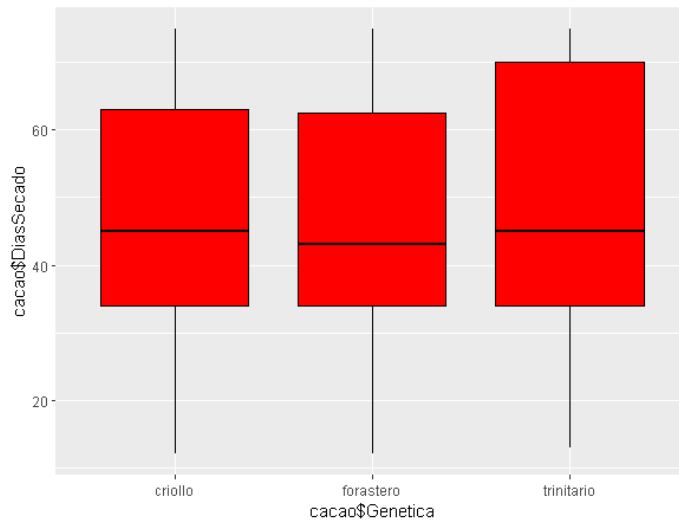
Gráfico evolutivo Días de secado vs Cultivo



En la figura 23, Un gráfico evolutivo que compara los días de secado del cacao con el cultivo de cacao puede proporcionar información valiosa sobre mejoras en eficiencia, impacto climático, calidad del producto y sostenibilidad en la industria del cacao. En primer lugar, es una disminución en los días de secado a lo largo del tiempo, podría indicar mejoras en los métodos de secado del cacao, está relacionado con la implementación de tecnologías más avanzadas, cambios en las prácticas agrícolas o mejoras en la infraestructura. Por ejemplo, una disminución en los días de secado podría sugerir prácticas agrícolas más eficientes o variedades de cacao que requieren menos tiempo de secado.

Figura 24

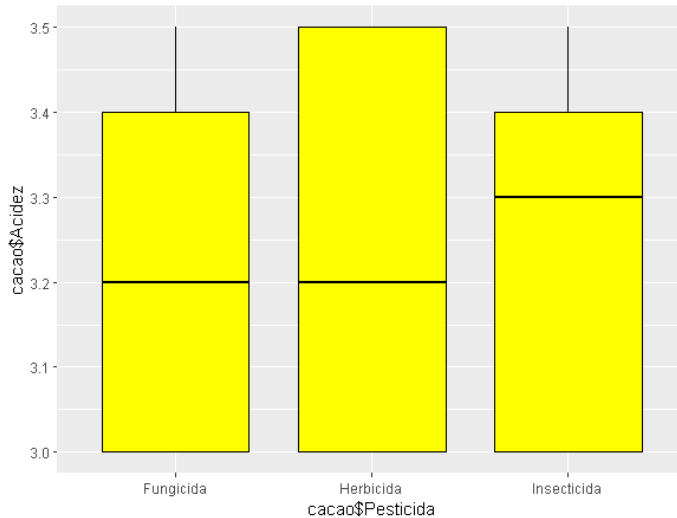
Gráfico evolutivo Días de secado vs Genética



En la figura 24, la caja criolla se encuentra más simétrica, que las demás cajas. Al igual que todos los gráficos evolutivos vistos, estas tampoco poseen valores atípicos. ciertas variedades genéticas podrían desarrollarse para tener características que faciliten un proceso de secado más rápido y eficiente. Además, la genética del cacao también puede influir en la resistencia de las plantas a enfermedades y plagas, es decir, si hay una mejora genética en este sentido, podría haber una correlación con una disminución en los días de secado, ya que las plantas más saludables pueden secarse más rápidamente. también, la genética del cacao puede afectar directamente las características físicas de los granos, el gráfico evolutivo que compara la genética con los días de secado podría indicar si ciertas características genéticas están asociadas con granos que se secan más rápidamente o que requieren un proceso de secado más prolongado. Compara los días de secado del cacao con la genética del cacao podría proporcionar información valiosa sobre la relación entre la variabilidad genética y la eficiencia del proceso de secado, así como sus posibles implicaciones en términos de calidad y resistencia.

Figura 25

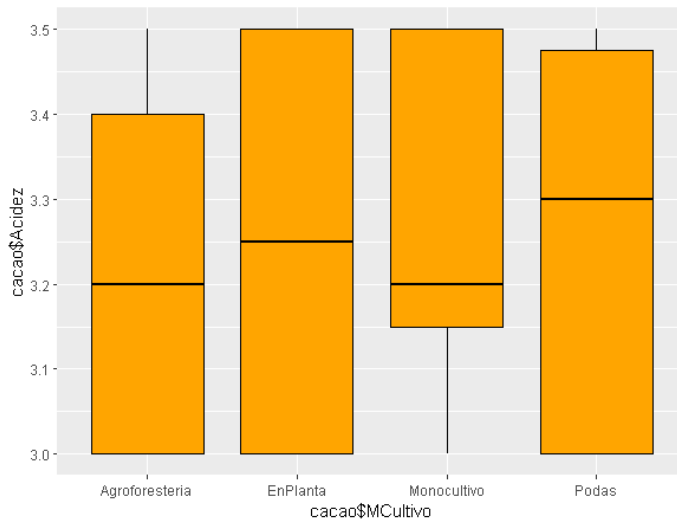
Gráfico evolutivo Acidez vs Pesticida



En la figura 25, se observa que el fungicida tiene mayor simetría a diferencias de las demás cajas. Además, estos gráficos de cajas no poseen valores atípicos. La presencia de pesticidas en el cultivo de cacao podría tener un impacto en la acidez del producto final, un aumento en la acidez podría estar relacionado con el uso de ciertos pesticidas, es decir, analizar una disminución en el uso de pesticidas a lo largo del tiempo, podría indicar un cambio hacia prácticas agrícolas más sostenibles, este cambio podría tener un impacto positivo en la acidez del cacao, ya que el uso reducido de pesticidas puede contribuir a un mejor equilibrio del suelo y a la salud general del cultivo. Sin duda, la acidez del cacao es una característica importante que afecta su calidad organoléptica, el análisis del gráfico evolutivo podría revelar si hay una correlación entre el uso de pesticidas y la calidad de la acidez del cacao. Por ejemplo, observar si la reducción del uso de pesticidas se asocia con una acidez más equilibrada y deseable.

Figura 26

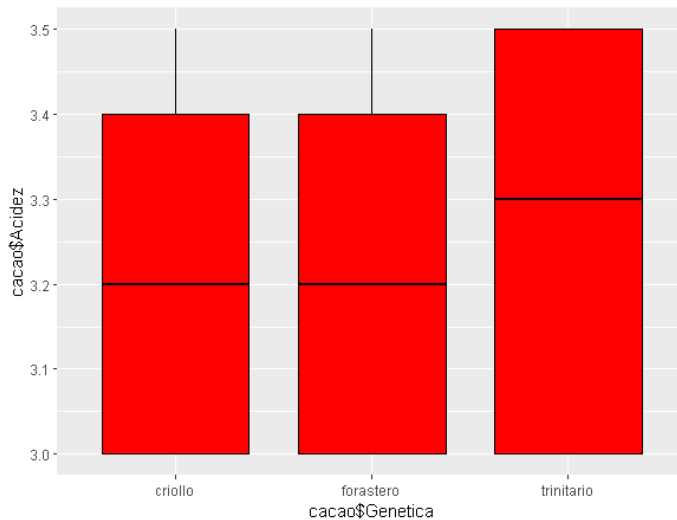
Gráfico evolutivo Acidez vs Cultivo



En la figura 26, el gráfico evolutivo trabaja con la variable acidez y método de cultivo, en la cual agroforestería tiene mayor simetría a diferencia de las demás cajas. Además, estos gráficos de cajas no poseen valores atípicos. La acidez del cacao puede variar estacionalmente debido a factores climáticos y condiciones de cultivo, el gráfico evolutivo puede identificar patrones estacionales en la acidez y cómo estos coinciden con períodos específicos de cultivo. Por ejemplo, podría observarse si hay estacionalidades en la acidez relacionadas con las condiciones de lluvia, temperatura u otros factores ambientales. Además, cambios en las prácticas de cultivo a lo largo del tiempo podrían influir en la acidez del cacao, es decir, si hay una disminución o aumento en la acidez, correlaciones con cambios en las técnicas de cultivo, como el uso de fertilizantes, métodos de poda, sistemas de riego, etc.

Figura 27

Gráfico evolutivo Acidez vs genética



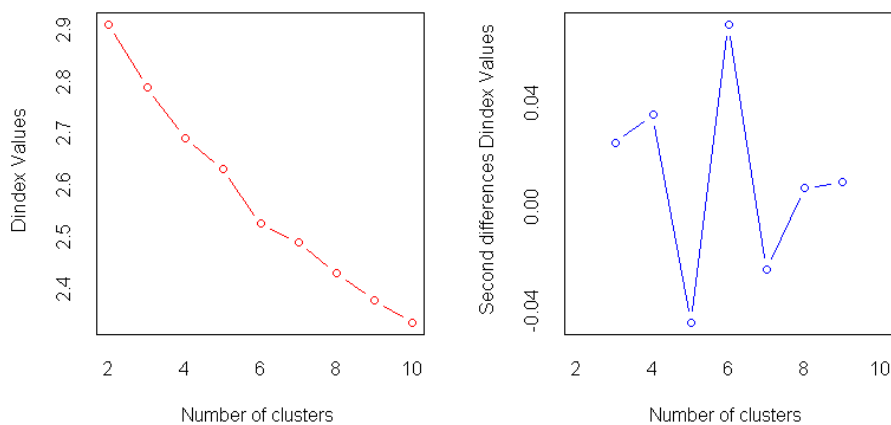
En la figura 27, Se observa un gráfico evolutivo entre la variable acidez y genética, en la cual criollo y forastero presentan simétricas y las cajas no poseen valores atípicos. un gráfico evolutivo que compara la acidez del cacao con la genética del cacao podría ser esclarecedor para entender cómo las variaciones genéticas influyen en los perfiles de sabor y la calidad del cacao. En primer lugar, diferentes variedades genéticas de cacao pueden tener perfiles de sabor únicos, incluyendo niveles de acidez, el gráfico revela cómo ciertas variaciones genéticas se asocian con cambios en la acidez del cacao. Por ejemplo, podrías observar si ciertas variedades genéticas tienden a producir cacao con perfiles de acidez más altos o bajos. Además, ciertas variedades genéticas pueden adaptarse mejor a ciertas condiciones de cultivo, lo que podría afectar la acidez del cacao; el gráfico evolutivo podría indicar si ciertas variedades genéticas se han vuelto más prevalentes debido a su capacidad para producir cacao con niveles de acidez deseados bajo ciertas condiciones de crecimiento. Por último, el gráfico podría mostrar cómo las tendencias en genética del cacao se alinean con las preferencias cambiantes de acidez en el mercado.

Una vez analizada los gráficos descriptivos, podemos entender de mejor manera los resultados arrojados por el modelo

En primer lugar, con el propósito de utilizar el paquete *clúster* cuya finalidad es agrupar variables tratando de lograr la máxima homogeneidad en cada grupo y la mayor diferencia entre los grupos, es decir, el tipo de cacao cuyas dos variedades, el cacao nacional fino de aroma y el CCN51, dos especies que van para diferentes mercados pero que tiene muy buen prestigio internacional.

Figura 28

Gráfico de clústeres

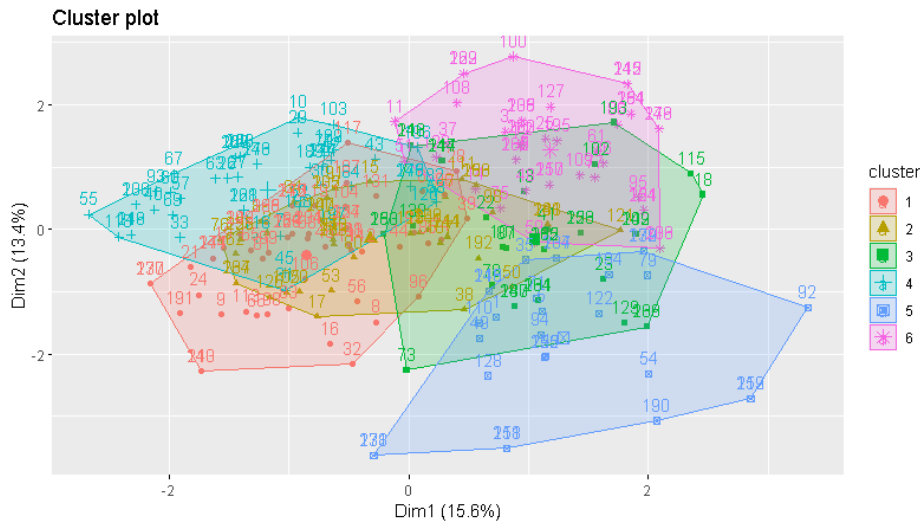


Nota. Obtenido de RStudio a partir de la codificación escrita por autor.

En la figura 28, que representa el primer gráfico de clústeres, se puede identificar el punto en el cual la disminución de la medida se estabiliza y los valores tienden a mantenerse prácticamente constantes. Este cambio en la tendencia sugiere un punto óptimo para la elección del número de clústeres. Notablemente, en este caso específico, el punto crítico se sitúa en 6 clústeres, lo que indica que este valor podría ser el más adecuado para la estructuración de los datos en grupos distintos.

Figura 29

Clúster conglomerados

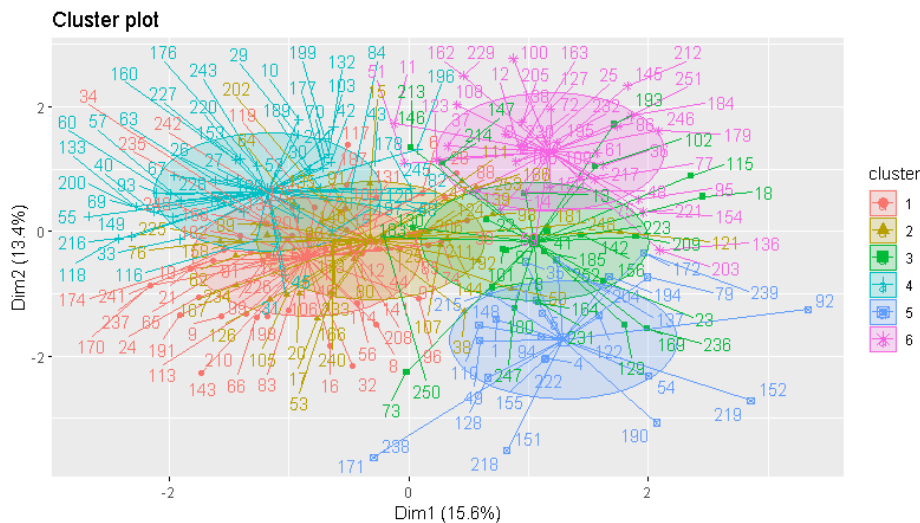


Nota. Obtenido de RStudio a partir de la codificación escrita por autor.

En este grafico obtenido mediante el algoritmo de K-means y el código `fviz_clúster`, se presenta la asignación de cada dato a un clúster específico. Estos clústeres se forman en base a similitudes en las características y requisitos de los datos.

Figura 30

Gráfico elíptico de tipo euclidiana

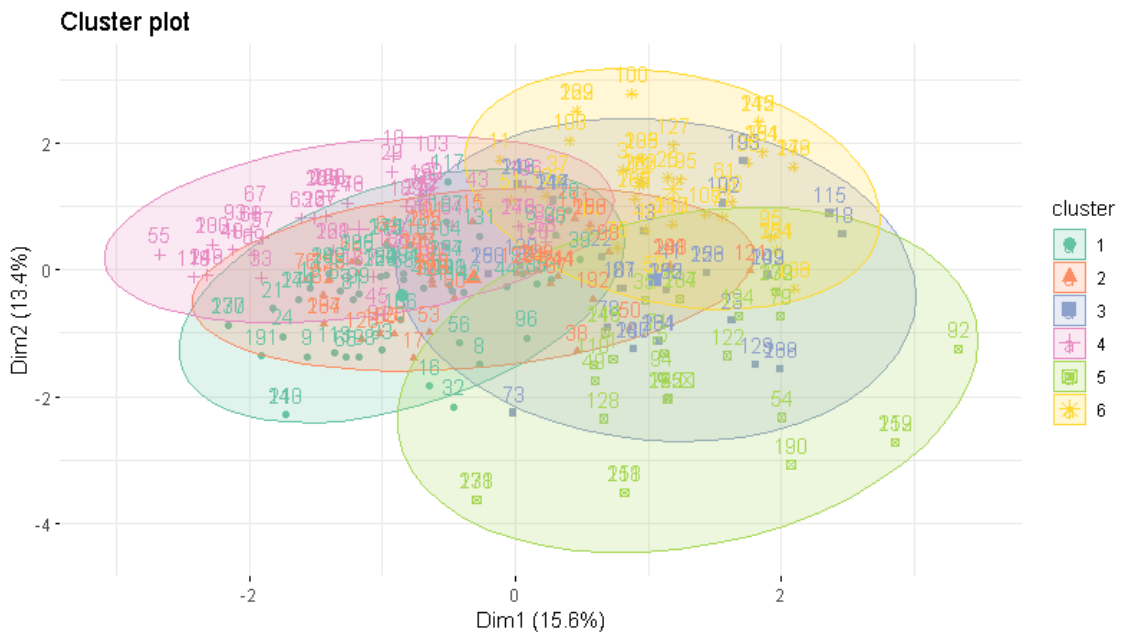


Nota. Obtenido de RStudio a partir de la codificación escrita por autor.

En la figura 30, los datos se escapan de los clústeres, dando como resultado más distinciones de tipos de cacao, es una representación visual de datos que muestra cómo se agrupan o dividen los puntos de datos en grupos distintos. Como puede comprenderse fácilmente el análisis clúster tiene una extraordinaria importancia y eficiencia para la clasificación.

Figura 31

Gráfico de aplicación de *k-means* en los clústeres



Nota. Obtenido de RStudio a partir de la codificación escrita por autor.

En la figura 31, se convergen en una sola y la mayoría del tipo de cacao tienen algo en común, dando como resultado muy poco las representaciones de cacao que salen de los clústeres, es decir, hay algunas instancias que no encajan tan claramente en estos clústeres, posiblemente debido a características distintivas o atípicos.

Tabla 3*Nivel de importancia*

	MeanDecreaseGini
HectareasNA	23.795.356
HectareasCCN51	22.101.604
EdadPlantacion	24.877.272
Lotes	10.423.500
Diasfermentacion	9.685.832
DiasSecado	15.161.867
Pesticida	6.165.658
MCultivo	12.371.632
Genetica	21.100.642
Acidez	22.447.701

En la tabla 3, se denota la importancia relativa de las variables en el modelo de clasificación RandomForest, dando como resultado la importancia de la contribución de cada variable a reducir la impureza en los nodos del modelo de árbol de decisión, sin duda, la edad de plantación es un factor clave para el modelo, con una importancia de aproximadamente 24,877, por otra parte, pesticida aporta con una importancia relativa de alrededor de 6,2 siendo menos influyente en comparación con otras variables.

Tabla 4*Matriz de Confusión*

Actual	Predicho					
	Exportación 1	Nacional	Exportación 3	Exportación 4	Exportación 5	Exportación 6
Exportación 1	55	0	1	0	0	0
Nacional	0	43	0	0	0	1
Exportación 3	0	0	32	0	0	0
Exportación 4	0	0	0	50	0	0
Exportación 5	0	0	0	0	29	0
Exportación 6	0	0	0	0	0	42

Nota. Obtenido de RStudio a partir de la codificación escrita por autor.

La finalidad de la matriz de confusión es proporcionar una visión detallada del rendimiento del modelo de clasificación RandomForest. A partir de los valores en la matriz, se pueden calcular

diversas métricas de evaluación, como la precisión, la sensibilidad, la especificidad y la tasa de error, entre otras. Se puede observar, un solo error dentro de las predicciones.

Tabla 5

Predicciones a través de K-means y RandomForest

K_Means	Prediccion	RandProb[, "Exportacion 1"]	RandProb[, "Nacional"]	RandProb[, "Exportacion 3"]	RandProb[, "Exportacion 4"]	RandProb[, "Exportacion 5"]	RandProb[, "Exportacion 6"]
Exportacion 5	Exportacion 5	0.028666667	0.179000000	0.058666667	0.042333333	0.663000000	0.028333333
Nacional	Nacional	0.297000000	0.582333333	0.005666667	0.070000000	0.022000000	0.023000000
Exportacion 6	Exportacion 6	0.005000000	0.025666667	0.009000000	0.067333333	0.006333333	0.886666667
Exportacion 5	Exportacion 5	0.135666667	0.022333333	0.064000000	0.001000000	0.755666667	0.021333333
Exportacion 5	Exportacion 5	0.005000000	0.041333333	0.071333333	0.015333333	0.861333333	0.005666667
Exportacion 1	Exportacion 1	0.530666667	0.054333333	0.022000000	0.148333333	0.021000000	0.223666667
Exportacion 4	Exportacion 4	0.112333333	0.020666667	0.020000000	0.833333333	0.007666667	0.006000000
Exportacion 1	Exportacion 1	0.916333333	0.057666667	0.000666667	0.000666667	0.013333333	0.011333333
Exportacion 1	Exportacion 1	0.801000000	0.076333333	0.014333333	0.096666667	0.005666667	0.006000000
Exportacion 4	Exportacion 4	0.023333333	0.024000000	0.061666667	0.883000000	0.000333333	0.007666667
Exportacion 6	Exportacion 6	0.313000000	0.024000000	0.023000000	0.089000000	0.011333333	0.539666667
Exportacion 6	Exportacion 6	0.011666667	0.103000000	0.096333333	0.035000000	0.003000000	0.751000000
Exportacion 3	Exportacion 3	0.002333333	0.009666667	0.933000000	0.008000000	0.041000000	0.006000000
Exportacion 6	Exportacion 6	0.076333333	0.035666667	0.039666667	0.027333333	0.005666667	0.815333333
Nacional	Nacional	0.015333333	0.783333333	0.006000000	0.152333333	0.003666667	0.039333333
Exportacion 1	Exportacion 1	0.875666667	0.077000000	0.000666667	0.004000000	0.020000000	0.022666667
Nacional	Nacional	0.032333333	0.926666667	0.009666667	0.019666667	0.010333333	0.001333333
Exportacion 3	Exportacion 3	0.016666667	0.003666667	0.706000000	0.065333333	0.187666667	0.020666667
Exportacion 1	Exportacion 1	0.850333333	0.067333333	0.000000000	0.066666667	0.007333333	0.008333333

Nota. Obtenido de RStudio a partir de la codificación escrita por autor. Breve visualización de los resultados.

Como se observa en la tabla no solamente da valores de quien es de exportación y quien es nacional pero también da información sobre los diferentes grupos de exportación y se puede encontrar un modelo de negocio mejor al que se tiene.

En el marco de nuestro proyecto, es fundamental asegurar la integridad de los resultados y la meticulosidad de las metodologías aplicadas. Por ende, explicaremos brevemente un artículo científico, donde se implementó el algoritmo Random Forest para la clasificación del cacao. Este empeño busca garantizar no solo la exactitud d ellos resultados obtenidos, sino también las transparencias y confiabilidad de cada paso metodológico a lo largo de la investigación.

El artículo científico titulado “Clasificación de granos de cacao fermentado (Prueba de corte) mediante visión por ordenador, escrito por Marciano M. Oliveria, Breno V. Cerqueira, Sylvio

Barbon Jr., Douglas F. Barbin (2021), aborda la fermentación de los granos de cacao como un paso crucial en la fabricación de chocolate, dado su impacto en el desarrollo del sabor y en componentes fundamentales como aminoácidos libres, péptidos y azúcares. La determinación del grano fermentado se lleva a cabo mediante la inspección visual de cambios en el color interno y la textura de los granos, a través de la prueba de corte. Aunque este método se considera estándar para evaluar la fermentación en granos de cacao, su limitación radica en su lentitud y dependencias de personal especializado (p. 1).

La captura de imágenes y su análisis generan características elaboradas a mano a partir de ellos granos, las cuales se emplearon como predictores en bosque aleatorios para la clasificación de muestras. Un total de 1800 granos se clasificaron en cuatro grados de fermentación. Respecto a todas las características de la imagen, se obtuvo una precisión del 0.93 en la validación del conjunto de datos desequilibrado con una precisión de 0.85 y una recuperación de 0.81. A pesar de que el conjunto de datos desequilibrado refleja la variación real de la fermentación, se probó el método con un conjunto de datos equilibrados para explorar la influencia de un menor número de muestras por clase, logrando 0.92 hasta 0.90 en precisión y recuperación, respectivamente. La técnica podría evolucionar hacia una aplicación industrial con un marco de integración adecuado, sustituyendo el método tradicional para clasificar los granos de cacao fermentados (M. Oliveira, V. Cerqueira, Barbon Jr., & F. Barbin, 2021).

Conclusiones

Después de realizar una revisión literaria exhaustiva, se evidencia la complejidad y la amplitud de los conceptos relacionados con el machine learning. Se ha logrado una comprensión profunda de las diversas dimensiones que abarca el estudio de la calidad del aroma del cacao fino, desde factores agronómicos hasta aspectos sensoriales. Esta revisión proporciona una base sólida para el desarrollo del proyecto, permitiendo una contextualización adecuada y una identificación de lagunas de conocimiento que pueden abordarse.

Debido a las normas de los requisitos de calidad para los granos de cacao y los criterios para su clasificación para ser considerado materia prima apta para elaboración de derivados de cacao destinados al consumo humano y fines de comercialización; nace la necesidad de la exportadora Cocholato S.A aplicar algoritmos de machine learning para la clasificación de cacao de fino aroma, utilizando los modelos clasificatorios RandomForest, K-means y Mean Decrease in Gini, presentan capacidades significativas para la identificación precisa de las características que determinan la calidad del cacao. Este análisis proporciona una base sólida para la selección del modelo más adecuado para el proyecto, considerando factores como la complejidad del modelo y la disponibilidad de datos.

La aplicación del algoritmo de machine learning seleccionado demuestra una eficacia considerable en la clasificación de la calidad del cacao fino aroma. Los resultados obtenidos indican una capacidad satisfactoria para discernir entre distintas categorías de calidad, lo que sugiere un potencial real para la implementación práctica de esta tecnología en la industria cacaotera. Sin embargo, es esencial tener en cuenta la necesidad continua de ajustes y optimizaciones, así como la importancia de la validación y la interpretación de los resultados para garantizar la confiabilidad de la clasificación.

Recomendaciones

Para la exportadora Cocholato S.A la mejora de la calidad del cacao es un proceso continuo que requiere un compromiso a largo plazo y la colaboración de todos los actores involucrados en la cadena de suministro, desde la plantación hasta la exportación. Con referencia a, las buenas prácticas agrícolas, técnicas de fermentación y secado, y métodos de gestión de calidad se recomienda formación y capacitación continua a los agricultores y trabajadores involucrados en la producción de cacao, es decir, establecer una colaboración estrecha con los agricultores locales, proporcionándoles apoyo técnico y financiero para implementar prácticas de alta calidad. En efecto, con la presente investigación se busca adoptar nuevas prácticas innovadoras en la producción y procesamiento de cacao que puedan mejorar la calidad del producto final, implementando tecnologías avanzadas para monitorear tales como: sensores de humedad, sistemas de trazabilidad, y otros dispositivos que ayuden a garantizar un control de calidad más preciso.

Analizando detenidamente los resultados arrojados por el modelo de Random Forest, sugerimos a la empresa Cocholate que explore la posibilidad de incorporar, a largo plazo, estrategias basadas en aprendizaje automático. Esta iniciativa podría resultar altamente beneficiosa al mejorar de manera significativa la capacidad de clasificación de la calidad de sus granos de cacao. Al lograr una mayor precisión en esta evaluación, la empresa estaría en condiciones de fijar precios más ajustados y competitivos, lo que contribuiría a consolidar su posición en el mercado. La adopción de tecnología avanzadas no solo optimizaría la gestión interna de la calidad, sino también diferenciar a Cocholate de sus competidores, consolidando su reputación como productor de chocolate de alta calidad.

Referencias Bibliográficas

- (CEPAL), C. E. (2021). *Tecnologías digitales para un nuevo futuro*. Santiago. Obtenido de https://repositorio.cepal.org/bitstream/handle/11362/46816/1/S2000961_es.pdf
- Abdala, M. A. (30 de Marzo de 2020). *RPubs*. Obtenido de <https://rpubs.com/arquez9512/592295>
- Aguilar, H. (2016). *Manual para la Evaluación de la Calidad del Grano de Cacao*. Obtenido de FUNDACIÓN HONDUREÑA DE INVESTIGACIÓN AGRÍCOLA.
- Alcívar Córdova, K., Quezada Campoverde, J. M., Barrezuela Unda, S., Garzón Montealegre, V. J., & Carvaja Romero, H. (2021). Análisis económico de la exportación del cacao en el Ecuador durante el periodo 2014 – 2019. *Polo del Conocimiento*, 6(3), 2430-2444. doi:10.23857/pc.v6i3.2522
- Araneda, P. (27 de Abril de 2021). *Rpubs*. Obtenido de <https://rpubs.com/paraneda/tidyverse>
- Arias, A. S. (2020). Rentabilidad. *Economipedia*.
- Ariza-López, F. J., Rodríguez-Avi, J., & Alba-Fernández, V. (2018). Control estricto de matrices de confusión por medio de distribuciones multinomiales. *Geofocus. Revista Internacional de Ciencia y Tecnología de la Información Geográfica*(21), 215-226. doi:<http://dx.doi.org/10.21138/GF.591>
- Auguie, B. (2017). *R-Project.org*. Obtenido de <https://cran.r-project.org/web/packages/gridExtra/index.html>
- Bajaña Morán, L. A., & Paladines Gallardo, J. A. (2021). *Repositorio UPS*. Obtenido de <chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://dspace.ups.edu.ec/bitstream/123456789/21405/1/UPS-GT003527.pdf>
- Ballari, D. (2018). *RPubs*. Obtenido de <https://rpubs.com/daniballari/ggplot>
- Baviera, T. (2016). Técnicas para el Análisis de Sentimiento en Twitter: Aprendizaje Automático Supervisado y SentiStrength. *Revista Dígitos*, 1(3), 33-50. Obtenido de <https://revistadigitos.com/index.php/digitos/article/view/74/39>
- Berrendero, J. R. (11 de Agosto de 2017). *RPubs*. Obtenido de <https://rpubs.com/joser/caret>
- Berrendero, J. R. (s.f.). Introducción a CARET. *Rpubs*.
- Camacho Rosero, M. P., & Reyna Muentes, C. (21 de Febrero de 2022). *Repositorio UCSG*. Obtenido de <chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/http://repositorio.ucsg.edu.ec/bitstream/3317/18340/1/T-UCSG-PRE-CEAC-CNI-24.pdf>

- CEFA. (2023). *CEFA Ecuador, La semilla de la solidaridad*. Obtenido de <https://cefaecuador.org/productos/cacao/>
- CEPAL. (2023). *Índice Latinoamericano de inteligencia artificial*. Obtenido de <https://indicelatam.cl/fichLa%20tecnolog%C3%ADa%20ha%20emergido%20como%20una%20herramienta%20fundamental%20en%20la%20sociedad%20moderna,%20transformando%20completamente%20nuestra%20vida%20cotidiana,%20entorno%20laboral%20y%20relaciones%20interpersonales.%>
- Chamorro, K., Laza, N., Noriega, H., Rojano, J., & Heredia, D. (2021). Aplicación de Machine Learning para análisis de los fenómenos de violencia intrafamiliar en el departamento del Atlántico. *Investigación y Desarrollo en TIC*, 12(1), 1-12. Obtenido de <https://bonga.unisimon.edu.co/handle/20.500.12442/8101>
- Corporación Financiera Nacional. (2022). *Ficha Sectorial Cacao y Chocolate*. Obtenido de <chrome-extension://efaidnbmninnibpcajpegclclefindmkaj/https://www.cfn.fin.ec/wp-content/uploads/downloads/biblioteca/2022/fichas-sectoriales-2-trimestre/Ficha-Sectorial-Cacao.pdf>
- Corvalán, J. G. (2018). Inteligencia artificial: retos, desafíos y oportunidades - Prometea: la primera inteligencia artificial de Latinoamérica al servicio de la Justicia. *Revista de Investigações Constitucionais*, 5(1), 295-316. doi:<https://doi.org/10.5380/rinc.v5i1.55334>
- Franco, P. T. (7 de Julio de 2023). Día Mundial del Cacao: en Ecuador se espera que efectos de El Niño no se sientan en 2023 para poder romper récord de exportación. Obtenido de <https://www.eluniverso.com/noticias/economia/dia-mundial-del-cacao-en-ecuador-se-espera-que-efectos-de-el-nino-no-se-sientan-en-2023-para-poder-romper-record-de-exportacion-nota/>
- Galán, J. S. (2020). Economipedia.
- Gobierno de la República del Ecuador. (2015). *Cultura y Patrimonio*. Obtenido de <https://www.culturaypatrimonio.gob.ec/el-intenso-color-del-cacao-ecuadoriano/>
- González Rosado, L. L., Moreira Sornoza, W. G., & Dueñas Rivadeneira, A. A. (2022). La cadena de comercialización del cacao fino de aroma, cantón Pichincha, Ecuador. *ECA Sinergia*, 13(3), 86-95. doi:<https://doi.org/10.33936/ecasinergia.v13i3.4689>
- Haan, K. (25 de Abril de 2023). *Forbes Advisor*. Obtenido de <https://www.forbes.com/advisor/business/ai-statistics/>

- IBM. (2022). *IBM Global AI Adoption Index 2022*. Nueva York. Obtenido de <https://www.ibm.com/downloads/cas/GVAGA3JP>
- IBM. (2023). *¿Qué es el aprendizaje supervisado?* Obtenido de <https://www.ibm.com/es-es/topics/supervised-learning>
- IBM. (2023). *¿Qué es el random forest?* Obtenido de <https://www.ibm.com/mx-es/topics/random-forest>
- IBM. (2023). *IBM*. Obtenido de <https://www.ibm.com/es-es/topics/unsupervised-learning>
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2023). *An Introduction to Statistical Learning: with applications in R*. New York: Springer New York Heidelberg Dordrecht London. doi:10.1007/978-1-4614-7138-7
- Jurafsky, D., & H. Martin, J. (2008). *Speech and Language Processing*. Prentice Hall. Obtenido de <https://web.stanford.edu/~jurafsky/slp3/5.pdf>
- León, C. R. (26 de Febrero de 2020). *Repositorio UCSG*. Obtenido de <http://repositorio.ucsg.edu.ec/handle/3317/14343?mode=full>
- M. Oliveira, M., V. Cerqueira, B., Barbon Jr., S., & F. Barbin, D. (2021). Classification of fermented cocoa beans (cut test) using computer vision. *Journal of Food Composition and Analysis*, 97. doi:<https://doi.org/10.1016/j.jfca.2020.103771>
- Medina Merino, R. F., & Ñique Chacón, C. i. (2017). Bosques aleatorios como extensión de los árboles de clasificación con los programas R y Python. *Interfases*(10), 165-189. Obtenido de <https://dialnet.unirioja.es/servlet/articulo?codigo=6230447>
- Munárriz, L. Á. (1994). *Fundamentos de Inteligencia Artificial* (2 ed.). Murcia. Obtenido de [https://www.google.com.ec/books/edition/Fundamentos_de_inteligencia_artificial/UfccXvwzIOUC?hl=es-419&gbpv=1&dq=Mun%C3%A1rriz+\(1994\)+La+inteligencia+artificial+es+una+ciencia+que+tiene+como+objetivo+el+dise%C3%B1o+y+construcci%C3%B3n+de+m%C3%A1quinas+capa](https://www.google.com.ec/books/edition/Fundamentos_de_inteligencia_artificial/UfccXvwzIOUC?hl=es-419&gbpv=1&dq=Mun%C3%A1rriz+(1994)+La+inteligencia+artificial+es+una+ciencia+que+tiene+como+objetivo+el+dise%C3%B1o+y+construcci%C3%B3n+de+m%C3%A1quinas+capa)
- Murphy, K. P. (2012). *Machine Learning: A probabilistic Perspective*. Massachusetts Institute of Technology. Obtenido de chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/http://noiselab.ucsd.edu/ECE228/Murphy_Machine_Learning.pdf
- Oca, J. M. (2020). Exportaciones . *Economipedia* .

- Organization), I. (. (2022). Obtenido de <https://www.icco.org/>
- Pérez González, Á. R., Villegas Estévez, C. J., Cabascango Jaramillo, J. C., & Soria Flores, E. R. (2023). Inteligencia artificial como estrategia de innovación en empresas de servicios: Una revisión bibliográfica. *Revista Publicando*, 10(38), 74 - 82.
doi:<https://doi.org/10.51528/rp.vol10.id2359>
- Perugachi, M. (2017). *Republica del Cacao*. Obtenido de <https://republicadelcacao.com/es/blogs/news/certifications-in-the-food-industry>
- Pilay, M. F. (2020). *Repositorio Universidad de Guayaquil*. Obtenido de <http://repositorio.ug.edu.ec/handle/redug/61827?mode=full>
- Plazas, J. E., López, I. D., & Corrales, J. C. (07 de July de 2017). A Tool for Classification of Cacao Production in Colombia Based on Multiple Classifier Systems. *Springer*, 10405. doi:10.1007/978-3-319-62395-5_5
- Portillo, E., Graziani De Farinas, L., & Betancourt, E. (2005). Efecto de los tratamiento post-cosecha sobre la temperatura y el indice de fermentación en la calidad del cacao criollo porcela (*Theobroma cacao* L.) en el Sur del Lago Macaraibo. *Scielo*, 22(4), 3394-406. Obtenido de https://ve.scielo.org/scielo.php?pid=S0378-78182005000400007&script=sci_arttext
- RDocumentation*. (2014). Obtenido de <https://www.rdocumentation.org/packages/NbClust/versions/3.0.1/topics/NbClust>
- Roldan, P. N. (2020). Cadena de suministro . *Economipedia* .
- Romero, Ramirez, Rojas, & Darghan. (27 de Junio de 2020). *Rpubs*. Obtenido de <https://rpubs.com/lhromeroj/analisisdeclusterR>
- Rouhiainen, L. (2018). *Inteligencia artificial 101 cosas que debes saber hoy sobre nuestro futuro*. Barcelona: Editorial Planeta. Recuperado el 2023, de <https://www.planetadelibros.com/libro-inteligencia-artificial/280581>
- R-Project*. (2020). Obtenido de <https://cran.r-project.org/web/packages/factoextra/index.html>
- Rueda, A., Orozco, O., Petit, V., Villamizar, J., & Lizarazo, S. (08 de Junio de 2022). *Rpubs*. Obtenido de https://rpubs.com/andresss20/codigo_pca
- Russell, S., & Norvig, P. (2004). *Inteligencia Artificial: Un Enfoque Moderno*. Madrid: PEARSON EDUCACIÓN S.A. Obtenido de

<https://luismejias21.files.wordpress.com/2017/09/inteligencia-artificial-un-enfoque-moderno-stuart-j-russell.pdf>

Servicio de acreditacion ecuatoriano . (21 de 03 de 2018). Obtenido de

<https://www.acreditacion.gob.ec/cacao-en-ecuador-mejora-su-calidad/>

Universidad Europea. (07 de 06 de 2022). Obtenido de

<https://universidadeuropea.com/blog/aprendizaje-supervisado-no-supervisado/>

Vite Cevallo, H., Carvajal Romero, H., & Barrezueta Unda, S. (2022). Aplicación de algoritmos

de aprendizaje automático para clasificar la fertilidad de un suelo bananero. *Revista*

Universidad y Sociedad, 29-33. Obtenido de

http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S2218-36202022000500029

Zendesk. (2022). Perfil de cliente para una empresa: crea el tuyo en 3 pasos.

Anexos

Anexo 1

Script Completo

```
getwd()
setwd("C:/Users/angel/OneDrive/Escritorio/TESIS")

cacao<-read.csv2("DATOS/CACAOMYA.csv")

#INSTALACIÓN DE PAQUETES
install.packages("caret")
install.packages("gridExtra")
install.packages("ggplot2")
install.packages("randomForest")
install.packages("tidyverse")
install.packages("factoextra")
install.packages("cluster")
install.packages("NbClust")
install.packages("GGally")
library(GGally)
library(caret)
library(gridExtra)
library(ggplot2)
library(randomForest)
library(tidyverse)
library(cluster)
library(factoextra)
library(NbClust)

#K MEDIAS PARA SEPARAR ENTRE EXPORTACION Y NO EXPORTACION

#ESCALAMOS
SKcacao<- scale(cacao)

#ESTIMAR EL NUMERO DE CLÚSTERS
NumClust<-NbClust(SKcacao, distance = "euclidean",
                  min.nc=2, max.nc=10, method = "kmeans",
                  index = "alllong")
fviz_nbclust(NumClust)
```

```

#CALCULAMOS LOS DOS CLÚSTERS
k6 <- kmeans(SKcacao, centers = 6, nstart = 25)
k6
str(k6)

#PLOTAR LOS CLÚSTER
fviz_cluster(k6, data = SKcacao)
fviz_cluster(k6, data = SKcacao, ellipse.type = "euclid",
             repel = TRUE, star.plot = TRUE) #ellipse.type= "t", "norm", "euclid"
fviz_cluster(k6, data = SKcacao, ellipse.type = "norm")
fviz_cluster(k6, data = SKcacao, ellipse.type = "norm",
             palette = "Set2", ggtheme = theme_minimal())

#PASAR CLUSTER A LA BASE
cacao$K_Means<-as.factor(k6$cluster)
table(cacao$K_Means)

#CODIFICO LA VARIABLE DE RESPUESTA
cacao$K_Means <- factor(cacao$K_Means,
                      levels = c(1,2,3,4,5,6),
                      labels = c("Exportacion 1","Nacional",
                                  "Exportacion 3","Exportacion 4",
                                  "Exportacion 5","Exportacion 6"))

cacao$Pesticida <- factor(cacao$Pesticida,
                        levels = c(1,2,3),
                        labels = c("Fungicida","Herbicida","Insecticida"))

cacao$MCultivo <- factor(cacao$MCultivo,
                       levels = c(1,2,3,4),
                       labels = c("Agroforesteria","EnPlanta",
                                   "Monocultivo","Podas"))

cacao$Genetica <- factor(cacao$Genetica,
                       levels = c(1,2,3),
                       labels = c("criollo","forastero","trinitario"))

#CREAMOS UN ARBOL
#para un arbol o cualquier derivado de un arbol de decision, se planta una semilla
set.seed(123)

#ENTRENAMOS LOS DATOS
dataentrenamiento <- createDataPartition(cacao$K_Means, p=0.80, list = FALSE)

#ARMAMOS EL BOSQUE ALEATORIO
RandomTreeModel <- randomForest(x=cacao[dataentrenamiento, 1:10],
                               y=cacao[dataentrenamiento,11],
                               ntree = 3000, keep.forest = TRUE)

RandomTreeModel$importance

#REALIZO LA PREDICCION
Prediccion <- predict(RandomTreeModel,cacao)
Prediccion

#ARMO LA MATRIZ DE CONFUSION
Matriz <- table(cacao$K_Means, Prediccion, dnn = c("Actual", "Predicho"))
Matriz

#VALORES CLASIFICADOS
cacao$Prediccion <- predict(RandomTreeModel,cacao)

#PROBABILIDAD DE EXITO
cacao$RandProb <- predict(RandomTreeModel,cacao, type = "prob")

```

```

#ANALISIS EXPLORATORIOS DE DATOS
#EXPLORANDO VARIABLES
summary(cacao$Diasfermentacion)
summary(cacao$DiasSecado)
summary(cacao$Pesticida)
summary(cacao$MCultivo)
summary(cacao$Genetica)
summary(cacao$Acidez)

#GRAFICO DE NORMALIDAD Y COMBINADO
C.CA=ggplot(cacao,
            aes(x=cacao$Diasfermentacion,fill=cacao$Pesticida))+
  geom_histogram(aes(y=..density..),colour="black")+
  geom_density(alpha=.3,fill="green")
C.CA

C.CA2=ggplot(cacao,
            aes(x=cacao$DiasSecado,fill=cacao$Genetica))+
  geom_histogram(aes(y=..density..),colour="purple")+
  geom_density(alpha=.3,fill="pink")
C.CA2

C.CA3=ggplot(cacao,
            aes(x=cacao$EdadPlantacion,fill=cacao$MCultivo))+
  geom_histogram(aes(y=..density..),colour="brown")+
  geom_density(alpha=.3,fill="orange")
C.CA3

grid.arrange(C.CA,C.CA2,C.CA3)

#HISTOGRAMAS
hist(cacao$Diasfermentacion)
hist(cacao$DiasSecado)
hist(cacao$Acidez)

```

```

#GRAFICO DE BARRAS
C.PESTI <- ggplot(Cacao,
                 aes(cacao$Pesticida))+
  geom_bar(fill="blue",
           colour="yellow")
C.PESTI+coord_flip()

C.CULT <- ggplot(cacao,
                 aes(cacao$MCultivo))+
  geom_bar(fill="green",
           colour="yellow")
C.CULT+coord_flip()

C.GEN <- ggplot(cacao,
                 aes(cacao$Genetica))+
  geom_bar(fill="purple",
           colour="blue")
C.GEN+coord_flip()

grid.arrange(C.PESTI+coord_flip(),C.CULT+coord_flip(),C.GEN+coord_flip())

#DATOS NO AGRUPADOS
#TABLA DE FRECUENCIA
CAFER<-as.data.frame(table(cacao$DiasFermentacion))
CAFER
TabCAFER<-transform(CAFER,
                    FreRel=round(prop.table(CAFER$Freq),3),#round es para redondear
                    FreAcu=cumsum(CAFER$Freq),
                    FreRelAcu=round(cumsum(prop.table(CAFER$Freq)),3))
TabCAFER

GraFER<-ggplot(TabCAFER,aes(x=TabCAFER$Var1,
                           y=TabCAFER$Freq))+geom_bar(stat="identity",fill="pink",
                           ,colour="blue")

GraFER

CASEC<-as.data.frame(table(cacao$DiasSecado))
CASEC
TabCASEC<-transform(CASEC,
                    FreRel=round(prop.table(CASEC$Freq),3),#round es para redondear
                    FreAcu=cumsum(CASEC$Freq),
                    FreRelAcu=round(cumsum(prop.table(CASEC$Freq)),3))
TabCASEC

```

```

GraSEC<-ggplot(TabCASEC,aes(x=TabCASEC$Var1,
                           y=TabCASEC$Freq))+geom_bar(stat="identity",fill="purple",
                           ,colour="blue")

GraSEC

CAACI<-as.data.frame(table(cacao$Acidez))
CAACI
TabCAACI<-transform(CAACI,
                    FreRel=round(prop.table(CAACI$Freq),3),#round es para redondear
                    FreAcu=cumsum(CAACI$Freq),
                    FreRelAcu=round(cumsum(prop.table(CAACI$Freq)),3))

TabCAACI

GraACI<-ggplot(TabCAACI,aes(x=TabCAACI$Var1,
                           y=TabCAACI$Freq))+geom_bar(stat="identity",fill="yellow",
                           ,colour="blue")

GraACI

#DATOS AGRUPADOS
#TABLA DE FRECUENCIA
FREPEST<-as.data.frame(table(cacao$Pesticida))
FREPEST
TabFREPEST<-transform(FREPEST,
                      FreRel=round(prop.table(FREPEST$Freq),3),#round es para redondear
                      FreAcu=cumsum(FREPEST$Freq),
                      FreRelAcu=round(cumsum(prop.table(FREPEST$Freq)),3))

TabFREPEST

GraPEST<-ggplot(TabFREPEST,aes(x=TabFREPEST$Var1,
                               y=TabFREPEST$Freq))+geom_bar(stat="identity",fill="skyblue",
                               ,colour="blue")

GraPEST

FRECVLT<-as.data.frame(table(cacao$MCultivo))
FRECVLT
TabFRECVLT<-transform(FRECVLT,
                      FreRel=round(prop.table(FRECVLT$Freq),3),#round es para redondear
                      FreAcu=cumsum(FRECVLT$Freq),
                      FreRelAcu=round(cumsum(prop.table(FRECVLT$Freq)),3))

TabFRECVLT

```



```

GraCULT<-ggplot(TabFREULT,aes(x=TabFREULT$Var1,
                             y=TabFREULT$Freq))+geom_bar(stat="identity",fill="grey",
                             ,colour="blue")

GraCULT

FREGEN<-as.data.frame(table(cacao$Genetica))
FREGEN
TabFREGEN<-transform(FREGEN,
                     FreRel=round(prop.table(FREGEN$Freq),3),#round es para redondear
                     FreAcu=cumsum(FREGEN$Freq),
                     FreRelAcu=round(cumsum(prop.table(FREGEN$Freq)),3))

TabFREGEN

GraGEN<-ggplot(TabFREGEN,aes(x=TabFREGEN$Var1,
                             y=TabFREGEN$Freq))+geom_bar(stat="identity",fill="green",
                             ,colour="blue")

GraGEN
#GRÁFICO DE EVOLUCIÓN
GEPEST<-ggplot(cacao,
               aes(x=cacao$Pesticida,
                   y=cacao$EdadPlantacion))+
  geom_boxplot(fill="yellow",colour="black")
GEPEST

GECULT<-ggplot(cacao,
               aes(x=cacao$MCultivo,
                   y=cacao$EdadPlantacion))+
  geom_boxplot(fill="orange",colour="black")
GECULT

GEGEN<-ggplot(cacao,
               aes(x=cacao$Genetica,
                   y=cacao$EdadPlantacion))+
  geom_boxplot(fill="red",colour="black")
GEGEN

GESECL<-ggplot(cacao,
               aes(x=cacao$Pesticida,
                   y=cacao$DiasSecado))+
  geom_boxplot(fill="yellow",colour="black")
GESECL

```

```

GESEC2<-ggplot(cacao,
               aes(x=cacao$MCultivo,
                   y=cacao$DiasSecado))+
  geom_boxplot(fill="orange",colour="black")
GESEC2

GESEC3<-ggplot(cacao,
               aes(x=cacao$Genetica,
                   y=cacao$DiasSecado))+
  geom_boxplot(fill="red",colour="black")
GESEC3

GEAC1<-ggplot(cacao,
               aes(x=cacao$Pesticida,
                   y=cacao$Acidez))+
  geom_boxplot(fill="yellow",colour="black")
GEAC1

GEAC2<-ggplot(cacao,
               aes(x=cacao$MCultivo,
                   y=cacao$Acidez))+
  geom_boxplot(fill="orange",colour="black")
GEAC2

GEAC3<-ggplot(cacao,
               aes(x=cacao$Genetica,
                   y=cacao$Acidez))+
  geom_boxplot(fill="red",colour="black")
GEAC3
ggpairs(cacao[,1:10])

```

Anexo 2

Análisis descriptivo a partir de la función Summary

```

> summary(cacao$Diasfermentacion)
  Min. 1st Qu.  Median    Mean 3rd Qu.  Max.
3.000  5.000  7.000  6.917  9.000 10.000
> summary(cacao$DiasSecado)
  Min. 1st Qu.  Median    Mean 3rd Qu.  Max.
12.0  34.0  45.0  48.1  65.0  75.0
> summary(cacao$Pesticida)
  Fungicida  Herbicida Insecticida
        78         89         86
> summary(cacao$MCultivo)
  Agroforesteria  EnPlanta  Monocultivo  Podas
        37         74         84         58
> summary(cacao$Genetica)
  criollo  forastero  trinitario
        85         99         69
> summary(cacao$Acidez)
  Min. 1st Qu.  Median    Mean 3rd Qu.  Max.
3.000  3.000  3.200  3.243  3.500  3.500

```


Anexo 4

Base de datos (cacao)

HectareasNa	HectareasCCN51	Edad Plantación	lotes	Días fermentación	Días Secado	Pesticida	MCultivo	Genetica	Acidez
0	15	6	3	4	30	1	2	1	3
2	3	23	1	3	75	3	4	2	3
18	2	14	4	4	45	2	1	1	3,5
4	20	16	3	5	75	1	4	3	3
10	13	4	2	6	67	1	2	1	3
7	3	23	6	7	34	2	3	3	3,2
9	4,5	40	7	8	41	3	1	2	3,3
4	4	21	10	9	75	1	4	3	3
2	3	34	12	9	70	3	3	2	3,2
6	2	23	2	9	15	2	2	1	3,5
7,5	2	12	1	8	40	2	2	3	3,5
12	5	3	4	7	50	2	3	1	3,5
4	12	4	4	10	23	3	3	2	3,4
11	2,5	9	5	8	56	1	3	2	3,3
5	2	10	3	6	34	3	2	1	3,2
1	3	14	8	4	75	2	1	3	3
1	2	15	10	6	63	2	3	1	3
10	10	1	12	5	12	3	2	2	3,5
4	3	40	2	4	43	1	2	2	3
3	13	38	4	10	30	3	3	1	3
2	3	35	5	9	61	2	2	3	3,2
5	10	12	6	9	30	3	3	2	3,3
14	12	23	9	9	45	2	4	2	3
3	3	34	8	8	67	1	3	1	3,2

20	5	23	4	7	75	1	4	3	3,5
7	2	36	5	10	43	2	2	2	3,5
4	4	40	7	8	23	2	3	1	3,5
2	1	10	9	6	30	1	4	3	3,4
4	2	26	2	4	34	3	4	1	3,3
5	5	34	3	6	26	2	3	2	3,2
4	1,5	40	12	5	34	3	2	2	3
6	1	27	15	4	75	3	1	3	3
4	2	35	10	10	63	3	3	1	3,5
9	3	34	2	10	30	1	1	2	3
10	12	24	3	4	32	3	1	3	3
12	4	4	5	4	35	2	3	2	3,2
3	3	2	1	5	39	1	4	2	3,3
6	4	6	10	8	46	1	2	1	3
4	8	10	3	9	48	1	4	3	3,2
1	5	40	4	6	50	3	2	2	3,5
10	13	4	5	7	75	2	4	1	3,5
7	3	23	3	3	43	3	1	2	3,5
9	4,5	40	8	4	23	2	4	1	3,4
4	4	21	10	5	45	3	2	3	3,3
2	3	34	12	6	34	2	3	1	3,2
6	2	23	2	7	46	1	1	3	3
7,5	2	12	4	8	44	2	4	2	3
12	5	3	5	9	75	1	3	3	3,5
4	12	4	6	9	63	3	2	2	3
11	2,5	9	12	9	43	2	2	1	3
5	2	10	2	8	42	3	3	3	3,2
1	3	14	5	7	45	3	3	1	3,3
1	2	15	9	10	49	2	3	2	3
10	10	1	17	8	65	1	2	2	3,2
4	3	40	2	6	66	3	1	1	3,5

3	13	38	10	4	66	2	3	3	3,5
2	3	35	5	6	40	2	2	1	3,5
5	10	12	2	5	45	3	2	2	3,4
14	12	23	4	4	67	1	3	2	3,3
3	3	34	1	10	30	3	2	1	3,2
20	5	23	2	9	32	2	3	3	3
7	2	36	5	9	75	3	4	2	3
4	4	40	5	9	38	2	3	2	3,5
2	1	10	1	8	56	3	4	1	3
4	2	26	1	7	65	2	2	3	3
5	5	34	10	10	75	2	3	2	3,2
4	1,5	40	4	8	45	3	4	1	3,3
6	1	27	3	6	75	1	4	3	3
4	2	35	2	4	67	3	3	1	3,2
9	3	34	5	6	34	2	2	2	3,5
10	12	24	14	5	41	2	1	2	3,5
12	4	4	3	4	75	3	3	3	3,5
3	3	2	20	10	70	3	1	1	3,4
6	4	6	7	10	34	1	1	2	3,3
4	8	10	4	4	40	3	3	3	3,2
1	5	40	2	4	50	2	4	2	3
18	2	14	4	5	33	1	2	2	3
4	20	16	5	8	56	1	4	1	3,5
10	13	4	4	9	34	1	2	3	3
7	3	23	6	6	75	3	4	2	3
9	4,5	40	4	7	63	2	1	3	3,2
4	4	21	9	3	34	3	4	2	3,3
2	3	34	10	4	43	2	2	2	3
6	2	23	4	5	32	1	3	1	3,2
7,5	2	12	10	6	61	1	1	3	3,5
12	5	3	7	7	33	2	4	2	3,5

4	12	4	9	8	45	3	3	1	3,5
11	2,5	9	4	9	67	1	2	3	3,4
5	2	10	2	9	75	3	2	1	3,3
1	3	14	6	9	43	2	3	2	3,2
1	2	15	2	8	23	3	3	2	3
10	10	1	18	7	13	2	3	3	3
4	3	40	4	10	34	1	2	1	3,5
10	12	24	10	8	26	1	1	2	3
12	4	4	7	6	34	3	3	3	3
3	3	2	9	4	75	2	2	2	3,2
6	4	6	4	6	63	1	2	2	3,3
4	8	10	2	5	12	1	3	1	3
1	5	40	6	4	32	1	2	3	3,2
18	2	14	1	10	35	3	3	2	3,5
4	20	16	4	9	39	2	4	1	3,5
10	13	4	3	9	34	3	3	2	3,5
7	3	23	2	9	32	2	4	1	3,4
9	4,5	40	6	8	30	1	2	3	3,3
4	4	21	7	7	75	1	3	1	3,2
2	3	34	10	10	43	2	4	3	3
6	2	23	12	8	30	3	4	2	3
7,5	2	12	2	6	30	1	3	3	3,5
12	5	3	1	4	34	3	2	2	3
4	12	4	4	6	45	2	1	1	3
11	2,5	9	4	5	55	3	3	1	3,2
5	2	10	5	4	75	2	1	3	3,3
1	3	14	3	10	63	1	1	2	3
1	2	15	8	10	45	1	3	3	3,2
10	10	1	10	4	34	3	4	2	3,5
4	3	40	12	4	35	2	2	1	3,5
4	4	21	2	5	56	2	4	3	3,5

2	3	34	4	8	66	3	2	1	3,4
6	2	23	5	9	62	1	4	2	3,3
7,5	2	12	6	6	35	3	1	2	3,2
12	5	3	9	7	40	2	4	1	3
4	12	4	8	3	45	3	2	3	3
11	2,5	9	4	4	67	2	3	1	3,5
5	2	10	5	5	30	3	1	2	3
1	3	14	7	6	32	2	4	2	3
1	2	15	9	7	75	2	3	1	3,2
18	2	14	2	8	38	3	2	3	3,3
4	20	16	3	9	56	1	2	2	3
10	13	4	12	9	65	3	3	2	3,2
7	3	23	15	9	30	2	3	1	3,5
9	4,5	40	10	8	34	1	3	3	3,5
4	4	21	2	7	26	1	2	2	3,5
2	3	34	3	10	34	3	1	1	3,4
6	2	23	5	8	75	1	3	3	3,3
7,5	2	12	1	6	63	3	2	1	3,2
12	5	3	10	4	30	2	2	2	3
4	12	4	3	6	32	2	3	2	3
11	2,5	9	4	5	35	2	2	3	3,5
5	2	10	5	4	39	1	3	1	3
1	3	14	3	10	46	3	4	2	3
1	2	15	8	9	48	2	3	3	3,2
10	10	1	10	9	50	3	4	2	3,3
4	3	40	12	9	75	2	2	2	3
10	12	24	2	8	43	1	3	1	3,2
12	4	4	4	7	23	1	4	3	3,5
3	3	2	5	10	45	3	4	2	3,5
6	4	6	6	8	34	2	3	1	3,5
4	8	10	12	6	46	1	2	2	3,4

1	5	40	2	4	44	3	1	1	3,3
18	2	14	5	6	75	2	3	3	3,2
4	20	16	9	5	63	1	1	1	3
10	13	4	17	4	43	1	1	3	3
2	3	23	2	10	42	2	3	2	3,5
18	2	14	10	10	45	1	4	3	3
4	20	16	5	4	49	3	2	2	3
10	13	4	2	4	65	2	4	1	3,2
7	3	23	4	5	66	1	2	3	3,3
9	4,5	40	1	8	66	1	4	1	3
4	4	21	2	9	40	1	1	2	3,2
2	3	34	5	6	45	3	4	2	3,5
6	2	23	5	7	67	2	2	1	3,5
7,5	2	12	1	3	30	3	3	3	3,5
12	5	3	1	4	32	2	1	1	3,4
4	12	4	10	5	75	3	4	2	3,3
11	2,5	9	4	6	38	2	3	2	3,2
5	2	10	3	7	56	3	2	1	3
1	3	14	2	8	65	2	2	1	3
1	2	15	5	9	75	1	3	3	3,5
10	10	1	14	9	45	3	3	1	3
4	3	40	3	9	75	2	3	2	3
3	13	38	20	8	67	2	2	2	3,2
10	13	4	7	7	34	3	1	3	3,3
7	3	23	4	10	41	1	3	1	3
9	4,5	40	2	8	75	1	2	2	3,2
4	4	21	4	6	70	3	2	3	3,5
2	3	34	5	4	34	2	3	2	3,5
6	2	23	4	6	40	1	2	2	3,5
7,5	2	12	6	5	50	1	3	1	3,4
12	5	3	4	4	33	1	4	3	3,3

4	12	4	9	10	56	3	3	2	3,2
11	2,5	9	10	9	34	2	4	1	3
6	2	23	4	9	75	3	2	2	3,3
7,5	2	12	10	9	63	2	3	1	3,4
12	5	3	7	8	34	3	4	3	3,5
4	12	4	9	7	43	2	4	1	3,4
11	2,5	9	4	10	32	3	3	1	3
5	2	10	2	8	61	2	2	3	3,4
1	3	14	6	6	33	1	1	1	3,5
1	2	15	2	4	45	3	3	2	3,3
10	10	1	18	6	67	2	1	2	3
4	3	40	4	5	75	2	1	3	3
7,5	2	12	10	4	43	2	3	1	3
12	5	3	7	10	23	3	4	2	3,4
4	12	4	9	10	13	1	2	3	3,2
11	2,5	9	4	4	34	3	4	2	3,1
5	2	10	2	4	26	2	2	2	3,3
1	3	14	6	5	34	1	2	1	3,3
3	13	38	6	8	75	1	2	3	3,5
4	4	21	2	7	26	1	2	2	3,5
2	3	34	3	10	34	3	1	1	3,4
6	2	23	5	8	75	1	3	3	3,3
7,5	2	12	1	6	63	3	2	1	3,2
12	5	3	10	4	30	2	2	2	3
4	12	4	3	6	32	2	3	2	3
11	2,5	9	4	5	35	2	2	3	3,5
5	2	10	5	4	39	1	3	1	3
1	3	14	3	10	46	3	4	2	3
1	2	15	8	9	48	2	3	3	3,2
10	10	1	10	9	50	3	4	2	3,3
4	3	40	12	9	75	2	2	2	3

10	12	24	2	8	43	1	3	1	3,2
12	4	4	4	7	23	1	4	3	3,5
3	3	2	5	10	45	3	4	2	3,5
6	4	6	6	8	34	2	3	1	3,5
4	8	10	12	6	46	1	2	2	3,4
1	5	40	2	4	44	3	1	1	3,3
18	2	14	5	6	75	2	3	3	3,2
4	20	16	9	5	63	1	1	1	3
10	13	4	17	4	43	1	1	3	3
2	3	23	2	10	42	2	3	2	3,5
18	2	14	10	10	45	1	4	3	3
4	20	16	5	4	49	3	2	2	3
10	13	4	2	4	65	2	4	1	3,2
7	3	23	4	5	66	1	2	3	3,3
9	4,5	40	1	8	66	1	4	1	3
4	4	21	2	9	40	1	1	2	3,2
2	3	34	5	6	45	3	4	2	3,5
6	2	23	5	7	67	2	2	1	3,5
7,5	2	12	1	3	30	3	3	3	3,5
12	5	3	1	4	32	2	1	1	3,4
4	12	4	10	5	75	3	4	2	3,3
11	2,5	9	4	6	38	2	3	2	3,2
5	2	10	3	7	56	3	2	1	3
1	3	14	2	8	65	2	2	1	3
1	2	15	5	9	75	1	3	3	3,5
10	10	1	14	9	45	3	3	1	3
4	3	40	3	9	75	2	3	2	3
3	13	38	20	8	67	2	2	2	3,2
10	13	4	7	7	34	3	1	3	3,3
7	3	23	4	10	41	1	3	1	3
9	4,5	40	2	8	75	1	2	2	3,2

4	4	21	4	6	70	3	2	3	3,5
2	3	34	5	4	34	2	3	2	3,5
6	2	23	4	6	40	1	2	2	3,5
7,5	2	12	6	5	50	1	3	1	3,4
12	5	3	4	4	33	1	4	3	3,3
4	12	4	9	10	56	3	3	2	3,2
11	2,5	9	10	9	34	2	4	1	3
6	2	23	4	9	75	3	2	2	3,3
7,5	2	12	10	9	63	2	3	1	3,4
12	5	3	7	8	34	3	4	3	3,5
4	12	4	9	7	43	2	4	1	3,4
11	2,5	9	4	10	32	3	3	1	3

Anexo 5

Predicción RandomForest

K_Means	Predicción	RandProb [, Exportación 1]	RandProb [, Nacional]	RandProb [, Exportación 2]	RandProb [, Exportación 3]	RandProb [, Exportación 4]	RandProb [, Exportación 5]
Exportacion 5	Exportacion 5	0.030666667	0.172000000	0.053666667	0.036666667	0.678000000	0.029000000
Nacional	Nacional	0.286000000	0.598333333	0.004666667	0.065000000	0.024000000	0.022000000
Exportacion 6	Exportacion 6	0.008666667	0.023000000	0.008333333	0.076666667	0.005333333	0.878000000
Exportacion 5	Exportacion 5	0.128333333	0.026000000	0.057666667	0.001000000	0.763000000	0.024000000
Exportacion 5	Exportacion 5	0.007000000	0.043000000	0.067000000	0.016666667	0.861000000	0.005333333
Exportacion 1	Exportacion 1	0.523333333	0.046666667	0.016333333	0.157333333	0.019333333	0.237000000
Exportacion 4	Exportacion 4	0.109666667	0.023000000	0.018666667	0.833000000	0.009666667	0.006000000
Exportacion 1	Exportacion 1	0.918000000	0.054333333	0.001666667	0.001333333	0.014666667	0.010000000
Exportacion 1	Exportacion 1	0.801666667	0.073333333	0.015000000	0.096666667	0.008666667	0.004666667
Exportacion 4	Exportacion 4	0.027000000	0.024000000	0.058666667	0.884000000	0.001333333	0.005000000
Exportacion 6	Exportacion 6	0.304333333	0.023333333	0.028666667	0.075333333	0.012000000	0.556333333
Exportacion 6	Exportacion 6	0.009666667	0.096666667	0.110000000	0.027666667	0.004000000	0.752000000
Exportacion 3	Exportacion 3	0.005000000	0.014000000	0.920666667	0.009333333	0.040666667	0.010333333

Exportacion 6	Exportacion 6	0.080666667	0.031333333	0.042333333	0.023000000	0.004000000	0.818666667
Nacional	Nacional	0.013666667	0.781333333	0.005666667	0.150333333	0.006666667	0.042333333
Exportacion 1	Exportacion 1	0.878666667	0.073333333	0.001000000	0.004333333	0.021333333	0.021333333
Nacional	Nacional	0.030333333	0.926333333	0.011666667	0.018000000	0.010666667	0.003000000
Exportacion 3	Exportacion 3	0.015000000	0.005000000	0.711000000	0.065333333	0.180666667	0.023000000
Exportacion 1	Exportacion 1	0.844000000	0.070666667	0.000000000	0.074333333	0.006333333	0.004666667
Nacional	Nacional	0.041000000	0.699000000	0.086000000	0.071666667	0.096333333	0.006000000
Exportacion 1	Exportacion 1	0.974000000	0.003666667	0.000000000	0.016000000	0.003333333	0.003000000
Exportacion 3	Exportacion 3	0.016000000	0.026000000	0.833000000	0.078666667	0.027666667	0.018666667
Exportacion 3	Exportacion 3	0.012000000	0.109333333	0.685666667	0.002333333	0.085000000	0.105666667
Exportacion 1	Exportacion 1	0.719666667	0.134000000	0.006000000	0.133000000	0.004333333	0.003000000
Exportacion 6	Exportacion 6	0.289666667	0.025666667	0.025000000	0.039000000	0.016666667	0.604000000
Exportacion 4	Exportacion 4	0.099666667	0.033333333	0.088666667	0.761666667	0.004333333	0.012333333
Exportacion 4	Exportacion 4	0.014333333	0.009666667	0.059666667	0.914333333	0.000000000	0.002000000
Exportacion 1	Exportacion 1	0.781000000	0.013000000	0.017333333	0.039666667	0.017666667	0.131333333
Exportacion 4	Exportacion 4	0.010333333	0.024000000	0.005666667	0.947333333	0.001000000	0.011666667
Exportacion 4	Exportacion 4	0.102000000	0.059333333	0.005333333	0.808000000	0.005666667	0.019666667
Exportacion 4	Exportacion 4	0.169333333	0.106333333	0.000000000	0.704666667	0.011333333	0.008333333
Exportacion 1	Exportacion 1	0.882000000	0.049666667	0.005000000	0.024333333	0.020666667	0.018333333
Exportacion 4	Exportacion 4	0.054000000	0.020666667	0.074333333	0.844333333	0.003000000	0.003666667
Exportacion 1	Exportacion 1	0.772000000	0.107333333	0.003000000	0.087333333	0.018000000	0.012333333
Exportacion 5	Exportacion 5	0.066666667	0.025333333	0.009333333	0.024000000	0.836000000	0.038666667
Exportacion 6	Exportacion 6	0.008333333	0.007333333	0.017000000	0.011000000	0.002666667	0.953666667
Exportacion 6	Exportacion 6	0.068666667	0.025333333	0.122000000	0.080333333	0.014333333	0.689333333
Nacional	Nacional	0.059333333	0.814666667	0.033000000	0.017666667	0.057000000	0.018333333
Exportacion 1	Exportacion 1	0.760333333	0.024333333	0.040333333	0.003333333	0.054000000	0.117666667
Exportacion 4	Exportacion 4	0.064333333	0.025666667	0.004333333	0.891333333	0.011666667	0.002666667
Exportacion 3	Exportacion 3	0.010000000	0.009000000	0.925000000	0.007333333	0.029333333	0.019333333
Exportacion 4	Exportacion 4	0.091000000	0.031000000	0.048000000	0.771333333	0.016000000	0.042666667
Exportacion 4	Exportacion 4	0.020000000	0.019666667	0.041666667	0.899666667	0.007000000	0.012000000

Exportacion 1	Exportacion 1	0.874333333	0.002000000	0.002333333	0.071000000	0.013000000	0.037333333
Exportacion 4	Exportacion 4	0.085000000	0.069333333	0.006666667	0.830666667	0.004000000	0.004333333
Exportacion 1	Exportacion 1	0.891000000	0.043000000	0.000000000	0.012666667	0.017333333	0.036000000
Nacional	Nacional	0.036000000	0.895666667	0.009666667	0.019333333	0.004666667	0.034666667
Exportacion 6	Exportacion 6	0.093666667	0.010333333	0.039666667	0.001666667	0.005666667	0.849000000
Exportacion 5	Exportacion 5	0.011000000	0.018000000	0.105000000	0.004333333	0.858333333	0.003333333
Nacional	Nacional	0.014666667	0.899333333	0.008000000	0.009666667	0.009666667	0.058666667
Exportacion 6	Exportacion 6	0.145000000	0.041666667	0.012000000	0.020333333	0.004333333	0.776666667
Exportacion 4	Exportacion 4	0.013333333	0.089666667	0.041333333	0.825666667	0.003666667	0.026333333
Nacional	Nacional	0.107333333	0.863000000	0.010000000	0.008666667	0.005666667	0.005333333
Exportacion 5	Exportacion 5	0.033666667	0.013333333	0.108333333	0.008666667	0.822333333	0.013666667
Exportacion 4	Exportacion 4	0.042333333	0.020333333	0.011333333	0.920000000	0.002000000	0.004000000
Exportacion 1	Exportacion 1	0.735666667	0.025333333	0.056333333	0.030666667	0.114666667	0.037333333
Exportacion 4	Exportacion 4	0.007000000	0.002666667	0.000000000	0.989666667	0.000333333	0.000333333
Exportacion 4	Exportacion 4	0.012333333	0.025000000	0.120000000	0.677000000	0.148333333	0.017333333
Exportacion 6	Exportacion 6	0.041000000	0.050000000	0.089000000	0.027000000	0.051666667	0.741333333
Exportacion 4	Exportacion 4	0.100333333	0.081666667	0.004000000	0.799666667	0.003333333	0.011000000
Exportacion 6	Exportacion 6	0.078000000	0.044333333	0.032666667	0.007666667	0.022000000	0.815333333
Nacional	Nacional	0.212666667	0.758666667	0.004333333	0.023000000	0.000333333	0.001000000
Exportacion 4	Exportacion 4	0.066000000	0.010666667	0.095333333	0.822000000	0.001000000	0.005000000
Nacional	Nacional	0.014000000	0.949666667	0.008000000	0.017000000	0.001666667	0.009666667
Exportacion 1	Exportacion 1	0.940666667	0.024666667	0.000000000	0.013666667	0.007333333	0.013666667
Exportacion 1	Exportacion 1	0.666666667	0.144333333	0.031333333	0.125666667	0.019333333	0.012666667
Exportacion 4	Exportacion 4	0.035000000	0.029000000	0.016333333	0.916000000	0.000666667	0.003000000
Exportacion 1	Exportacion 1	0.903333333	0.055666667	0.000000000	0.007333333	0.018666667	0.015000000
Exportacion 4	Exportacion 4	0.278666667	0.263666667	0.006666667	0.421000000	0.011000000	0.019000000
Exportacion 4	Exportacion 4	0.022000000	0.005666667	0.016333333	0.940000000	0.003333333	0.012666667
Exportacion 5	Exportacion 5	0.015666667	0.021666667	0.106333333	0.098666667	0.740666667	0.017000000
Exportacion 6	Exportacion 6	0.035333333	0.006333333	0.018000000	0.001666667	0.004000000	0.934666667
Exportacion 3	Exportacion 3	0.058333333	0.059333333	0.772666667	0.053666667	0.037333333	0.018666667

Exportacion 1	Exportacion 3	0.238666667	0.069333333	0.367000000	0.154333333	0.062666667	0.108000000
Exportacion 6	Exportacion 6	0.124000000	0.024333333	0.030666667	0.024666667	0.055333333	0.741000000
Nacional	Nacional	0.199333333	0.711333333	0.004000000	0.070000000	0.007666667	0.007666667
Exportacion 6	Exportacion 6	0.022333333	0.081000000	0.003000000	0.018333333	0.003666667	0.871666667
Exportacion 3	Exportacion 3	0.017333333	0.041666667	0.812000000	0.053333333	0.057333333	0.018333333
Exportacion 5	Exportacion 5	0.014333333	0.010333333	0.028000000	0.000666667	0.925000000	0.021666667
Nacional	Nacional	0.111333333	0.847000000	0.004000000	0.024333333	0.009333333	0.004000000
Exportacion 1	Exportacion 1	0.945333333	0.007666667	0.000333333	0.018000000	0.009000000	0.019666667
Exportacion 4	Exportacion 4	0.040000000	0.032333333	0.029666667	0.874333333	0.005333333	0.018333333
Exportacion 1	Exportacion 1	0.793666667	0.085000000	0.000333333	0.103000000	0.014000000	0.004000000
Exportacion 4	Exportacion 4	0.037666667	0.185666667	0.006000000	0.757000000	0.000000000	0.013666667
Exportacion 1	Exportacion 1	0.842666667	0.006000000	0.016333333	0.037666667	0.016666667	0.080666667
Exportacion 6	Exportacion 6	0.006000000	0.029333333	0.124000000	0.010333333	0.003000000	0.827333333
Exportacion 3	Exportacion 3	0.002666667	0.009666667	0.960333333	0.007666667	0.016000000	0.003666667
Exportacion 1	Exportacion 1	0.688333333	0.022333333	0.011333333	0.002000000	0.010000000	0.266000000
Nacional	Nacional	0.110333333	0.729000000	0.055666667	0.080000000	0.005666667	0.019333333
Nacional	Nacional	0.245333333	0.317333333	0.116666667	0.250666667	0.005333333	0.064666667
Nacional	Nacional	0.024000000	0.892333333	0.002333333	0.056333333	0.004333333	0.020666667
Exportacion 5	Exportacion 5	0.026000000	0.026000000	0.082333333	0.001333333	0.817000000	0.047333333
Exportacion 4	Exportacion 4	0.039000000	0.016666667	0.010000000	0.932333333	0.000333333	0.001666667
Exportacion 5	Exportacion 5	0.043000000	0.069333333	0.031666667	0.021000000	0.828000000	0.007000000
Exportacion 6	Exportacion 6	0.020333333	0.038333333	0.015333333	0.001666667	0.025666667	0.898666667
Exportacion 1	Exportacion 1	0.736666667	0.076333333	0.072333333	0.023333333	0.039000000	0.052333333
Exportacion 1	Exportacion 1	0.717333333	0.049333333	0.053333333	0.069666667	0.036666667	0.073666667
Nacional	Nacional	0.014000000	0.807333333	0.028666667	0.039000000	0.088000000	0.023000000
Exportacion 1	Exportacion 1	0.861333333	0.015000000	0.000000000	0.068333333	0.012666667	0.042666667
Exportacion 6	Exportacion 6	0.013666667	0.036000000	0.071666667	0.046333333	0.001666667	0.830666667
Exportacion 3	Exportacion 3	0.006666667	0.028000000	0.873000000	0.058000000	0.024666667	0.009666667
Exportacion 3	Exportacion 3	0.003666667	0.017333333	0.905333333	0.016000000	0.039000000	0.018666667
Exportacion 4	Exportacion 4	0.042666667	0.058666667	0.261000000	0.609333333	0.002333333	0.026000000

Exportacion 1	Exportacion 1	0.847666667	0.004666667	0.004666667	0.070333333	0.018000000	0.054666667
Nacional	Nacional	0.179333333	0.730666667	0.006666667	0.075000000	0.003666667	0.004666667
Exportacion 1	Exportacion 1	0.866000000	0.068666667	0.001666667	0.020333333	0.010666667	0.032666667
Nacional	Nacional	0.054333333	0.808333333	0.013333333	0.108000000	0.009333333	0.006666667
Exportacion 6	Exportacion 6	0.116333333	0.003000000	0.006000000	0.035333333	0.001333333	0.838000000
Exportacion 6	Exportacion 6	0.005666667	0.027666667	0.010000000	0.007333333	0.010000000	0.939333333
Exportacion 5	Exportacion 5	0.004666667	0.029333333	0.052000000	0.008333333	0.901333333	0.004333333
Nacional	Exportacion 6	0.019333333	0.273666667	0.027666667	0.044333333	0.004333333	0.630666667
Exportacion 1	Exportacion 1	0.896666667	0.013000000	0.006000000	0.020666667	0.007666667	0.056000000
Exportacion 1	Exportacion 1	0.739666667	0.223666667	0.005333333	0.011666667	0.011666667	0.008000000
Exportacion 1	Exportacion 1	0.733333333	0.062333333	0.012000000	0.053666667	0.015000000	0.123666667
Exportacion 3	Exportacion 3	0.019000000	0.007333333	0.802000000	0.034666667	0.067666667	0.069333333
Exportacion 4	Exportacion 4	0.024666667	0.004666667	0.021333333	0.928333333	0.019333333	0.001666667
Exportacion 1	Exportacion 1	0.878333333	0.008666667	0.007000000	0.056333333	0.007000000	0.042666667
Exportacion 4	Exportacion 4	0.059666667	0.023000000	0.010333333	0.903000000	0.002666667	0.001333333
Exportacion 1	Exportacion 1	0.441666667	0.120333333	0.097333333	0.306000000	0.003333333	0.031333333
Exportacion 4	Exportacion 4	0.043333333	0.177000000	0.017000000	0.697666667	0.006666667	0.058333333
Nacional	Nacional	0.002666667	0.723666667	0.034666667	0.000666667	0.008333333	0.230000000
Exportacion 5	Exportacion 5	0.026666667	0.003666667	0.019000000	0.004666667	0.929666667	0.016333333
Exportacion 6	Exportacion 6	0.022666667	0.043666667	0.038000000	0.045000000	0.004333333	0.846333333
Nacional	Nacional	0.061333333	0.783666667	0.002000000	0.101000000	0.018000000	0.034000000
Nacional	Nacional	0.042666667	0.885333333	0.004333333	0.045666667	0.005000000	0.017000000
Nacional	Nacional	0.086000000	0.817666667	0.020000000	0.065000000	0.004000000	0.007333333
Exportacion 6	Exportacion 6	0.053000000	0.012333333	0.005333333	0.009666667	0.008666667	0.911000000
Exportacion 5	Exportacion 5	0.081333333	0.071333333	0.080666667	0.019666667	0.742333333	0.004666667
Exportacion 3	Exportacion 3	0.008666667	0.011000000	0.916000000	0.002000000	0.057666667	0.004666667
Exportacion 3	Exportacion 3	0.011666667	0.029333333	0.692333333	0.239666667	0.020666667	0.006333333
Exportacion 1	Exportacion 1	0.589333333	0.011666667	0.038666667	0.166000000	0.023000000	0.171333333
Exportacion 4	Exportacion 4	0.036000000	0.003000000	0.004000000	0.950000000	0.003000000	0.004000000
Exportacion 4	Exportacion 4	0.006666667	0.004666667	0.005000000	0.983000000	0.000000000	0.000666667

Exportacion 1	Exportacion 1	0.964666667	0.005666667	0.002000000	0.008333333	0.001000000	0.018333333
Nacional	Nacional	0.040666667	0.817000000	0.013666667	0.091333333	0.005666667	0.031666667
Exportacion 6	Exportacion 6	0.005666667	0.038000000	0.005333333	0.000666667	0.003666667	0.946666667
Exportacion 5	Exportacion 5	0.005333333	0.051000000	0.144666667	0.008333333	0.775333333	0.015333333
Exportacion 6	Exportacion 6	0.043000000	0.002000000	0.000333333	0.006333333	0.005000000	0.943333333
Nacional	Nacional	0.007333333	0.967666667	0.001333333	0.014666667	0.002000000	0.007000000
Nacional	Nacional	0.031000000	0.952000000	0.005333333	0.005666667	0.001666667	0.004333333
Exportacion 1	Exportacion 1	0.926000000	0.023000000	0.004666667	0.007000000	0.001666667	0.037666667
Exportacion 3	Exportacion 3	0.004000000	0.005000000	0.955000000	0.007333333	0.021666667	0.007000000
Exportacion 1	Exportacion 1	0.917666667	0.047000000	0.001666667	0.022000000	0.011333333	0.000333333
Nacional	Nacional	0.012000000	0.885333333	0.044000000	0.013000000	0.036333333	0.009333333
Exportacion 6	Exportacion 6	0.006666667	0.001333333	0.021000000	0.001000000	0.002000000	0.968000000
Exportacion 3	Exportacion 3	0.009666667	0.012000000	0.918000000	0.029666667	0.001000000	0.029666667
Exportacion 3	Exportacion 3	0.011666667	0.013000000	0.908333333	0.040333333	0.005000000	0.021666667
Exportacion 5	Exportacion 5	0.081333333	0.030000000	0.064333333	0.153000000	0.649000000	0.022333333
Exportacion 4	Exportacion 4	0.041666667	0.031666667	0.002000000	0.912666667	0.003666667	0.008333333
Exportacion 6	Exportacion 6	0.035666667	0.003666667	0.001000000	0.002333333	0.001000000	0.956333333
Exportacion 5	Exportacion 5	0.008666667	0.019000000	0.012666667	0.004333333	0.954000000	0.001333333
Exportacion 5	Exportacion 5	0.004333333	0.000333333	0.003333333	0.000666667	0.987666667	0.003666667
Exportacion 4	Exportacion 4	0.016333333	0.008000000	0.028333333	0.942000000	0.000000000	0.005333333
Exportacion 6	Exportacion 6	0.078000000	0.075333333	0.046666667	0.001333333	0.009333333	0.789333333
Exportacion 5	Exportacion 5	0.006666667	0.018000000	0.011000000	0.011333333	0.950000000	0.003000000
Exportacion 3	Exportacion 3	0.006000000	0.021333333	0.916666667	0.004000000	0.039666667	0.012333333
Exportacion 1	Exportacion 1	0.979000000	0.003000000	0.000666667	0.010666667	0.001333333	0.005333333
Nacional	Nacional	0.387666667	0.475000000	0.016333333	0.101666667	0.010666667	0.008666667
Exportacion 1	Exportacion 1	0.901333333	0.021333333	0.003000000	0.066333333	0.004333333	0.003666667
Exportacion 4	Exportacion 4	0.009000000	0.007666667	0.009333333	0.972000000	0.000000000	0.002000000
Exportacion 4	Exportacion 4	0.035333333	0.016333333	0.014000000	0.931000000	0.001000000	0.002333333
Exportacion 6	Exportacion 6	0.016666667	0.004333333	0.004666667	0.013666667	0.001333333	0.959333333
Exportacion 6	Exportacion 6	0.004333333	0.019333333	0.011000000	0.009000000	0.001666667	0.954666667

Exportacion 3	Exportacion 3	0.014666667	0.002666667	0.948333333	0.001333333	0.022666667	0.010333333
Exportacion 6	Exportacion 6	0.011666667	0.049333333	0.005333333	0.023333333	0.001666667	0.908666667
Nacional	Nacional	0.017000000	0.955666667	0.002333333	0.011000000	0.010000000	0.004000000
Nacional	Nacional	0.036333333	0.944333333	0.002000000	0.013000000	0.002666667	0.001666667
Exportacion 1	Exportacion 1	0.751000000	0.031000000	0.035000000	0.070666667	0.006333333	0.106000000
Exportacion 3	Exportacion 3	0.001333333	0.077000000	0.764000000	0.001333333	0.153000000	0.003333333
Exportacion 1	Exportacion 1	0.934333333	0.056333333	0.000333333	0.006000000	0.001333333	0.001666667
Exportacion 5	Exportacion 5	0.122000000	0.040000000	0.036666667	0.025000000	0.770333333	0.006000000
Exportacion 5	Exportacion 5	0.010000000	0.001000000	0.043333333	0.002666667	0.928666667	0.014333333
Nacional	Nacional	0.034000000	0.931000000	0.003333333	0.027333333	0.001666667	0.002666667
Exportacion 1	Exportacion 1	0.949666667	0.026666667	0.000333333	0.017333333	0.003666667	0.002333333
Exportacion 1	Exportacion 1	0.956000000	0.001333333	0.000333333	0.029000000	0.002000000	0.011333333
Exportacion 4	Exportacion 4	0.013333333	0.004333333	0.004000000	0.973333333	0.001333333	0.003666667
Exportacion 4	Exportacion 4	0.056333333	0.009666667	0.003333333	0.918666667	0.003333333	0.008666667
Exportacion 4	Exportacion 4	0.038666667	0.068000000	0.103333333	0.719666667	0.014666667	0.055666667
Exportacion 6	Exportacion 6	0.019333333	0.002666667	0.011000000	0.001000000	0.005000000	0.961000000
Exportacion 3	Exportacion 3	0.010333333	0.019666667	0.902000000	0.002333333	0.062333333	0.003333333
Nacional	Nacional	0.001333333	0.955666667	0.012000000	0.001333333	0.001000000	0.028666667
Exportacion 1	Exportacion 1	0.752666667	0.096666667	0.014666667	0.127333333	0.003333333	0.005333333
Exportacion 3	Exportacion 3	0.042333333	0.085000000	0.709000000	0.141666667	0.004666667	0.017333333
Exportacion 6	Exportacion 6	0.038000000	0.012666667	0.143666667	0.001333333	0.028000000	0.776333333
Exportacion 3	Exportacion 3	0.004333333	0.008000000	0.924000000	0.008333333	0.046666667	0.008666667
Nacional	Nacional	0.011666667	0.849000000	0.012666667	0.008000000	0.003000000	0.115666667
Exportacion 1	Exportacion 1	0.861000000	0.021000000	0.019000000	0.019666667	0.006666667	0.072666667
Exportacion 4	Exportacion 4	0.039333333	0.038333333	0.037666667	0.840666667	0.013666667	0.030333333
Exportacion 4	Exportacion 4	0.025333333	0.053333333	0.014666667	0.848333333	0.002333333	0.056000000
Exportacion 5	Exportacion 5	0.015000000	0.011000000	0.075000000	0.007666667	0.883333333	0.008000000
Exportacion 1	Exportacion 1	0.979666667	0.008000000	0.000000000	0.004333333	0.007000000	0.001000000
Nacional	Nacional	0.022333333	0.924333333	0.010000000	0.027666667	0.005000000	0.010666667
Exportacion 3	Exportacion 3	0.013333333	0.020666667	0.675666667	0.004000000	0.001666667	0.284666667

Exportacion 5	Exportacion 5	0.044666667	0.007000000	0.065333333	0.005000000	0.864333333	0.013666667
Exportacion 6	Exportacion 6	0.004000000	0.178000000	0.008666667	0.008333333	0.003666667	0.797333333
Exportacion 4	Exportacion 4	0.054666667	0.113000000	0.020000000	0.713666667	0.010333333	0.088333333
Exportacion 4	Exportacion 4	0.079333333	0.133333333	0.022333333	0.703000000	0.021666667	0.040333333
Exportacion 1	Exportacion 1	0.516666667	0.029333333	0.055333333	0.063333333	0.310333333	0.025000000
Exportacion 4	Exportacion 4	0.036000000	0.003000000	0.004000000	0.950000000	0.003000000	0.004000000
Exportacion 4	Exportacion 4	0.006666667	0.004666667	0.005000000	0.983000000	0.000000000	0.000666667
Exportacion 1	Exportacion 1	0.964666667	0.005666667	0.002000000	0.008333333	0.001000000	0.018333333
Nacional	Nacional	0.040666667	0.817000000	0.013666667	0.091333333	0.005666667	0.031666667
Exportacion 6	Exportacion 6	0.005666667	0.038000000	0.005333333	0.000666667	0.003666667	0.946666667
Exportacion 5	Exportacion 5	0.005333333	0.051000000	0.144666667	0.008333333	0.775333333	0.015333333
Exportacion 6	Exportacion 6	0.043000000	0.002000000	0.000333333	0.006333333	0.005000000	0.943333333
Nacional	Nacional	0.007333333	0.967666667	0.001333333	0.014666667	0.002000000	0.007000000
Nacional	Nacional	0.031000000	0.952000000	0.005333333	0.005666667	0.001666667	0.004333333
Exportacion 1	Exportacion 1	0.926000000	0.023000000	0.004666667	0.007000000	0.001666667	0.037666667
Exportacion 3	Exportacion 3	0.004000000	0.005000000	0.955000000	0.007333333	0.021666667	0.007000000
Exportacion 1	Exportacion 1	0.917666667	0.047000000	0.001666667	0.022000000	0.011333333	0.000333333
Nacional	Nacional	0.012000000	0.885333333	0.044000000	0.013000000	0.036333333	0.009333333
Exportacion 6	Exportacion 6	0.006666667	0.001333333	0.021000000	0.001000000	0.002000000	0.968000000
Exportacion 3	Exportacion 3	0.009666667	0.012000000	0.918000000	0.029666667	0.001000000	0.029666667
Exportacion 3	Exportacion 3	0.011666667	0.013000000	0.908333333	0.040333333	0.005000000	0.021666667
Exportacion 5	Exportacion 5	0.081333333	0.030000000	0.064333333	0.153000000	0.649000000	0.022333333
Exportacion 4	Exportacion 4	0.041666667	0.031666667	0.002000000	0.912666667	0.003666667	0.008333333
Exportacion 6	Exportacion 6	0.035666667	0.003666667	0.001000000	0.002333333	0.001000000	0.956333333
Exportacion 5	Exportacion 5	0.008666667	0.019000000	0.012666667	0.004333333	0.954000000	0.001333333
Exportacion 5	Exportacion 5	0.004333333	0.000333333	0.003333333	0.000666667	0.987666667	0.003666667
Exportacion 4	Exportacion 4	0.016333333	0.008000000	0.028333333	0.942000000	0.000000000	0.005333333
Exportacion 6	Exportacion 6	0.078000000	0.075333333	0.046666667	0.001333333	0.009333333	0.789333333
Exportacion 5	Exportacion 5	0.006666667	0.018000000	0.011000000	0.011333333	0.950000000	0.003000000
Exportacion 3	Exportacion 3	0.006000000	0.021333333	0.916666667	0.004000000	0.039666667	0.012333333

Exportacion 1	Exportacion 1	0.979000000	0.003000000	0.000666667	0.010666667	0.001333333	0.005333333
Nacional	Nacional	0.387666667	0.475000000	0.016333333	0.101666667	0.010666667	0.008666667
Exportacion 1	Exportacion 1	0.901333333	0.021333333	0.003000000	0.066333333	0.004333333	0.003666667
Exportacion 4	Exportacion 4	0.009000000	0.007666667	0.009333333	0.972000000	0.000000000	0.002000000
Exportacion 4	Exportacion 4	0.035333333	0.016333333	0.014000000	0.931000000	0.001000000	0.002333333
Exportacion 6	Exportacion 6	0.016666667	0.004333333	0.004666667	0.013666667	0.001333333	0.959333333
Exportacion 6	Exportacion 6	0.004333333	0.019333333	0.011000000	0.009000000	0.001666667	0.954666667
Exportacion 3	Exportacion 3	0.014666667	0.002666667	0.948333333	0.001333333	0.022666667	0.010333333
Exportacion 6	Exportacion 6	0.011666667	0.049333333	0.005333333	0.023333333	0.001666667	0.908666667
Nacional	Nacional	0.017000000	0.955666667	0.002333333	0.011000000	0.010000000	0.004000000
Nacional	Nacional	0.036333333	0.944333333	0.002000000	0.013000000	0.002666667	0.001666667
Exportacion 1	Exportacion 1	0.751000000	0.031000000	0.035000000	0.070666667	0.006333333	0.106000000
Exportacion 3	Exportacion 3	0.001333333	0.077000000	0.764000000	0.001333333	0.153000000	0.003333333
Exportacion 1	Exportacion 1	0.934333333	0.056333333	0.000333333	0.006000000	0.001333333	0.001666667
Exportacion 5	Exportacion 5	0.122000000	0.040000000	0.036666667	0.025000000	0.770333333	0.006000000
Exportacion 5	Exportacion 5	0.010000000	0.001000000	0.043333333	0.002666667	0.928666667	0.014333333
Nacional	Nacional	0.034000000	0.931000000	0.003333333	0.027333333	0.001666667	0.002666667
Exportacion 1	Exportacion 1	0.949666667	0.026666667	0.000333333	0.017333333	0.003666667	0.002333333
Exportacion 1	Exportacion 1	0.956000000	0.001333333	0.000333333	0.029000000	0.002000000	0.011333333
Exportacion 4	Exportacion 4	0.013333333	0.004333333	0.004000000	0.973333333	0.001333333	0.003666667
Exportacion 4	Exportacion 4	0.056333333	0.009666667	0.003333333	0.918666667	0.003333333	0.008666667
Exportacion 4	Exportacion 4	0.038666667	0.068000000	0.103333333	0.719666667	0.014666667	0.055666667
Exportacion 6	Exportacion 6	0.019333333	0.002666667	0.011000000	0.001000000	0.005000000	0.961000000
Exportacion 3	Exportacion 3	0.010333333	0.019666667	0.902000000	0.002333333	0.062333333	0.003333333
Nacional	Nacional	0.001333333	0.955666667	0.012000000	0.001333333	0.001000000	0.028666667
Exportacion 1	Exportacion 1	0.752666667	0.096666667	0.014666667	0.127333333	0.003333333	0.005333333
Exportacion 3	Exportacion 3	0.042333333	0.085000000	0.709000000	0.141666667	0.004666667	0.017333333
Exportacion 6	Exportacion 6	0.038000000	0.012666667	0.143666667	0.001333333	0.028000000	0.776333333
Exportacion 3	Exportacion 3	0.004333333	0.008000000	0.924000000	0.008333333	0.046666667	0.008666667
Nacional	Nacional	0.011666667	0.849000000	0.012666667	0.008000000	0.003000000	0.115666667

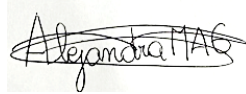
DECLARACIÓN Y AUTORIZACIÓN

Nosotras, **Alejandra Margarita Abad Garzón**, con C.C. #0930892393; **Huaxing Angela Hong Luo**, con C.C. #0924844061, autores del trabajo de titulación: **Aplicación de algoritmo de Machine Learning para clasificar la calidad de un cacao fino aroma de exportación para la exportadora Cocholato S.A.** Previo a la obtención del título de **Licenciado en Negocios Internacionales** en la Universidad Católica de Santiago de Guayaquil.

1. Declaro tener pleno conocimiento de la obligación que tienen las instituciones de educación superior, de conformidad con el Artículo 144 de la Ley Orgánica de Educación Superior, de entregar a la SENESCYT en formato digital una copia del referido trabajo de titulación para que sea integrado al Sistema Nacional de Información de la Educación Superior del Ecuador para su difusión pública respetando los derechos de autor.

2. Autorizo a la SENESCYT a tener una copia del referido trabajo de titulación, con el propósito de generar un repositorio que democratice la información respetando las políticas de propiedad intelectual vigentes.

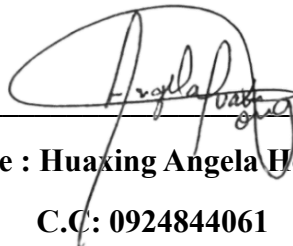
Guayaquil, 15 de febrero del 2024



f. _____

Nombre: Abad Garzón, Alejandra Margarita

C.C: 0930892393



f. _____

Nombre : Huaxing Angela Hong Luo

C.C: 0924844061

REPOSITORIO NACIONAL EN CIENCIA Y TECNOLOGÍA

FICHA DE REGISTRO DE TESIS/TRABAJO DE TITULACIÓN

TEMA Y SUBTEMA:	Aplicación de algoritmo de Machine Learning para clasificar la calidad de un cacao fino aroma de exportación para la exportadora Cocholato S. A.		
AUTOR(ES)	Abad Garzón, Alejandra Margarita; Hong Luo Huaxing Angela		
REVISOR(ES)/TUTOR(ES)	Ing. Félix Miguel, Carrera Buri, Mgs.		
INSTITUCIÓN:	Universidad Católica de Santiago de Guayaquil		
FACULTAD:	Facultad de Economía y Empresas		
CARRERA:	Negocios Internacionales		
TÍTULO OBTENIDO:	Licenciadas en Negocios Internacionales		
FECHA DE PUBLICACIÓN:	15 de febrero del 2024	No. DE PÁGINAS:	111
ÁREAS TEMÁTICAS:	Inteligencia Artificial, Inteligencia de negocios, Machine Learning, Calidad del cacao, Producción y exportación de cacao, análisis de datos e impacto en los agricultores.		
PALABRAS CLAVES/ KEYWORDS:	Palabras claves: Inteligencia de negocios, Machine learning, cacao fino de aroma, exportación, RandomForest, Calidad. / Keywords: Business intelligence, Machine learning, fine aroma cocoa, export, RandomForest, Quality.		
RESUMEN/ABSTRACT (150-250 palabras):			
<p>Resumen: En Ecuador la producción y exportación de cacao fino o de aroma representa una gran oportunidad en mercados europeos debido a la tendencia de los consumidores y su preocupación por la sostenibilidad del medioambiente. El proyecto de investigación tiene como propósito analizar la aplicación de algoritmos de machine learning para clasificar la calidad de un cacao fino aroma de exportación para la exportadora Cocholato S.A, una empresa familiar exportadora de cacao, ubicada en el cantón Rio Verde de la provincia de Esmeraldas, la clasificación del cacao es esencial para garantizar la calidad, el sabor y la consistencia de los productos de cacao, así como para facilitar la comercialización y el cumplimiento de estándares de calidad en la industria. Además, puede tener un impacto positivo en los ingresos de los agricultores de cacao y en la sostenibilidad de la producción. Cuya metodología se basó en Restudio con el algoritmo RandomForest. Los resultados arrojados son prometedores no solo para el presente sino también para el futuro sugiere que los hallazgos del estudio tienen aplicaciones prácticas y beneficios a largo plazo para los agricultores.</p> <p>Abstract: In Ecuador, the production and export of fine or aromatic cocoa represents a wonderful opportunity in European markets due to the trend of consumers and their concern for environmental sustainability. The research project aims to analyze the application of machine learning algorithms to classify the quality of a fine aroma cocoa for export for the exporter Cocholato S.A., a family cocoa exporting company, located in the Rio Verde canton of the province of Esmeraldas, the classification of cocoa is essential to ensure the quality, flavor and consistency of cocoa products, as well as to facilitate marketing and compliance with quality standards in the industry. In addition, it can have a positive impact on cocoa farmers' incomes and on the sustainability of production. Whose methodology was based on Restudio with the RandomForest algorithm. The results are promising not only for the present but also for the future, suggesting that the study's findings have practical applications and long-term benefits for farmers.</p>			
ADJUNTO PDF:	<input checked="" type="checkbox"/> SI	<input type="checkbox"/> NO	
CONTACTO CON AUTOR/ES:	Teléfono: +593 95 900 0965	E-mail: alejita.abad@hotmail.com	
CONTACTO CON LA INSTITUCIÓN (COORDINADOR DEL PROCESO UTE):	Nombre: Freire Quintero, Cesar Enrique		
	Teléfono: 593 99 009 0702		
	E-mail: cesar.freire@cu.ucsg.edu.ec		
SECCIÓN PARA USO DE BIBLIOTECA			
Nº. DE REGISTRO (en base a datos):			
Nº. DE CLASIFICACIÓN:			
DIRECCIÓN URL (tesis en la web):			