

**ALGORITMO PARA MEJORAR LA CALIDAD DEL CAFÉ COLOMBIANO Y UNA APROXIMACIÓN A
BUENAS PRÁCTICAS EN LA LOGÍSTICA DE APROVISIONAMIENTO**

**MIGUEL ÁNGEL ROJAS HIDALGO
VALERIA ANDREA GAITÁN DÍAZ**

**Proyecto integral de grado para optar al título de
INGENIERÍA INDUSTRIAL**

**Orientador
Julio Mario Daza Escorcía
Ingeniero Industrial**

**FUNDACIÓN UNIVERSIDAD DE AMÉRICA
FACULTAD DE INGENIERÍAS
PROGRAMA DE INGENIERÍA INDUSTRIAL
BOGOTÁ
2022**

NOTA DE ACEPTACIÓN

Nombre
Firma del Director

Nombre
Firma del Presidente Jurado

Nombre
Firma del Jurado

Nombre
Firma del Jurado

Bogotá D.C. febrero de 2022

DIRECTIVOS DE LA UNIVERSIDAD

Presidente de la Universidad y Rector del Claustro

Dr. Mario Posada García-Peña

Consejero Institucional

Dr. Luis Jaime Posada García-Peña

Vicerrectora Académica y de Investigaciones

Dra. Alejandra Mejía Guzmán

Vicerrector Administrativo y Financiero

Dr. Ricardo Alfonso Peñaranda Castro

Secretario General

Dr. José Luis Macías Rodríguez

Decana de la Facultad de Ingenierías

Dr. Naliny Patricia Guerra Prieto

Director del programa

Dr. Julio Aníbal Moreno Galindo

AGRACEDIMIENTOS

Agradecemos a Dios, por permitirnos vivir esta increíble experiencia del conocimiento. A nuestras familias, por confiar totalmente en nuestras habilidades y capacidades para ofrecer soluciones al mundo real desde el profesionalismo y la ética. A la universidad por aportar incondicionalmente a nuestra formación como ingenieros industriales íntegros. Y finalmente, a nosotros mismos por tanto amor, pasión, perseverancia, constancia, disciplina, disposición y autoexigencia para renovar cada día lo que somos.

Las directivas de la Universidad de América, los jurados calificadores y el cuerpo docente no son responsables por los criterios e ideas expuestas en el presente documento. Estos corresponden únicamente a los autores.

TABLA DE CONTENIDO

	pág.
1. PLANTEAMIENTO DEL PROBLEMA	12
1.1. Antecedentes	16
1.2. Justificación	19
1.3. Objetivo general	20
1.4. Objetivos específicos	20
1.5. Delimitación	20
2. METODOLOGÍA	22
2.1. Las nuevas tendencias y aplicaciones del ML en el sector cafetero	22
2.1.1. Algoritmo K-Means KNN	24
2.1.2. Máquina de vectores de soporte SVM	27
2.1.3. Redes neuronales artificiales RNN	30
2.1.4. Utilización de la espectroscopia de infrarrojo cercano NIR con diferentes algoritmos del ML	35
2.2. La cadena de suministro del sector cafetero colombiano: Diferenciales a nivel mundial y una mirada desde los efectos del COVID-19	38
2.2.1. Elementos principales de la logística de aprovisionamiento en el sector cafetero colombiano	41
2.2.2. Visualización del modelo SCOR desde la perspectiva de los proveedores	49
2.2.3. Parámetros para clasificar un café colombiano como especial o excelso	54
2.2.4. Factores con mayor impacto en la calidad del café según Total Cup Points	58
2.3. Metodología y datos	59
2.3.1. Selección de Variables	61
2.3.2. Lenguaje de programación e importación de módulos y base de datos	62
2.3.3. Análisis exploratorio de datos	63
2.3.4. Limpieza de datos	64
2.3.5. Diagramas de cajas y bigotes e histogramas	65
2.3.6. Diagrama de Violín	69
2.3.7. Data de entrenamiento y data de prueba	71
2.3.8. Modelo Predictivo	71

2.3.9. Resultados y validación del modelo predictivo.	75
2.3.10. Factores de alto impacto en la LA a partir de los lineamientos de calidad del café según el total cup points de la data en estudio	75
2.3.11. Recomendaciones selección y evaluación de proveedores	77
3. MARCO REFERENCIAL	80
3.1. Marco conceptual	80
3.1.1. Calidad de Café	80
3.1.2. Algoritmo	81
3.1.3. Inteligencia artificial	81
3.1.4. Cadena de suministros	84
3.1.5. Productividad	88
3.2. Marco Teórico	89
3.2.1. Producción y calificación del café	89
3.2.2. Métodos de aplicación del ML	92
3.2.3. Cadena de suministro del sector cafetero colombiano	100
3.3. Marco Histórico	102
3.3.1. Evolución de la IA y el ML	102
3.4. Marco Normativo	105
4. RESULTADOS	109
5. CONCLUSIONES	111
BIBLIOGRAFÍA	113

LISTA DE FIGURAS

	pág.
Figura 1. Herramienta causal – Diagrama Ishikawa 8M	15
Figura 2. Efectos del cambio climático en regiones agrícolas a nivel mundial	23
Figura 3. Plano de iteraciones k-means: agrupamiento con la minería de datos	25
Figura 4. Interacción del algoritmo SVM con los datos suministrados	28
Figura 5. Asignación de pesos de neuronas que realizaron sinapsis entre ellas	32
Figura 6. Afectación covid-19 en operaciones intermediarias de la CS del café	40
Figura 7. Flujo de información de la logística de aprovisionamiento	42
Figura 8. Descripción del flujo de información de la LA del SC colombiano	43
Figura 9. Representación gráfica del modelo SCOR	53
Figura 10. Requisitos del café verde en almendra excelso dictados por la FNC	56
Figura 11. Matriz de correlación de las variables evaluadas en los Cup Points	59
Figura 12. Diagrama del proceso de tratamiento del dataset para aplicar el ML	61
Figura 13. Resultado de la línea de código df.info()	64
Figura 14. Histogramas de las variables de interés	65
Figura 15. Diagramas cajas y bigotes de las variables más influyentes en el estudio	67
Figura 16. Diagrama de cajas y bigotes con el método Tukey	68
Figura 17. Histogramas consolidados una vez implementado el método Tukey	69
Figura 18. Diagrama de violín. Total cup points vs. Flavor (Sabor)	70
Figura 19. Diagrama de violín. Total cup points vs. Sweetness (Dulzor)	71
Figura 20. Gráfica Flavor vs. total cup points	74
Figura 21. Recomendaciones para la LA desde la prolongación y percepción sabor	78
Figura 22. Procedimiento para la aplicación del ML	84
Figura 23. Actores principales de la cadena de suministros	86
Figura 24. Diferenciación entre Clasificación-Regresión según el algoritmo esperado	95
Figura 25. Esquema básico de los componentes del RL	97
Figura 26. Red de cinco capas adaptada por Nielsen	99
Figura 27. Caracterización de la cadena de abastecimiento local de café	101
Figura 28. Hitos del ML a nivel mundial	103

RESUMEN

Colombia se posiciona como uno de los tres productores de café más importantes a nivel mundial. Por lo tanto, es un sector económico primordial para la economía del país. Sin embargo, en los últimos 10 años ha disminuido su productividad y calidad del grano debido a la migración de mano de obra a otros sectores, arraigo a las prácticas tradicionales, baja inversión en desarrollos tecnológicos y altos costos de producción.

En consecuencia, en el presente estudio se utilizó el *machine learning ML* a partir de data de acceso libre dispuesta por la *CQI* con más de 52.000 registros, acotados por la variedad arábica y países con condiciones similares a Colombia, manipulada a través de *Python*. Inicialmente, se exploró sobre las nuevas tendencias del *ML* en la agricultura y caficultura, en donde el *algoritmo k-means* y *neural networks* son los más utilizados. Seguidamente, se caracterizó la logística de aprovisionamiento *LA* del sector, la cual arrojó los factores más influyentes en la calidad del café y la siendo estos el almacenamiento, adecuada fertilización y buenas prácticas en los cultivos. Finalmente, se utilizó el algoritmo de *ML* de regresión lineal bajo el estudio de las variables evaluadas en la catación de una taza referente a la calidad, como aroma, sabor, cuerpo, acidez, *after taste*; con un porcentaje de confiabilidad del 91.94% tanto el sabor como el *after taste* son en más del 80% representativas en la calificación total de calidad del café. Finalmente, se plantean diversas recomendaciones para el caficultor en su *LA* para preservar dichas propiedades.

PALABRAS CLAVE: *Machine learning*, logística de entrada, cadena de abastecimiento, calidad del café, productividad, regresión lineal

INTRODUCCIÓN

El sector cafetero SC en Colombia surge como principal sustento económico de la agricultura y desde entonces ha sido un punto de referencia de calidad a nivel mundial, compitiendo con grandes exponentes del café excelso como Brasil, Vietnam y Malasia. Sin embargo, el SC ha perdido competitividad, en gran parte por la denominada transformación digital la cual contempla como eje principal la implementación de la inteligencia artificial en el marco del *machine learning*. Gran variedad de tareas o actividades se han automatizado, como lo es la selección y recolección autónoma de granos de café maduros. Por otra parte, los caficultores se ven fuertemente afectados entre muchos otros aspectos; en términos de logística de aprovisionamiento, ya que en gran cantidad de fincas existe un fuerte arraigo a las tradiciones, las cuales son en algunas ocasiones inquebrantables, llegando al punto de evitar la actualización de máquinas, procesos o el uso de nuevo productos en cultivos, lo que desencadena repercusiones negativas en la economía de los agricultores.

Por todo lo anterior, el presente trabajo de investigación busca implementar el *ML* como aquella herramienta de la *IA* que permita mejorar las prácticas en la logística de aprovisionamiento *LA* de los caficultores colombianos y así obtener una excelente calidad del grano. Inicialmente, se identifican aquellas aplicaciones del *ML* tanto en la agricultura como la caficultura a partir del año 2019 con la finalidad de reconocer los algoritmos de *ML* en tendencia y visualizar las problemáticas en las cuales el *ML* obtuvo un excelente rendimiento. Luego de esto, se caracteriza en rasgos generales la *LA* del SC colombiano por medio el abordaje de la NTC 5181 que establece las buenas prácticas de manufactura para la industria cafetera [1]. Además, a través de la información recopilada en fuentes confiables como la *FNC* y *cenicafé*, se construye un diagrama del flujo de información y el modelo *SCOR*; esto con el objetivo de detectar aquellos factores con mayor impacto en la *LA*.

Para enlazar el modelo de *ML* y la *LA* bajo la perspectiva de la calidad del café, se expone la *resolución 02 del 2016* [2] la cual consolida los *lineamientos que debe cumplir el grano verde para ser considerado de alta calidad y así mismo se retoman los factores*

más influyentes en la LA. Finalmente, el algoritmo se desarrolla en el reconocido lenguaje de programación *Python*, a partir de una base de datos dispuesta por el *CQI* en la plataforma *Kaggle*. El resultado del algoritmo de aprendizaje supervisado de regresión lineal obtuvo una calificación mayor al 80% determinando que las dos variables que más influyen al momento de determinar la calidad de una taza de café son el sabor y la prolongación de sabor del café; para la base de datos estudiada. Siendo un punto de partida para enfocar el proyecto desde la *LA* sin olvidar estas dos características claves en la calidad de un café excelso por medio de la consolidación de recomendaciones fehacientes para evaluar factores en los proveedores que impacten directamente la prolongación y sabor del café.

1. PLANTEAMIENTO DEL PROBLEMA

El café es el producto de producción agrícola con mayor importancia para la economía colombiana y no por nada, el año cafetero 2019 cerró con la exportación de 13.5 millones de sacos de 60 kg de café verde. Lo que vincula laboralmente a más de 478.344 hogares caficultores en el territorio colombiano. Aun así, no todo es color de rosa, y más del 55.9% de estas familias no cuentan con sistema de acueducto en sus hogares. Cerca del 56.9% de ellas tienen un nivel educativo de primaria, 27.7% con escuela secundaria, 12.6% no cuentan con ningún tipo de escolaridad y tan solo el 2.8% accede a los programas de educación superior [3]. Bajos niveles educativos y pobreza castigan y debilitan al SC, cediendo participación económica frente a otros sectores como la floricultura, el constructor y energético. En adición, la falta de planeación de los pequeños caficultores impacta de forma negativa su capacidad de reaccionar frente a escenario imprevistos. En el mismo ámbito, es claro que el SC carece vertiginosamente de investigación, desarrollo e innovación en gran medida por el fuerte arraigo hacia las tradiciones heredadas de generación en generación.

Otro factor externo que ha contribuido a la baja productividad del sector ha sido la contundente falta de modernización de las entidades cafeteras a nivel nacional para adoptar nuevas prácticas que permitan adaptarse a las tendencias de la economía mundial a partir de la aplicación de la tecnología [4]. A simple vista se observa cómo loscafeteros colombianos CC actualmente continúan realizando operaciones rutinarias dentro del ciclo productivo del café de forma totalmente arcaica. Esto es aún más evidente al comparar el SC nacional con el SC brasilero, quienes cuentan con un nivel de automatización mucho más avanzado, logrando reducir costos, tiempos no productivos, errores humanos y en general recursos que al final del día castigan a los CC, desacelerando el crecimiento del sector y su valioso impacto como motor en la economía colombiana. Demostrando y reafirmando así que con la implementación netamente de sistemas robotizados no es suficiente para generar mejoras en la productividad ya que estos se basan en la mecanización. Mientras que, las aplicaciones de la IA permiten la generación de procesos con cierto grado de

razonamiento al desarrollar las actividades productivas y de cierta manera, asegurando una retroalimentación en tiempo real sobre las posibles fallas. Es en este punto, donde el *ML* entra a jugar un papel crucial a través de los modelos que brindan monitoreo y predicción sobre el desempeño, los cuales se ajustan a los *dos desafíos claves del SC colombiano: los espacios heterogéneos y la variabilidad temporal* [5].

Si bien es cierto que algunas instituciones como la *Corporación Nacional de investigación Agropecuaria Corpoica*, desde el año 2019 vienen trabajando de la mano con el análisis de datos y la *IA* para la observación de las condiciones de la tierra y así llevar a cabo los cultivos, es preciso mencionar que aún existen falencias en cuanto a la investigación y fortalecimiento de herramientas tecnológicas que efectivamente agreguen valor a la cadena de suministro del *SC* en Colombia [6].

En lo que corresponde a la mano de obra, en Colombia más del 40% de los trabajadores del gremio se encuentran en el rango de los 29 y los 59 años [3], lo que de cierta forma retira inmediatamente el interés por parte de las nuevas generaciones de participar activamente en la caficultura, además dichos jóvenes cuentan con la percepción de que los trabajos u operaciones desarrolladas son monótonas o repetitivas, y peor aún sin mucha proyección a futuro. Si a lo mencionado anteriormente se le adiciona salarios que en ocasiones son poco representativos frente a otros sectores económicos, se tiene un panorama preocupante frente al futuro del sector.

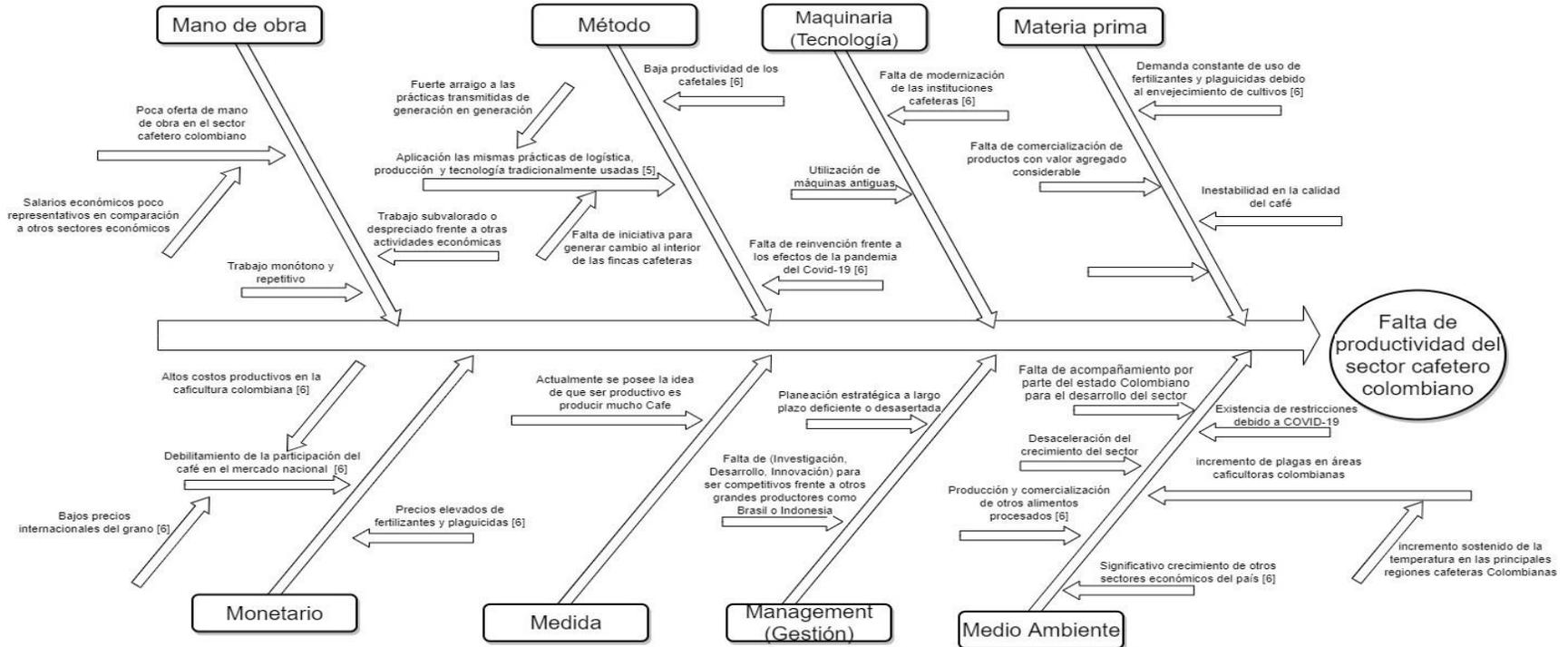
Un factor que ha sido mencionado a lo largo del gremio es que desafortunadamente los *CC*, en gran medida por su falta de innovación y planeación estratégica, no han generado valor en sus productos. En palabras prácticas, producen un café excelso, ellos no perciben el grueso de los beneficios económicos de la actividad. Mientras que organizaciones internacionales si logran capitalizar las ganancias por medio de la adición de valor al producto comercializado. Algunos ejemplos de productos de valor a partir del café son: licor, cerveza, postres, jabones, hilos e inclusive como elementos decorativos a base de café, entre muchos otros.

Para realizar el planteamiento del problema, se utilizó la herramienta causal denominada Diagrama *Ishikawa* desde el enfoque 8M's la cual abarca: Mano de obra, método, maquinaria, materia prima, monetario, método, medida, *management* (gestión) y medio ambiente. Se seleccionó este tipo de diagrama *Ishikawa* debido a que permite identificar la causa raíz del problema a partir de las diferentes categorías manejadas por el modelo y a su vez, visualizar concretamente las causas y los efectos para su posterior análisis, tal y como se observa en la Figura 1.

Acorde con esto se planteó conjuntamente un problema, el cual se denominó **FALTA DE PRODUCTIVIDAD EN EL SECTOR CAFETERO COLOMBIANO** y de la misma forma se definió como pregunta de investigación **¿En qué medida los algoritmos del *ML* son aplicables en el sector cafetero dentro del marco de la logística de aprovisionamiento en Colombia?**

Figura 1.

Herramienta causal - Diagrama Ishikawa 8M



Nota. Se observa el desarrollo detallado de la herramienta causal Diagrama de Ishikawa 8 M. Tomado de: «La innovación llega al sector cafetero,» Siglodata.co, 2021. Available: <http://www.siglodata.co/s/observatorio/20181123LATINPYMEINNPULSA.pdf> , O. Cerquera, «El acompañamiento institucional en el desarrollo del sector cafetero colombiano,», Scielo, 2014. <http://www.scielo.org.co/pdf/fype/v7n1/v7n1a09.pdf>, «FNC se reúne con cafeteros de Antioquia, Caldas, Quindío y Risaralda para analizar efectos de covid-19 y retos del sector - Federación Nacional de Cafeteros,» <https://federaciondecafeteros.org/wp/listadonoticias/fnc-se-reune-con-cafeteros-de-antioquia-caldas-quindio-y-risaralda-para-analizar-efectos-de-covid-19-y-retos-del-sector/>

1.1. Antecedentes

El SC se ha consolidado como carta de presentación colombiana a nivel nacional e internacional. Sin embargo, ha sido una carrera larga y llena de magnos retos a los cuales se les han sorteado soluciones determinantes para continuar a la vanguardia, compitiendo mano a mano contra fuertes competidores como Brasil, Vietnam y Malasia; pioneros de talla mundial en el sector caficultor.

Dentro de los múltiples estudios e investigaciones realizados en el marco del SC colombiano y su lucha por ser cada vez más eficientes, es necesario resaltar la publicación realizada en el año 2014, por los másteres en economía Cerquera Losada y Orjuela [8]. Quienes revisaron detalladamente desde el origen, la evolución, posterior consolidación y la pérdida de energía del SC colombiano, el cual se vio fuertemente apalancado en su momento por la Federación Nacional de Cafeteros *FNC* fundada en 1927. En el documento se expone y describe la bonanza cafetera experimentada a partir del año 1930, en donde el café colombiano se consolidó como punto de referencia a nivel internacional, mientras actuaba como motor económico y suministraba fuertes ingresos para el desarrollo industrial colombiano, hasta llegar al punto de intervenir fuertemente sobre intereses políticos de la época, mostrando el nivel de importancia que poseía el grano.

No obstante, el SC se vio fuertemente afectado dentro de otras muchas razones por la violencia experimentada a nivel nacional por las profundas corrientes políticas influenciadas y la imborrable lucha contra el narcotráfico llevada a cabo a lo largo del territorio a partir de 1980. En adición, se presentó en paralelo un estancamiento tecnológico del sector alimentado por el envejecimiento de los cafeteros, lo que sumado a la falta de educación terminó en una inevitable pérdida de productividad y competitividad frente a sus principales rivales como lo es Brasil; quien aprovechó para ratificar su posición dominante como principal productor y exportador de café global.

El mundo está cambiando a pasos agigantados y parte de ello se debe a la cuarta

revolución industrial la cual se afectó fuertemente, energizada por la pandemia del *covid-19*; en donde la sociedad se encontró en la obligación de echar mano de las herramientas digitales.

Como punto focal se encuentra el *machine learning ML* o más conocido como *IA* según lo expuesto en el año 2020 por el economista egipcio *Heidy Ali* [10], quien estudió la relación entre la transformación digital y el desarrollo económico de la mano de la productividad laboral en economías emergentes. Como se esperaba, el artículo mostró positivas y estrechas correlaciones en el marco de la transformación digital por medio del Índice de evaluación digital *DEI*, y el desarrollo económico, describiendo una interesante conclusión en donde el sexo femenino, como actor protagonista toma mayor provecho de la transformación digital sobre el sexo masculino. Es entonces, en donde gana mayor interés la aplicación del *ML* en el *SC* colombiano; el cual surge, no sólo como herramienta de productividad en el sector agro, sino como herramienta de equidad de género, brindando las mismas oportunidades, condiciones y beneficios tanto para mujeres como para hombres. Actualmente, la balanza de equidad de género se ve levemente inclinada hacia los hombres quienes predominan en los hogares cafeteros con un 54.7% frente a las mujeres quienes ocupan el 45.9% según el informe de la *FNC* lanzado en diciembre de 2020 [3].

Por otra parte, la logística de aprovisionamiento *LA* ha adquirido inmensurable importancia sin importar el sector económico que se quiera abordar. Actualmente, bajo la reciente crisis económica originada por la pandemia del *covid-19*, la *LA* se ha exigido al límite en un entorno en el cual los márgenes económicos de las industrias se han vuelto cada vez más estrechos, obligando a optimizar recursos, reducir costos, tiempos de entrega, tiempos de producción y en general creando nuevas tendencias en el complejo mundo de la cadena de aprovisionamiento. En el marco del *SC*, llama la atención el documento publicado por la Organización de Naciones Unidas de Agricultura y Alimentos en el año 2020 [11], el cual realiza una revisión detallada de la mano de obra en el *SC* de Uganda, abordando la problemática generada por los altos índices de jóvenes y niños menores de edad que trabajan bajo ningún tipo de aseguramiento laboral

o social en fincas cafeteras por periodos vacacionales a cambio de bajas remuneraciones económicas. En gran medida los menores de edad son obligados a sumergirse en el sector debido a la falta de oportunidades y la fuerte crisis económica que atraviesa el país africano. En el estudio se evalúan los impactos del fenómeno y las pocas oportunidades que existen para los jóvenes de Uganda quienes buscan en el café una forma de independencia económica.

Entrando en materia, algunas de las aplicaciones que ha tenido el *ML* en el sector caficultor a nivel mundial han sido bastantes representativas ya que han aportado nuevas formas de realizar varios procesos y al mismo tiempo, contribuir al desarrollo de las nuevas tecnologías para el beneficio de los caficultores. En un estudio de caso realizado por William R. Eustaquio y *Jesusimo L. Dioses Jr.* [12], se analizó la práctica de cosechar todas las semillas del café cuando el 75% de las mismas se encuentran totalmente rojas en un cultivo, originando granos de café inmaduros y aportando así la disminución de la calidad del grano. Por esta razón, estos dos profesionales especializados en el desarrollo de tecnologías, realizaron un aplicativo *Classification Learner* de *Matlab*, en el cual usaron 23 clasificadores de *ML*, desde el levantamiento de información en campo de una cosecha comprendida en el periodo de enero a marzo de 2019, exactamente del café Robusta dividida en dos categorías: granos de café maduros e inmaduros. Esta información permitió el procesamiento de imágenes con el fin de extraer características puntuales de los granos de café en estudio por medio de la herramienta color *RGB*, por sus siglas en ingles *red, green, blue*. Obteniendo como resultado, en términos de precisión y velocidad en el entrenamiento un valor del 94% y 0,62 segundos respectivamente.

Es importante resaltar que la estadística es una herramienta fundamental para la predicción de las condiciones del entorno productivo de cualquier sector económico ya que, por medio del análisis, se pueden asumir y planificar recursos necesarios para responder satisfactoriamente en todos los ámbitos de las organizaciones. Haciendo uso de la estadística y el aprendizaje automático, la Universidad de Ciencia y Tecnología de *Wroclaw*, Polonia [13] realizó una evaluación de factores como el medio

ambiente, procesamiento postcosecha y el genotipo para determinar en qué medida pueden afectar la cadena de producción del café y así mismo identificar rasgos netamente agronómicos que conduzcan a la calidad del mismo teniendo en cuenta la variación del entorno. A lo largo del estudio, se utilizó data en campo de cultivos de café ubicados en tres municipios de la región *Matas de Minas*, del Estado de *Minas Gerais*, en Brasil; evaluando el procesamiento natural y despulpado correspondiente a la postcosecha. Al aplicar los métodos, se obtuvo que los más eficientes fueron las redes de perceptrón multicapa y función de base radial ya que estos 2 métodos arrojaron la discriminación satisfactoria de las muestras de café en varios entornos. Consecuentemente, el tamaño del fruto, altura de la planta y la producción de frijol son los principales rasgos influyentes y responsables de la diferenciación de los lugares de producción.

1.2. Justificación

La creciente carrera mundial por liderar la producción y exportación del mercado cafetero lleva a una constante reinención de la industria, en la cual Colombia se ha caracterizado como uno de los actores principales. Sin embargo, a lo largo del siglo XXI, el liderato en la producción y exportación del grano tiene nombre propio, y es que, con una producción anual en 2019 de 59 millones de sacos de 60 kg de café verde, Brasil ejerce gran presión como agente dominante en el mercado cafetero. El SC se vuelve inclusive más llamativo con los recientes pronósticos de demanda del grano, en donde se estima que la demanda del café ronde alrededor de los 163.3 millones de sacos de 60 kg de café verde para el año 2021 [3].

Es necesario tener claridad sobre los históricos obstáculos que se encuentran a lo largo de la cadena productiva del café, dentro de los que destacan: Incremento sostenido de la temperatura global, falta de acompañamiento por parte del estado colombiano hacia el sector, bajos niveles de educación en los principales productores, falta de planeación a largo plazo por partes de los caficultores colombianos, falta de investigación, desarrollo e innovación. En adición se encuentran los elevados costos de los múltiples insumos de las fincas cafeteras como lo son los herbicidas, plaguicidas y

funguicidas, entre otros. Por otra parte, preocupa la disminución de la mano de obra del sector, y la falta de motivación de los jóvenes colombianos por incursionar en el mercado cafetero el cual ha llegado a ser un sector de labores monótonas y repetitivas.

Las oportunidades de mejora mencionadas anteriormente son el motor del presente proyecto el cual se aborda desde la perspectiva de la ingeniería industrial, integrando algoritmos y la *LA*, en adición con las nuevas tendencias *ML* para mejorar las condiciones del *SC* y por ende su competitividad.

1.3. Objetivo general

Proponer un algoritmo de *machine learning* aplicado a la logística de aprovisionamiento del café arábica en el sector cafetero colombiano con la finalidad de identificar las variables que influyen en la producción del café con calificación 10/10.

1.4. Objetivos específicos

- Identificar las nuevas tendencias del *ML* por medio de fuentes secundarias con el fin de identificar cuáles son aplicables en el objeto de estudio.
- Explorar los lineamientos y elementos principales de la logística de aprovisionamiento por medio de la búsqueda de información, para reconocer los factores con mayor impacto dentro de la cadena de abastecimiento del sector cafetero colombiano
- Plantear un algoritmo de *ML* en el marco de la logística de aprovisionamiento por medio de la utilización de herramientas tecnológicas, que permita estructurar recomendaciones para mejorar la competitividad y calidad del café arábico cultivado en Colombia

1.5. Delimitación

El proyecto contempla el proceso exploratorio de algoritmos de *ML* aplicables a la mejora de la *LA* para el *SC* colombiano enfocado en la variedad arábica, desde una perspectiva innovadora, como lo son los algoritmos de aprendizaje supervisado y no supervisado, claros exponentes dentro de la popular transformación digital la cual ha impactado no solo al sector agro, sino a todos los sectores económicos, en gran medida

por la reciente pandemia del *covid-19*. Esta coyuntura ha brindado terreno para nuevos estudios o proyectos de grado en el marco de la *IA* aplicada al sector agro colombiano por medio del desarrollo de una de las ramas de la ingeniería industrial.

El algoritmo seleccionado se alimentará de bases de datos reales a nivel mundial y dará como resultado una propuesta de mejora en la logística de abastecimiento ajustada a la realidad de los caficultores colombianos quienes en la actualidad han perdido productividad y con ella participación en el mercado, impactando drásticamente el bienestar de la economía de la nación. En el marco de la *LA* se emplean nuevas tendencias que brindan competitividad al *SC* colombiano.

El proyecto se llevará a cabo en conjunto por parte de los dos investigadores en sus respectivos lugares de residencia. La duración estimada del trabajo investigativo es de cuatro meses a partir de la aprobación del anteproyecto. Finalmente, la Fundación Universidad de América se responsabiliza de brindar las asesorías con el personal especializado en los campos técnicos según sean requeridos a lo largo del proyecto y de facilitar el acceso a las bibliotecas de los campus y a todo el material bibliográfico existente en los repositorios que posean convenios vigentes con la universidad.

2. METODOLOGÍA

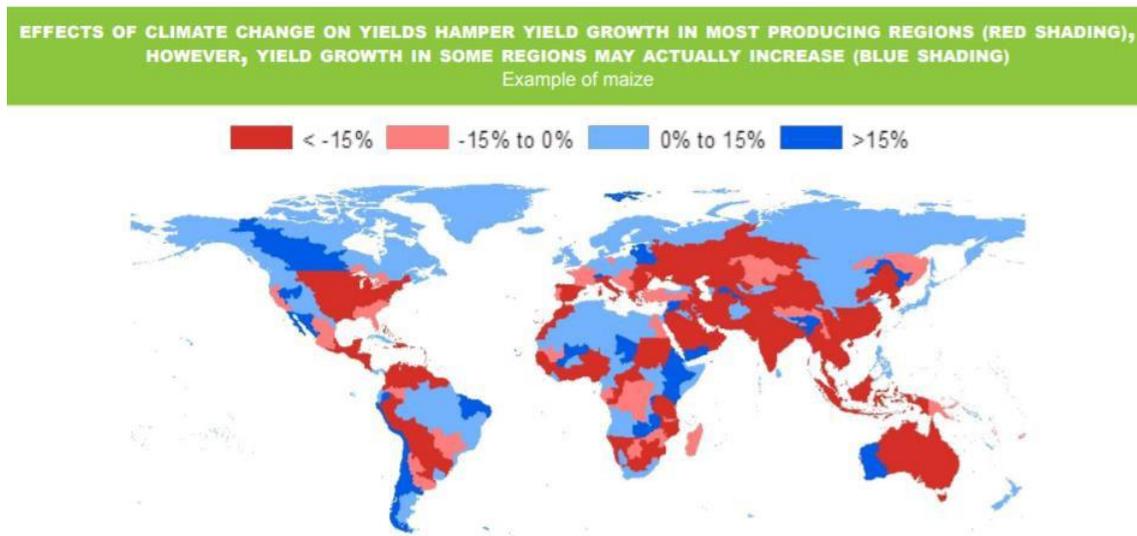
2.1. Las nuevas tendencias y aplicaciones del *ML* en el sector cafetero

Debido a los grandes cambios que genera la evolución de la tecnología y cómo los nuevos mercados incorporan todo este tipo de herramientas para predecir parámetros esenciales que aumenten su calidad y productividad simultáneamente, se visualiza el gran auge que ha tomado el *ML* en todo tipo de industrias.

Puntualizando en el sector al cual atañe la presente investigación, actualmente la industria agrícola presenta inestabilidad en los cultivos debido al cambio climático que impacta negativamente los componentes del subsuelo. Lo más preocupante ante esta problemática es el alza considerable de la temperatura y variaciones de las precipitaciones. Como bien se sabe, el cambio climático no tiene límites, sino que, por el contrario, con el paso del tiempo la frecuencia e intensidad de los fenómenos meteorológicos extremos aumentan las presiones sobre los sistemas agrícolas y alimentarios mundiales [14]. En el caso de Colombia, según la *OECD* los efectos del cambio climático actualmente disminuyen el rendimiento de la producción agrícola, como lo muestra la Figura 2.

Figura 2.

Efectos del cambio climático en regiones agrícolas a nivel mundial



Nota. Gráfico mundial sobre el impacto del cambio climático. *Tomado de:* OECD / IFPRI, *Modelling Adaptation to Climate Change in Agriculture, OECD Food, Agriculture and Fisheries Papers*, No. 70, 2014 <http://dx.doi.org/10.1787/5jxrclljnxbxq-en>.

Colombia se verá afectado en el rendimiento agrícola con la disminución del mismo en un 15%. Es ahí, donde el SC decide investigar sobre nuevas alternativas tecnológicas que involucren modelos de predicción para aproximarse a la realidad y así transformar sus cadenas de suministros a fin de obtener resultados fehacientes alineados a su objetivo principal: Producir café con los índices más altos de calidad y competir con esta insignia para mantenerse como uno de los países más representativos en la producción del café a nivel mundial.

Dentro de la exploración de las nuevas tendencias de los algoritmos del ML, se observa que la mayoría de aplicaciones se concentran en agrupar información en categorías o sectores a partir de un propósito inicialmente definido. Si bien es cierto que el mundo del ML incluye algoritmos que satisfacen los diferentes enfoques de las problemáticas y son flexibles a la hora de realizar la programación, varios casos de

estudio enlazan diversos algoritmos con la finalidad de analizar las distintas resoluciones y así elegir la solución óptima.

Para abordar cada código de *ML* en tendencia, a continuación, se exponen la variedad de algoritmos utilizados en el sector agrícola y cafetero, mencionando puntualmente en cuáles problemáticas han sido empleados.

2.1.1. Algoritmo K-Means KNN

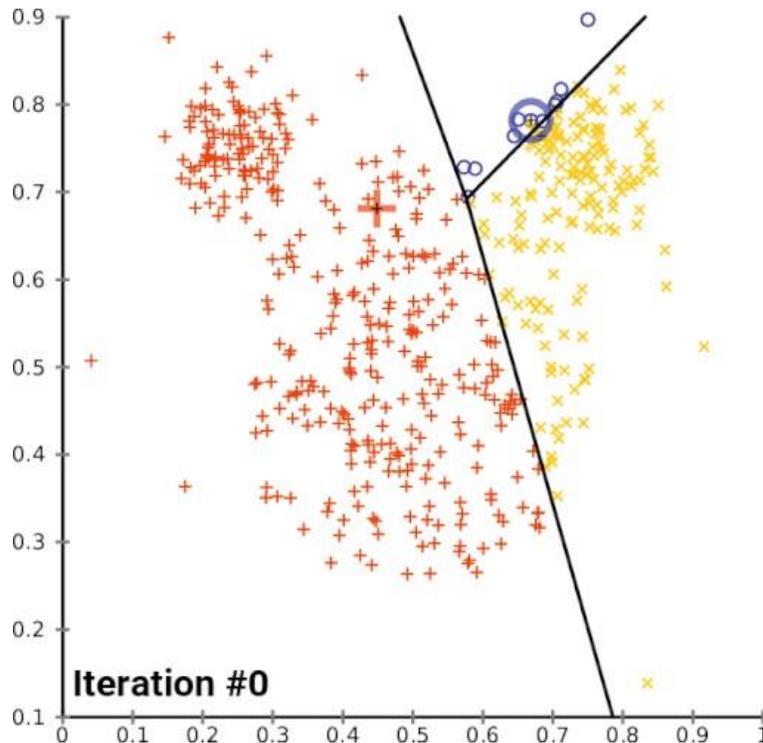
Hace parte de las aplicaciones de aprendizaje no supervisado en la categoría de *clustering* debido a que los datos de entrada no tienen ningún tipo de etiqueta. Por lo tanto, al entrar en funcionamiento el algoritmo agrupa los datos bajo la definición de un valor *k* que contiene ciertas características que el usuario previamente define. Una vez se tiene el valor *k*, se evalúa la similitud a partir de una función definida por la distancia a la cual se encuentra cierto dato en comparación al valor de *k*. Esta distancia puede ser de tipo euclidiana, *manhattan*, *minkowski* y *hamming*. Esta última función de *hamming* está basada a partir de variables categóricas, mientras que las primeras tres son funciones continuas. Comienza el programa a ejecutar las iteraciones hasta encontrar los grupos óptimos de datos que cumplan con las características planteadas y estén lo más cercanas posibles al valor *k*.

En adición, el modelo tiene un alto grado de exactitud cuándo maneja *dataset* relativamente medianas o pequeñas. [15].

Al ejecutar las respectivas iteraciones, arroja un plano similar al siguiente:

Figura 3.

Plano de iteraciones k-means: Agrupamiento con la minería de datos



Nota. Visualización de las iteraciones realizadas por el algoritmo *k-means* a partir de un *k* definido. Gráfico tomado de: *K-Means: Agrupamiento con Minería de datos [Introducción]*", ESTRATEGIAS DE TRADING, 2021. <https://estrategiastrading.com/k-means/>

A nivel de la industria agrícola, es pertinente mencionar el estudio realizado en cultivos de maíz, pues bien es sabido que este tipo de producto es uno de los más distribuidos en el mundo. Sin embargo, la cosecha es bastante vulnerable a la adquisición de hongos. Una de ellas y la más frecuente es la reproducción de aflatoxinas. En pocas palabras, este hongo es altamente cancerígeno, afectando la salud del ser humano y los animales.

Para implementar el algoritmo *k-means*, inicialmente se estableció una base de datos que contiene 600 bandas hiper espectrales equivalentes a un pixel, a los cuales se les asignó etiquetas con valor de contaminado o puro. Luego de esto, se agruparon las bandas con ciertas características y, al evaluar este grupo bajo el algoritmo arrojó una precisión del 98,77% y sin data que estuviera previamente agrupada, la probabilidad de precisión fue aún más alta [16].

Por otra parte, en los huevos se encuentra frecuentemente la patógena salmonela considerada como un grave problema de salud pública. Adicional a esto, es causante de diferentes brotes en la piel. Ante esta eventualidad, la creación de numerosos prototipos basados en el principio de detección inmediata del microbio ha venido en aumento a través de los años. Este estudio proporcionó un dispositivo móvil portátil con el espectro de detección de imágenes más amplio y con mayor intensidad de luz con respecto a los actuales en el mercado. El *k-means KNN* fue alimentado por varias imágenes de prueba y simultáneamente con espacios de color a fin de ser clasificados. Utilizando el algoritmo *KNN* y *SVM*, y espacio de color *L by RGB*, el modelo logró una exactitud del 95.56%. Además en pruebas experimentales se comprobó que la utilización del *ML* tuvo menos errores que la inspección visual [17].

Ya que se han abordado algunos estudios significativos del algoritmo a nivel agrícola, es vital indagar sobre qué aplicaciones ha tenido en la caficultura.

El *KNN* usualmente es implementado en la detección de plagas y enfermedades que afectan los cultivos. Una de las plagas más representativas debido a su incidencia y fuertes alteraciones que causa al sabor del grano es la plaga barrenador de café. Son una especie de larvas que taladran la parte central del tronco de los árboles, en las raíces o los tallos a la altura del nivel del suelo [18]. Por lo tanto, Lucas de *Oliveira* y José Cabral, miembros de *Science and Technology of Mato Grosso do Sul* de Brasil utilizaron este código para predecir la incidencia de esta plaga en los cultivos de la variedad arábica ubicados en el estado *Minas Gerais*, el cual al ser comparado con los bosques aleatorios *RFT*, el *k-means* obtuvo el rendimiento más bajo. Aunque no superó el rendimiento de los *RFT* su desempeño durante de las iteraciones generadas cumplió con la expectativa

de los investigadores, los cuales sugieren la combinación de varios algoritmos de *ML* para obtener evidencias reales y exitosas que satisfagan la problemática [19].

No siendo suficiente con la propagación de plagas y enfermedades, el sector ha venido luchando contra la falsa autenticación del café molido en el mercado, pues existen organizaciones ilícitas que se dedican a la suplantación de su origen. Por ello, expertos en el análisis de los entornos alimenticios de China ejecutaron un estudio que abarcó la observación detallada de los isótopos estables que puede tener un grano cultivado en tierras brasileñas, categorizándolo según el sistema orgánico o convencional. De acuerdo al isótopo, se establecieron ciertos parámetros que delimitaran de cierta forma la detección de categoría. Utilizaron varios algoritmos de *ML*, entre ellos el algoritmo en cuestión, el cual obtuvo el mejor rendimiento, llegando a clasificar el café entre orgánicoo convencional en un 87% y 67% de precisión [20].

Por todo lo anterior, el *KNN* es bastante versátil por su adaptabilidad a los requerimientos del usuario. Un factor importante es la correcta elección del valor de *k* ya que si este valor se encuentra desfasado o no presenta mayor coherencia, al momento de iterar con los datos de prueba difícilmente arrojará resultados confiables.

En conclusión, las aplicaciones mencionadas a nivel agrícola se consideran pertinentes para ser implementadas en el *SC* ya que su enfoque de solución de enfermedades y la usabilidad de data en imágenes es muy viable para ser aplicable, por ejemplo, en la detección prematura de la roya en las hojas del cafeto.

2.1.2. Máquina de vectores de soporte SVM

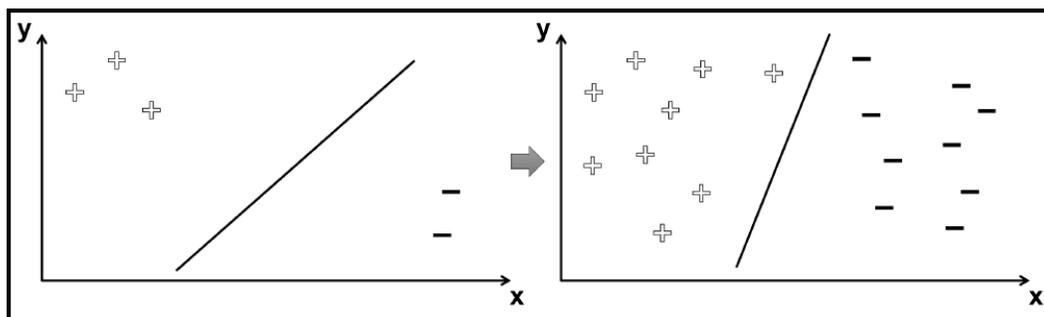
Este método de aplicación del aprendizaje no supervisado parte de la suposición y definición de dos límites para tomar decisiones a partir de la data que no contiene ningún tipo de etiqueta. Aunque sigue la filosofía de clasificación, su enfoque principal es maximizar los límites de decisión teniendo en cuenta la ubicación de los datos.

Antes de profundizar en las aplicaciones puntuales en el sector agrícola y cafetero, es fundamental comprender puntualmente el comportamiento de este método.

Se plantea una función que depende de dos variables A y B, donde estos valores deberán ser clasificados en dos etiquetas: + y -. Al definir únicamente dos posibles etiquetas de los datos, es un problema de clasificación binaria. En la siguiente imagen, se expone gráficamente el funcionamiento del algoritmo.

Figura 4.

Interacción del algoritmo SVM con los datos suministrados



Nota. Explicación explícita sobre la interacción de los datos, etiquetas y el límite de decisión. Tomado de: *M. Gopal, Applied Machine Learning. McGraw-Hill, 2019.*

En el plano cartesiano del costado izquierdo se puede observar cómo el *K-means* al ingresar progresivamente los datos de entrenamiento actualiza el límite de decisión según las etiquetas. Finalmente, establece un límite que clasifique óptimamente la data, minimizando el error en la precisión. Cabe aclarar que los vectores de soporte para delimitar el conjunto de datos son aquellos puntos que están ubicados en los márgenes de cada categoría (etiqueta) definida [21].

Ahora bien, teniendo claro el desarrollo que conlleva la implementación de este tipo de *ML* se pueden abordar aquellos casos de estudio en la agricultura y caficultura.

A nivel del sector agrícola la carrera por descubrir soluciones que mejoren el pronóstico del comportamiento de los genes en los animales y plantas, se realizó un estudio en los cerdos y el maíz. Como primera fase del estudio, se define una función con respecto al grano del maíz y con base en el modelo *SVM*, se establecieron los hiper parámetros alineados hacia el propósito del algoritmo que tuvo como entrada ocho conjuntos de datos sobre los genes de los cerdos y el maíz. Luego, para aplicar estructuras algorítmicas similares y posteriormente realizar una comparación de los resultados, se confrontaron la unión del *SVM* con Redes de función de base *RBF*, mejor predicción lineal insesgada genómica *GBLUP* y la regresión de Bayes *BR*. Finalmente, de los ocho conjuntos de datos el *SVM* obtuvo el mejor rendimiento en dos. Pero si se analiza el efecto en memoria empleada, el mejor fue el *SVM*, seguido por *GBLUP* y *BR*. Si bien el *BR* fue el peor de los tres algoritmos en cuestiones del rendimiento, en temas de la optimización del tiempo en sus interacciones fue mejor que el *GBLUP*, sin superar el *SVM*. Bajo la perspectiva de las dos variables, en términos generales todos arrojaron un comportamiento similar. Sin embargo, se puntualiza en la alta suficiencia del modelo al interactuar con datos de animales y plantas [22].

Otro enfoque de aplicación de esta metodología fue la realizada en los cultivos de moráceas, especialmente en las moreras cultivadas en los países asiáticos. Los cultivos de esta variedad se enfrentan a la presencia constante de un tipo de larva llamada *Diaphania pyloalis*. El estudio tuvo como objeto plantear una técnica efectiva de detección rápida de las larvas y sus impactos negativos en las matas, a través de imágenes hiper espectrales en la región e infrarroja cercana. Se obtuvo la información de 5 regiones principales sobre determinados espectros de la hoja, el mesófilo sano, niveles de daño: leve o grave; y finalmente de la larva. A partir de la data se estableció la implementación de métodos de análisis y clasificación como el análisis discriminante *PLS-DA* y *SVM* con mínimos cuadrados respectivamente. Entre los mejores métodos en la filtración y agrupamiento de los datos para distinguir cada muestra entre los espectros definidos, se encontró el *SVM* ya que demostró una efectividad del 100% basados en la espectroscopia de infrarrojo cercano *NIR* [23].

Se ha demostrado en el *SC* cómo la influencia de los orígenes geográficos del grano impacta directamente el sabor, aroma y la calidad, y así mismo su precio comercial. Es por esto que se ha indagado un poco más sobre la identificación del lugar de origen de cierto tipo de café para comprobar su autenticidad. Un grupo de investigadores especialistas en ingeniería de alimentos, química y manejo avanzado de instrumentos de alta precisión en China, desarrollaron un programa basado en la utilización de varias técnicas del *ML*, especialmente *SVM* y la espectroscopia de Tera hercios (THz), la cual consiste en unificar “todos los fenómenos electromagnéticos de ondas de radio a través de infrarrojos, rayos visibles, ultravioleta, rayos X, gamma y ondas electromagnéticas” [24].

Para lograr obtener conjuntos de muestras que abarcaran características puntuales, se utilizó el algoritmo *SVM* que, además de realizar el agrupamiento más sectorizado, redujo la complejidad para que al momento de implementar redes neuronales el sistema computacional obtuviera una mayor capacidad de entrenamiento. Sin embargo, a través del planteamiento de clasificación únicamente con *SVM* y el análisis de componentes principales *PCA* no alcanzó la precisión más alta ya que fue del 50 al 65%. Mientras que, al complementarlo con el algoritmo genético *GA* su rendimiento fue del 75%.

En conclusión, el algoritmo por sí solo es bastante funcional y recomendado para ser utilizado con otros que permitan una observación más profunda de los datos, ya que esto generará una limpieza y estructuración detallada, y así el algoritmo *SVM* al momento de asignar la etiqueta a cada dato sea más preciso. No obstante, el pilar de éxito del método está en la adecuada selección de las variables para modelar procesos robustos [25].

Independientemente del enfoque al aplicar el *SVM* se recomienda complementarlo con sub algoritmos que potencialicen su rendimiento. De igual manera, el método es eficiente por sí solo.

2.1.3. Redes neuronales artificiales RNN

Esta aplicación y actual tendencia del *ML* abarca códigos de programación un poco más robustos que los anteriormente mencionados, ya que fue creado con el propósito de enlazar infinidad de datos con múltiples características y así permitir la aplicación del mismo en sitios web del nivel de *Facebook* o *Twitter*. Por el nivel de complejidad de las *RNN* también son conocidas como aprendizaje profundo.

Todo inicia desde el concepto de las neuronas del ser humano las cuales analizan la información y al unificar todas las neuronas se obtienen deducciones alineadas al tipo de data que contiene cada una. A esta interacción se le denomina sinapsis.

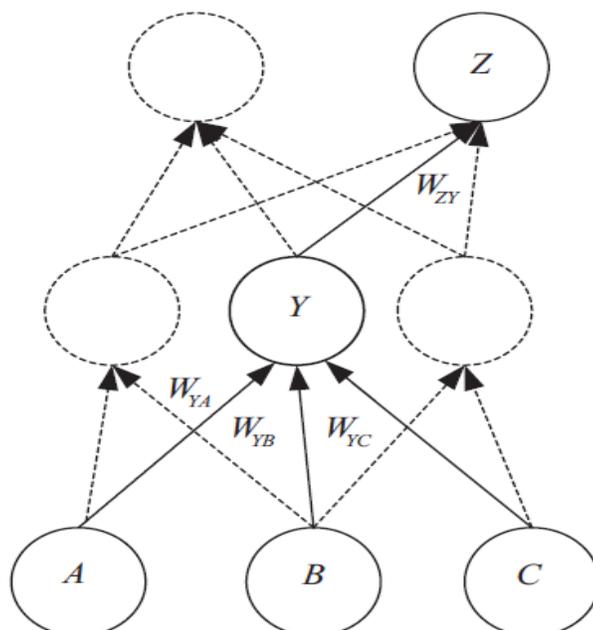
A nivel computacional, se menciona el término artificial ya que se pretende imitar la operatividad de las neuronas humanas en modelos de programación que obtengan un alto acercamiento a la lógica del cerebro humano.

Sobre este concepto básico, *Ethem Alpaydin* expone el método que emplean 2 tipos de *RNN*: Perceptrón y perceptrón multicapa.

El perceptrón contempla el uso de neuronas y define los respectivos efectos factibles en cuanto a la sinapsis o interacciones que ejecuten las neuronas a lo largo del algoritmo. Para interpretar óptimamente el funcionamiento, se suponen dos neuronas 1 y 2, donde cada una tiene asignado un valor de activación. Cuando las neuronas 1 y 2 se conectan, tienen un peso que determina el efecto de 1 sobre 2 o viceversa. Si en algún punto de las iteraciones la neurona 1 está activada también tratará de activar a 2, a este caso se le denomina sinapsis excitatoria. O si por el contrario, 1 está activada pero inactivará a 2 se le denomina sinapsis inhibitoria. Una vez se han realizado las iteraciones y conexiones entre ellas, cada una realizará la suma de todas aquellas neuronas que generaron algún tipo de sinapsis a partir de los pesos asignados inicialmente. Si se llega a presentar que el valor de activación total por neurona es menor al valor umbral, la neurona permanecerá inactiva. Si se presenta el caso contrario, la neurona se activará e inmediatamente enviará su activación a todas las neuronas que realizaron sinapsis con ella. Esta activación se visualiza en la siguiente figura:

Figura 5.

Asignación de pesos de neuronas que realizaron sinapsis entre ellas



Nota. Ejemplo de redes neuronales en donde asigna su peso W a cada neurona que tuvo sinapsis con la misma. Tomado de: *Ethem Alpaydin. E. Alpaydin, Machine Learning. ML: The new IA, Cambridge, MA: The MIT Press, 2016.*

En conclusión, el perceptrón calcula la suma ponderada anticipada a la decisión. Cuando se aborda el este tipo, difiere en la creación de capas de neuronas que interactúan entre sí [26].

Entrando en materia, para realizar la cartografía de cultivos ubicados en algunas regiones de Suiza, se determinó trazar la ubicación de los cultivos en mapas por medio de la clasificación de series de tiempo basados en imágenes de satélite. Para apropiarse del estudio, inicialmente se realizó la definición de las etiquetas a partir de los tipos de cultivos, dividido en clases gruesas (cultivos que abarcan gran variedad de productos

agrícolas) y finos (cultivo de un solo producto), configuradas en una *RNN*. En segunda instancia, se desarrolló la construcción de un método de sectorización de los tipos de cultivos considerados inusuales. Lo cual permitió definir el enfoque de la *RNN*, consiste en la evaluación de cada pixel de la imagen bajo las 3 etiquetas en diferentes niveles de granularidad, así es capaz de aprender a reconocer características inusuales del conjunto de datos, brindando así un nivel de sectorización más puntual y conciso. Finalmente, se compara el *RNN* con otros algoritmos, en donde el enfoque de la *RNN* al ser de nivel jerárquico superó a los demás algoritmos en un 9,9% [27].

Para la industria agrícola, las aceitunas son un producto clave ya que del fruto se produce el aceite de oliva. Por lo cual, el monitoreo para recopilar información sobre sus condiciones es fundamental. En Irán surge la creación de un programa computacional a través de *RGB* y *RNN* para el reconocimiento en tiempo real de las etapas de maduración de las aceitunas.

En primer lugar, se consideraron dos cultivos en donde se estandarizan las etapas de maduración del fruto (inmaduro, madurez verde, madurez negro y maduro). Este estudio presenta una peculiaridad al no contemplar el uso de un sistema de imágenes estructurado, esta decisión permitió que el desarrollo del modelo se mantuviera dentro de la naturalidad de los cultivos seleccionados. A partir del abordaje de los distintos métodos de las *RNN*, se efectúa una evaluación prueba y error bajo los parámetros de tiempo para así definir el pronóstico, precisión y el uso de la actividad computacional, entre varios algoritmos actuales y de vanguardia. Seguido de esto, se elaboró una red que incluyó seis clases de optimizadores: *Adagrad*, *SGD*, *SGDM*, *RMSProp*, *Adam* y *Nadam*, donde este último obtuvo el mayor grado de eficiencia siendo de 91,91%. Este maneja tasas de aprendizaje puntuales a partir de parámetros fácilmente adaptables a la frecuencia de actualización del mismo durante el entrenamiento, manejando un concepto similar al *Adagrad* pero *Nadam* utiliza el potencializador *Nesterov* [28]. En términos de la aplicación en el objeto de estudio, se expone la conformación de un programa fuerte con alto nivel de potencia para clasificar el grano en tiempo real a partir de las etapas de maduración. También, es bastante flexible al momento de ser

incorporado e implementado en el tratamiento de olivos [29].

Por otra parte, el entorno y la crisis ambiental por la cual atraviesa el mundo ha generado retos que conllevan a la implementación de prácticas agrícolas que no repercutan negativamente en el ambiente. Para lograr este objetivo desde la disminución de insumos químicos en los cultivos, es viable la creación de sistemas que optimicen el riego de pesticidas. Para esquematizar una solución efectiva, en el año 2020 el Departamento de Ingeniería, agricultura y Matemáticas de la universidad de *São Paulo State, Brasil UNESP* recolectó en plantaciones de café índices de vegetación como el borde rojo de diferencia normalizada y con respecto a la morfología del cultivo, se obtuvo datos de la altura y diámetro del cafeto. Bajo las variables definidas, utilizaron algoritmos bajo la filosofía de las redes neuronales (Perceptrón multicapa *PM* y función de base radial *RBF*) en conjunto con las técnicas de teledetección. Para evaluar la efectividad de los algoritmos al pronosticar las variables morfológicas, se tuvo en cuenta la precisión en R^2 y exactitud en cuanto a la evaluación del error cuadrático medio RMSE. Aunque no se especifica cuál algoritmo obtuvo la mejor precisión, se concluye que al utilizar el *PM* es factible pronosticar el volumen de una planta de café con un % de precisión coherente a partir de los índices de vegetación, diámetro y altura [30].

A partir de las investigaciones anteriormente expuesta, se puede afirmar que la utilización de las *RNN* es bastante robustas y efectivas al momento de manejar data con infinidad de características. Sin embargo, al definir el algoritmo que mejor se adapte a las necesidades del usuario puede ser muy engorroso. Por lo tanto, la experticia para implementar este tipo de *ML* deber ser excelente para así minimizar el error en la escogencia y se trabaje bajo escenarios realmente comprensibles por el programador. Así mismo, para aplicaciones en el entorno de tareas *PNL* o visión computacional las *RNN* son un modelo de referencia.

En cuanto a la perspectiva de las *RNN* dentro del sector agrícola y cafetero, definitivamente cumple perfectamente las funciones de evaluación impartidas y, como se pudo observar, tiene un comportamiento flexible y versátil ante cualquier problemática presentada en la industria.

2.1.4. Utilización de la espectroscopia de infrarrojo cercano NIR con diferentes algoritmos del ML

Aunque la *NIR* no hace parte de las aplicaciones del *ML*, es importante tener claridad sobre este concepto para comprender las aplicaciones de este método con el *ML*.

El análisis de infrarrojo cercano *NIR* hace parte de las técnicas de espectroscópica que usan el espectro electromagnético en un rango de longitudes de onda entre 800 y 2.500 nm. Este método ha tomado bastante fuerza para ser aplicado en entornos reales de las industrias debido a su practicidad, rapidez e inmersión de poco tiempo en cuanto a su preparación [31].

Las aplicaciones expuestas a continuación, enmarcan el *NIR* combinado con diversos algoritmos de *ML*.

Ante las exigencias del mercado actual por adquirir productos de excelente calidad y al mismo tiempo conserven la tradición de los cultivos para no perder su autenticidad, tanto organizaciones globales y emprendedores requieren de herramientas eficientes y económicas para clasificar su producción en el menor tiempo posible. Teniendo en cuenta esta necesidad que cada vez más se afianza en el mercado, Walter Lazo Aguirre desarrolla un modelo predictivo de clasificación de frutas a partir del *NIR*, *ML* y *Deep learning DL*. El producto objeto de estudio fueron los arándanos, la adquisición de las muestras fue en campo, obteniendo una cantidad de 1.000 datos. Seguido de esto, se realizaron simulaciones con la implementación del *NIR* con el propósito de recopilar la información de los espectros electromagnéticos digitalizados de cada una de las muestras y posteriormente construir la *RNN* para el entrenamiento del *dataset* a través de *keras* y *python*.

Como bien se conoce, cada aplicación del *ML* se divide en dos etapas: prueba y modelamiento del programa. Bajo ese criterio, se realizó el *testing* con nuevas muestras de *NIR*, arrojando un 92% de precisión [32].

Otra investigación integró la calidad y granos de cualquier tipo de fruto, fusionando las técnicas de *NIR*, imágenes rayos X y varios algoritmos de *ML* como *SVM*, método de *bayes* y bosque aleatorio, para estructurar un conjunto de algoritmos eficientes en cuanto a la clasificación de la calidad de semillas de la hierba *Urochloa brizantha*. En primer lugar, se predijo la composición y germinación de las semillas a través del *NIR* complementada con el método *Fourier FT* e imágenes de rayos X de muestras individuales de las semillas. Una vez se consolidaron las muestras, se crea el prototipo de *ML* a través de los algoritmos de bosque aleatorio *RF*, ingenuo *Bayes NB*, *PLS* y *SVM-r*.

El funcionamiento del *RF* consiste en clasificar *dataset's* robustos de manera más eficiente y precisa que los demás métodos existentes. Su diferencial es la capacidad de evaluar y entrenar la data rápidamente y simultáneamente. Además, cada árbol de decisión que construye funciona independientemente, facilitando así el entrenamiento. Como no existe un algoritmo perfecto, los *RF* no son la excepción, pues el interpretar los resultados pueden existir inconvenientes.

Desde el teorema de *Bayes*, el *NB* posee la habilidad de ser comprendido a partir de la probabilidad subyacente arrojada a partir de la base de datos. Como factor diferenciador del algoritmo, además de asignar etiqueta a cada dato, es capaz de crear nuevos registros manteniendo las mismas estadísticas. Por otra parte, el algoritmo contempla:

- Probabilidad: Valor de ocurrencia que puede obtener el modelo entre 0 o el 100%.
- Probabilidad condicional: A partir de la predicción anterior, se evalúa bajo otros parámetros la probabilidad de ocurrencia de un evento, es decir, define un valor más puntual.
- Distribución de la probabilidad: Permite visualizar las diferentes predicciones que puede obtener cierto escenario utilizando variables discretas o aleatorias
- Variable aleatoria: Se asigna un valor aleatorio a una variable [33].

Ya que se tiene claridad sobre el funcionamiento a grandes rasgos del algoritmo, es

preciso mencionar lo desarrollado en el estudio. En primera instancia, se utilizaron los algoritmos para el hallazgo de la germinación, donde cada uno obtuvo un 82% de exactitud y complementando su funcionamiento con las imágenes tomadas a través de rayos X y *FT-NIR* un 90%. Ahora bien, desde el enfoque de la composición de la semilla, los algoritmos individualmente, en promedio, obtuvieron un 61% de precisión y con las herramientas complementarias un 68%. Bajo los efectos obtenidos en el prototipo planteado, se concluye que al relacionar este tipo de técnicas en un mismo conjunto se comportan eficientemente a la hora de pronosticar la germinación de la semilla ya que sus resultados son rápidos, precisos y no destruyen la composición química del grano en estudio [34].

Por otra parte, la determinación de la calidad de los granos de café ha sido todo un misterio en cuanto a descubrir qué método o procedimiento es el más exacto, sobrepasando el umbral del 90% de exactitud. Claramente, desde catadores profesionales hasta analistas aficionados por catalogar el mejor café del mundo han invertido esfuerzos en descubrir posibles soluciones.

En el presente año, un grupo de investigadores a través del análisis de las proteínas y lípidos que componen a los granos de café verdes, formularon un prototipo de algoritmo de *ML* con *NIR*. Para llegar a ese resultado, se combinó *NIR* con la regresión de mínimos cuadrados parciales *PLS* el cual se basa en el error cuadrático medio *RMSE* donde evalúa el parámetro objetivo y el valor experimental de cada uno de los datos de entrenamiento y finalmente realiza un promedio. También, integra la evaluación de la variabilidad o dispersión existente entre los datos de prueba y finalmente el coeficiente de determinación *R2* [21].

A partir del *PLS*, se predice la cantidad de lípidos y proteínas presentes en varios tipos de granos estudiados. Pero antes de ingresar los datos *NIR* al *PLS*, se realizaron correcciones en la señal del espectro magnético para manejar data adecuada. Una vez se limpiaron los registros, a través del *PLS* se estandarizaron las variables con base en el resultado obtenido de la regresión, puntualmente del valor de *beta* β .

Al comparar los resultados con la corrección de la señal y sin esta, el prototipo que obtuvo mayor rendimiento fue el primero ya que mejoró radicalmente la eficiencia y calidad del mismo. También, influyó directamente la correcta selección de variables, ya que se obtuvo un *RMSE* del 0,10 y un *R2* de 0,98. El increíble desarrollo del *NIR* en conjunto con el *PLS* permite concluir que este prototipo puede implementarse fácilmente y tomar el liderazgo al precisar las proteínas y lípidos contenidos en una muestra de café verde [35].

Ya que en este apartado se exponen varias técnicas del *ML*, se evidencia cómo cada uno desde su esencia es realmente efectivo bajo distintas necesidades del entorno. Pues, fueron sometidos a clasificación de datos de tipo numérico e imágenes, y su capacidad de respuesta fue bastante acertada. Concluyendo así que, para su aplicabilidad en cualquier necesidad del SP es viable optar por el *ML* para proponer soluciones inmediatas y a largo plazo. Ahora bien, el porcentaje de efectividad y exactitud dependerá del análisis riguroso del objetivo al cual se pretende llegar y así realizar una elección alineada a este.

2.2. La cadena de suministro del sector cafetero colombiano: Diferenciales a nivel mundial y una mirada desde los efectos del COVID-19

Las cadenas de abastecimiento de las industrias, aunque se lideren bajo un mismo concepto, poseen un factor diferenciador gracias a los procesos productivos y administrativos en los que se incurren para la obtención de un producto final. Además, cada sector de la economía se enfrenta a diversos sucesos y retos en el día a día, permitiendo así la actualización o complementación de cada eslabón de la cadena según las necesidades vigentes del mercado.

Cada país posee su metodología para establecer la CS a partir del sector económico, y Colombia no es la excepción. Ya que el sector agrícola, principalmente el cafetero es el que atañe a la actual investigación, es pertinente mencionar aquellos puntos en los cuales discierne de la estructura tradicional de las CS a nivel mundial. En cuanto a las

temporadas de cosecha del grano, la mayoría de países optan por realizar periódicamente la programación de los cultivos, ya sea por periodos extensos o cortos. Por parte del tratamiento y transformación del café, comúnmente los agricultores comercializan el grano en bruto a ciertos procesadores con el fin de minimizar la inversión en recursos. Y por último, las cooperativas caficultoras presentes en los demás países no son representativas en cuanto a la participación activa y continua con los agricultores.

Por su parte, en Colombia la cosecha del café verde se mantiene durante todo el año generando la producción de alrededor de 14 millones de sacos de 60 kilos. Además, los caficultores realizan el procesamiento del grano para posteriormente comercializarlo, por lo cual es inusual encontrar en el mercado minorista y local el café verde. Finalmente, mientras que la presencia de cooperativas del café a nivel mundial es relativamente baja, en Colombia representan el 35% de la generación de ventas efectivas de los bultos del café a nivel local y en el exterior, convirtiéndose así en las entidades del gremio con la mayor participación en Latinoamérica.

A raíz de la emergencia sanitaria generada por el *COVID-19*, varias industrias sufrieron afectaciones en la capacidad de respuesta ágil ante eventualidades dentro de la CS. Si bien esta consecuencia es negativa en cuanto a la inmediatez de solucionar adversidades, se convierte en positiva una vez las organizaciones identifican sus puntos débiles, y a partir de ellos, construyen en conjunto estrategias enfocadas a la resiliencia en sus CS. Desde el SC, agremiaciones e instituciones ejemplares como la Federación Nacional de Cafeteros de Colombia *FNC* en conjunto el Gobierno Nacional Colombiano, realizaron la divulgación de protocolos de bioseguridad y cuidados personales a todo el gremio con el objeto de minimizar la tasa de contagio y así velar por la salud de los caficultores.

Aunque las intenciones de mitigar al máximo la transmisión del virus en los cafeteros fueron acertadas, el confinamiento estricto por el cual atravesó el país por más de 3 meses afectó considerablemente el transporte terrestre y marítimo, tanto al interior como al exterior, trayendo consigo aumentos en los costos de operación y almacenamiento

prolongado de los bultos. Una situación similar vivió a nivel mundial la industria cafetera, en donde las pérdidas fueron abismales y claramente la CS se vio fuertemente afectada. Desde la perspectiva del talento humano, se evidencia el gran impacto en la economía de los caficultores, pues se estima que los cultivos de café emplean a 743.000 trabajadores, representando esta cifra el 40% de la mano de obra rural del país [36].

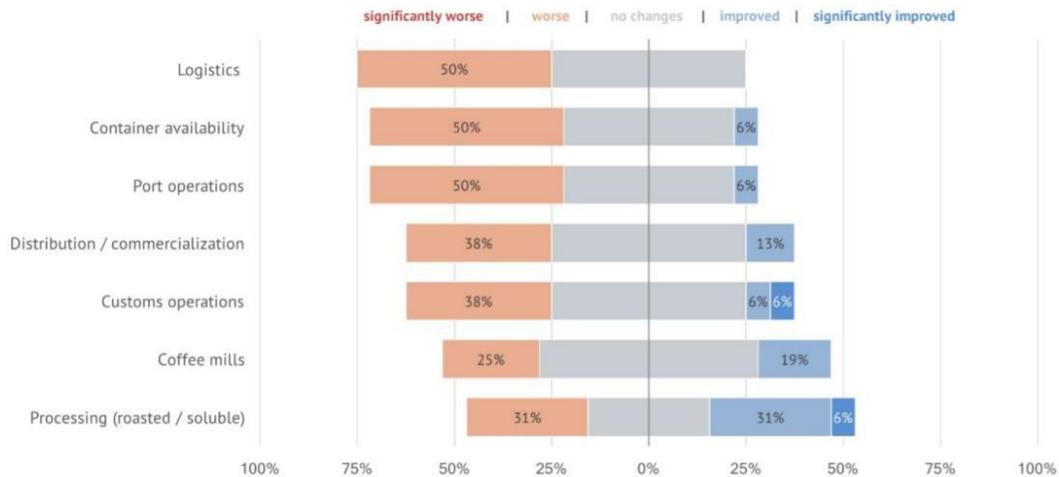
Así mismo el personal cualificado que no reside en la zona se vio impactado ya que por motivo de las restricciones de movilidad no pudieron desplazarse durante un largo tiempo a sus hogares de origen.

Desde una mirada global, la *OIC* realizó una encuesta a varios países latinoamericanos miembros de la organización por concepto de exportadores, entre ellos Colombia, sobre la percepción de los efectos del *COVID-19* en el sector mundial cafetero. Esta encuesta evidencia la gran preocupación por las repercusiones en la logística interna de cada organización y así mismo la afectación en los trámites de exportación que incluyen la usabilidad de contenedores, procesos aduaneros y finalmente la gestión de las operaciones portuarias.

Estos factores repercutieron considerablemente en las exportaciones ya que prolongaron los tiempos de entrega, haciendo que los costos por concepto de comercialización y negociación directamente presentaran un aumento. Aunque este estudio fue realizado a mediados del año 2020, aún estos factores son generadores del retraso en la gestión de las CS. A continuación, se registran los resultados de la valoración asignada a los eslabones de la CS bajo la escala considerablemente afectados hasta significativamente mejor [37].

Figura 6.

Afectación covid-19 en operaciones intermediarias de la CS del café



Nota. Resultados porcentuales sobre la percepción de los impactos del covid-19 en las CS de ciertos países latinoamericanos. Tomado de: Organización internacional del café. 2020.

Con respecto a la presente investigación, este tipo de información genera valor para el análisis del panorama actual al cual se ve expuesto la producción de café en el país porque permite detectar qué falencias se encuentran en la CS y así identificar las opciones de mejora de la misma en cuanto a la logística de abastecimiento para aportar soluciones a las necesidades del mercado cafetero.

2.2.1. Elementos principales de la logística de aprovisionamiento en el sector cafetero colombiano

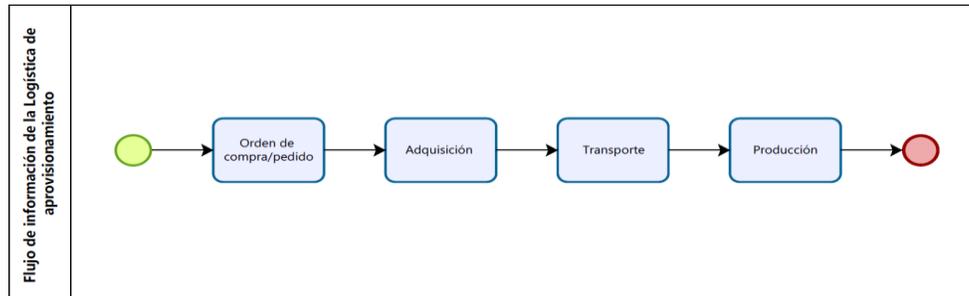
En la industria cafetera, Brasil, Vietnam y Colombia son los 3 exponentes a nivel mundial con índices de fabricación más altos, representando así una tercera parte de la producción global [38]. Este posicionamiento conlleva una gran responsabilidad alimentaria, social y ambiental, pues el debido manejo de los recursos y cultivos debe ser estricto para mantener un café con altos estándares de calidad.

El SC colombiano es consciente de los grandes retos que debe abordar para equilibrar sus CS y la inversión de capital que demanda la producción de los cafetos. Si bien es cierto que existen varias entidades que apoyan a los caficultores en cuanto a beneficios económicos, asesorías con expertos y canales de comercialización, es importante optar por brindar soluciones fehacientes de asesoría detallada en la estructuración de sus CS enfocadas a la sostenibilidad, optimización de recursos y tiempo, flexibilidad de adaptabilidad a cambios drásticos y mantenimiento de la calidad del café, ya que el eslabón crucial para iniciar la planeación de los requerimientos del cultivo es la LA. Cabe aclarar que en este apartado se expondrá a groso modo los componentes de la LA puntualmente en Colombia.

La logística de aprovisionamiento hace referencia a la integración completa del proceso de abastecimiento de materias primas hasta su transformación, es decir, producción. En este factor se puede evidenciar cómo este eslabón de la CS incluye actividades puntuales que conllevan a consolidar los pilares de la CS. A continuación, se puede observar el flujograma general de la información manejada dentro de la LA.

Figura 7.

Flujo de información de la logística de aprovisionamiento



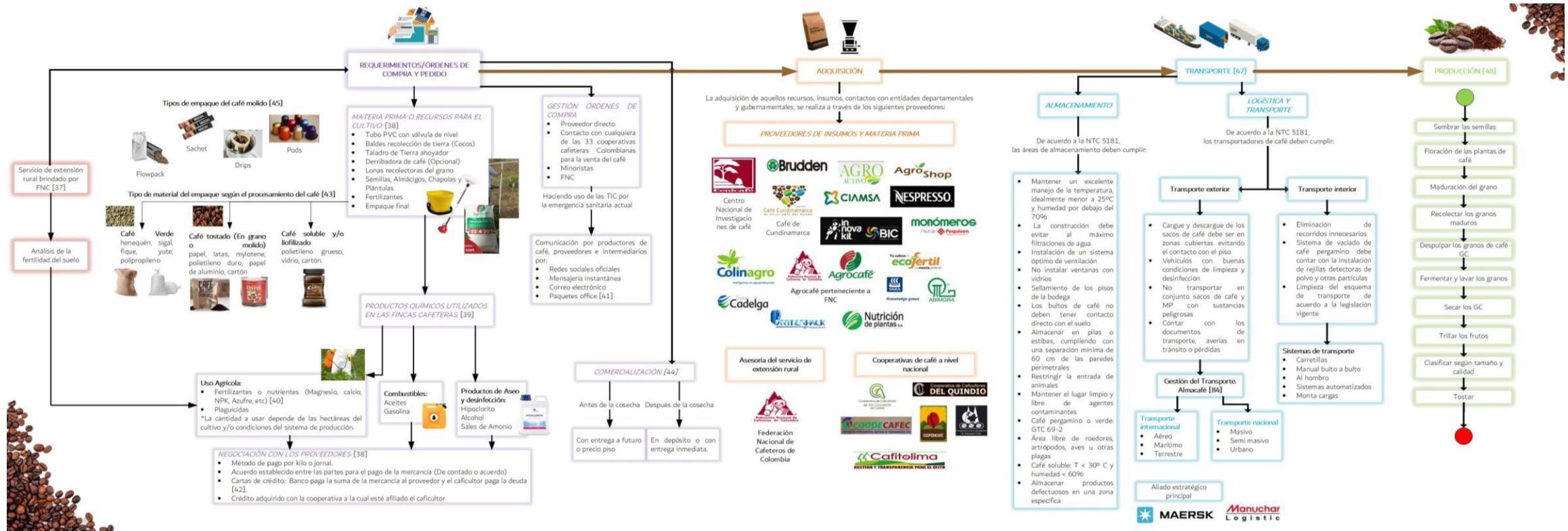
Powered by
bizagi
Modeler

Nota. Visualización de los procedimientos que incluye la LA. Tomado de: *Inbound Logistic in Coffee Industry. Transport Mode Assessment for Inbound Logistics: A Study Based on Coffee Industry.* Universidad Federal de Pernambuco <https://ieeexplore.ieee.org/document/6974097>

Según la figura anterior, la se compone de las órdenes de compra de las materias primas, entre otros; y de pedido abarca aquellas solicitudes de compra del producto anticipadamente. Una vez la CS tiene las órdenes, procede a adquirir aquellos materiales que son necesarios para la fabricación del producto. Cuando ya se tiene claridad sobre cuáles proveedores tomarán el papel de abastecedores, los insumos que provean serán transportados hasta el lugar de acopio o directamente al área de producción por vía terrestre, aérea o marítima. Y finalmente, la compañía se encuentra en condiciones de manufacturar. La Figura 8 describe la LA del SC colombiano.

Figura 8.

Descripción del flujo de información de la LA del SC colombiano



Nota. Descripción de cada punto que conforma la LA del SC colombiano.

El anterior diagrama esquematiza cada proceso que conforma la *LA* en la *CS* cafetera colombiana desde los requerimientos de órdenes de pedido y compras hasta el proceso de producción. A continuación, se explicará brevemente cada ítem que contiene el diagrama con la finalidad de puntualizar y aclarar otros aspectos fundamentales que debe tener en cuenta el caficultor.

2.2.1.a. Requerimientos u órdenes de compras y pedidos. En este ítem se expondrán algunos procedimientos claves para estructurar adecuadamente el inicio de una *LA* efectiva para los caficultores colombianos.

Para la obtención de cultivos eficientes, la *LA* debe iniciar desde el servicio de extensión rural que brinda la *FNC*, la cual consta de un análisis exhaustivo de la tierra o lote donde se ubicará el cultivo para determinar en qué condiciones químicas se encuentra el suelo y así determinar qué nutrientes como Azufre, magnesio y calcio, son necesarios para mantener el *PH* de la tierra balanceado. Además, se realiza un mapeo de la finca en donde se proporciona la ubicación de cada planta, según la altura e inclinación del lote, variedad y sombra imprescindible para un buen crecimiento del tallo. A través de este estudio, se establecen indicaciones explícitas sobre el cuidado de las plantas y fertilizantes.

Dentro del flujo de información de las órdenes de pedido principalmente se aborda la adquisición de materias primas *MP* básicas para conservar un excelente cultivo. Como mínimo el agricultor debe contar con un tubo *PVC* con una válvula de nivel para calcular la distancia entre cada cafeto, semillas, almácigos o plántulas para posteriormente ser sembradas en cada posición definida. La siembra de las plántulas, se puede efectuar por medio de un taladro de tierra ahoyador que permitirá tener mayor precisión a la hora de realizar cada agujero. Ahora bien, para la recolección de los granos maduros, existen dos métodos diferentes para ejecutar este procedimiento, la escogencia dependerá del factor económico y la productividad objetivo. A su vez, los baldes recolectores del grano se pueden utilizar en el método manual, o en caso de ser automático se

reemplazar este elemento por las derribadoras. Finalmente se adquieren los fertilizantes que contengan componentes como magnesio, calcio, azufre, etc. bajo el análisis del suelo.

Por otro lado, el funcionamiento de equipos y herramientas que intervienen en la producción demandan la utilización de combustibles que deben estar disponibles una vez se requieran para no afectar el ritmo de fabricación.

Para la salubridad tanto del personal como de las instalaciones, es indispensable contar con productos de aseo y desinfección, teniendo en cuenta que los componentes de los mismos no generen contaminación cruzada del producto y no afecten la salud de los colaboradores.

Teniendo claro el listado de *MP* requeridas para el proceso, se realiza la negociación con los proveedores sobre el método de pago. Para este acuerdo, se puede optar por pago de contado, cancelación por kilo del producto o si es por concepto de servicios, pagar al personal por jornal laborado. Una metodología bastante común consta en establecer un acuerdo entre las dos partes para el pago de la mercancía, donde se define la periodicidad y valor de la cuota. Si por alguna razón el caficultor no cuenta con los recursos económicos para cancelar los productos, tiene dos opciones de financiamiento: acceder a un crédito a través de la cooperativa cafetera a la cual se encuentre afiliado o solicitar una carta de crédito con un banco que cancele el valor total al proveedor, y el caficultor en ciertos plazos establecidos desembolse el dinero.

Desde el eje de compras, la gestión es vital para definir cómo se llevará a cabo la comercialización del café en cualquiera de sus presentaciones y así efectuar la venta de los productos, ya sea con la *FNC*, alma café, clientes directos y minoristas. Cuando el caficultor tenga claridad sobre cuáles son sus clientes y requerimientos de los mismos, puede realizar la negociación de los precios antes de la cosecha por medio de un acuerdo de entrega a futuro de la mercancía y venderla a precio piso, es decir, un precio definido que no obtendrá ninguna variación bajo la fluctuación del precio internacional del café.

En palabras prácticas, si oferta el kilo por X precio, ese valor no cambiará en el tiempo que transcurra hasta la entrega final, lo cual no incidirá si el precio sufre variaciones drásticas ya sea al alza o a la baja. U ofrecerle la alternativa de transferirle en depósito los sacos con entrega inmediata. La escogencia de cada tipo de comercialización dependerá de la prioridad del cliente y del caficultor. Adicional a esto, es importante definir muy bien cómo se obtendrán ganancias de la cosecha para así maximizar los ingresos netos del cultivo.

2.2.1.b. Adquisición. En este punto se exponen algunos proveedores a nivel nacional sobre insumos, recursos y MP recomendados directamente por las cooperativas cafeteras. Además, se menciona como entidad principal del servicio de extensión rural a la FNC.

A partir del levantamiento de información sobre los proveedores de distintos insumos como fertilizantes y almácigos existe mayor claridad sobre el amplio panorama de organizaciones que definitivamente aportan al buen desenvolvimiento de los cultivos de café colombianos. De igual manera, es crucial e indispensable contar con varias opciones de suministros con el fin de realizar una elección del proveedor que más se ajuste a las necesidades del cafeto y maneje una proporción equilibrada entre calidad y precio.

2.2.1.c. Transporte. El transporte es aquel factor considerado como uno de los más importantes dentro de una CS, pues si bien se coordina todo tipo de abastecimiento de productos bajo estrategias, el transporte es el que efectúa las entregas del mismo y a su vez provee a los clientes nacionales e internacionales. Por lo tanto, es indispensable contar con operadores logísticos efectivos que cumplan con los estándares de despacho y mantengan en excelentes condiciones el producto durante su distribución. Sin embargo, existen eventualidades imposibles de controlar, como por ejemplo los sucesos geológicos, irregularidades e inseguridad en vías, accidentes, entre otros. Lo importante es visualizar el grado de resiliencia de las CS para afrontar aquellos inconvenientes, evitando una elevada afectación en las operaciones logísticas.

En el diagrama, se divide el transporte desde 2 perspectivas consideradas fundamentales para la LA. Por un lado se encuentra el transporte exterior, en donde, de acuerdo a la *NTC 5181* que establece las *buenas prácticas de manufactura para la industria cafetera*, expone directrices generales sobre cómo manipular adecuadamente los sacos o productos de café. Algunas de estas son:

- Los lugares de cargue y descargue de los productos debe ser en un espacio abierto cubierto, apartado de zonas de transformación de la *MP* y evitando la exposición a condiciones climáticas que puedan afectar la integridad del producto
- Inspeccionar los vehículos antes de cargar el café, con el fin de garantizar las condiciones sanitarias adecuadas
- No transportar en conjunto productos de café con sustancias que puedan ser potencialmente contaminantes
- Utilizar estibas o canastillas para evitar el contacto directo del producto con el piso

En cuanto al transporte interior se mencionan aquellos lineamientos generales para movilizar la mercancía dentro del centro de acopio o bodega. Las más relevantes son:

- Eliminación de recorridos innecesarios
- Sistema de vaciado de café pergamino debe contar con la instalación de rejillas detectoras de polvo y otras partículas
- Limpieza del esquema de transporte de acuerdo a la legislación vigente [1]

Al interior del centro de acopio, el traslado de la mercancía se realiza por medio de carretillas, montacargas, bulto a bulto y al hombro. En este punto es donde se observa el gran potencial de innovación que puede ser implementado en las bodegas, por ejemplo, la utilización de sistemas automatizados como bandas transportadoras, montacargas autónomos dirigidos por robots, entre otros. Esto con la finalidad de mantener un flujo constante de la carga y así disminuir los tiempos de operación. Para llegar a tal grado de innovación, es importante que el sector caficultor esté abierto a nuevas oportunidades y visualice todas aquellas ventajas que pueden obtener a través

de la automatización.

Teniendo varias directrices dictadas sobre el transporte gracias a la norma técnica colombiana, entra a jugar un papel crucial la gestión del transporte. Con miras del transporte exterior, el operador logístico más importante y comúnmente usado por los caficultores para servicios de exportación es *Alma café* ya que posee el mayor complejo logístico de la industria ubicado en Soacha, Cundinamarca.

Cuenta con alrededor de 4.200 posiciones de estanterías y 2.100 en puestos de trabajo dedicados netamente a la preparación del despacho de pedidos. Adicionalmente, maneja el método *FIFO* que consiste en despachar el primer lote que entra al almacén y así sucesivamente según el orden de llegada [49]. Aparte de brindar soluciones logísticas internacionales, compra y vende el café verde. Este operador tiene como aliado estratégico principal *Maersk* el cual efectúa la movilización en territorio internacional de la carga, ya sea por vía terrestre, marítima o aérea. Por otro lado, para el traslado de mercancía nacional, *Alma café* posee alianzas estratégicas con distintas organizaciones logísticas para prestar el servicio de traslado bajo la modalidad masiva, semi masivo y urbano. El transporte masivo hace referencia a la utilización del cupo completo para varios despachos nacionales. Por su parte el semi masivo, transporta cierta carga a puntos de entrega definidos dentro del territorio nacional. Y finalmente el urbano abarca aquellos despachos que se deben cumplir dentro de una misma ciudad y/o tiene como destino final alguna ciudad principal de Colombia con cobertura en periferia [50].

Finalmente, el sector logístico colombiano se enfrenta a diversas problemáticas que estacan su productividad. Esta afirmación la sustenta la *Federación Colombiana de Transportadores de Carga y Logística Colfecar*, mencionando que las compañías transportadoras de carga en el territorio colombiano se ven obligadas a cancelar valores de peajes supremamente costosos, en donde no existe un equilibrio en la inversión vial del gobierno y los aportes tributarios por concepto de peajes [51]. Además, la infraestructura vial hasta este año tendrá la puesta en marcha de grandes proyectos que

aportarán a la movilidad del país, por lo tanto, el avance en cuanto a conexiones viales no se verá reflejada en el corto plazo [52].

2.2.1.d. Producción. Cuando se han llevado a cabo los procesos de adquisición de MP a través de proveedores y la gestión del transporte para que finalmente el cultivo tenga la disponibilidad de los recursos definidos, comienza el proceso productivo. Todo inicia desde la correcta selección de la semilla, plántulas o almácigos ya que de esta dependerá en gran parte la calidad del café. Se siembra en cada punto definido y en aproximadamente 4 a 5 meses florecerá el cafeto según las condiciones de la tierra y el clima. Una vez la planta se encuentre en su proceso de crecimiento o floración, los granos tardarán entre 6 a 8 meses en alcanzar su maduración dependiendo de las condiciones climáticas a las cuales está expuesto el cultivo. Ya que desde la siembra hasta la maduración se ha llevado una trazabilidad del tiempo de cosecha, el caficultor tiene un rango aproximado de tiempo para recolectar aquellos granos maduros, en donde lo podrá realizar de forma manual o con ayuda de equipos. Una vez se recolectan los frutos maduros, deberán ser procesados por la máquina despulpadora que contiene un rodillo que permite separar automáticamente la pulpa del grano de café. Teniendo los granos despulpados, se procede a fermentar, o más conocido entre los caficultores colombianos como el avinagramiento; este proceso puede durar entre 16 a 18 horas donde se descompone totalmente el mucílago (membrana del grano) y lavar el lote recolectado. Con el café debidamente procesado en los dos pasos anteriores, se ejecuta el secado por medio de la exposición a la luz solar, esto permitirá realizar la conversión a café pergamino, en donde a través de la trilla se retirará de cada grano la cascará de pergamino y se convertirá en café verde. Se clasifica según su calidad y tamaño para separar los diversos niveles de calidad del café (excelso, pasilla, entre otras). Finalmente, se realiza la tostión de los granos bajo un proceso de calor que permitirá potencializar las características como sabor y aroma.

En conclusión, para la presente investigación es fundamental el estudio detallado de la logística de entrada ya que es la que brinda las pautas de las variables indispensables para la construcción del algoritmo alineado a las necesidades de la industria actual y

adicionalmente, permite profundizar y ampliar el conocimiento del proceso productivo del café para así realizar una observación minuciosa que cada eslabón de la *LA*.

2.2.2. Visualización del modelo SCOR desde la perspectiva de los proveedores

Con el objeto de describir con simplicidad la *CS* de una empresa, organización o sector económico, varios autores han construido metodologías para suplir esta necesidad. Uno de ellos, es considerado como la entidad de mayor peso y confiabilidad en cuanto a *CS*, el *supply chain Council SCC* el cual planteó la herramienta del *modelo SCOR*, esta consiste en caracterizar todas aquellas actividades en las que se deben incurrir para satisfacer la demanda de los consumidores. Además, aborda la categorización de dichas actividades para jerarquizar y ordenar la explicación de la *CS*. Estas categorías son: planificar, proveer, fabricar, entregar y devolver.

La escogencia de este método fue para describir la *CS* del *SC* colombiano bajo la perspectiva únicamente de los proveedores se debe a la generación de valor a la presente investigación, así mismo se considera pertinente porque el *SCOR* posee la idoneidad de especificar concretamente todos los factores influyentes en la industria cafetera.

Con base en información recopilada a través de entidades sumamente importantes del café como la *FNC* y *cenicafé*, se unifica todas aquellas acciones que conforman una *CS* general en Colombia. A continuación, se encontrará su desarrollo.

Tabla 1.

Modelo SCOR de la CS de un caficultor tradicional colombiano

Descripción de la cadena de suministro de un caficultor tradicional colombiano bajo el Modelo SCOR		
	Cultivo	Comités Cafeteros regionales
Plan	TODOS	P4
Source	S1	N/A
Make	M1	N/A
Deliver	D1	D4
Return	Source Return	SR1
	Deliver Return	DR1
En algunos casos específicos el caficultor puede manejar <i>make to order</i> si un cliente puntual realiza una orden de compra previa de cierta cantidad de café una vez haya finalizado su proceso de maduración. De igual manera, la metodología <i>source make to order</i> es implementada. Sin embargo, para el caso de estudio se describe la CS desde la ejecución de los cultivos para posteriormente ser vendidos en el mercado sin tener una orden previa de compra.		

Nota. Modelo SCOR construido a partir de información recopilada de la FNC.

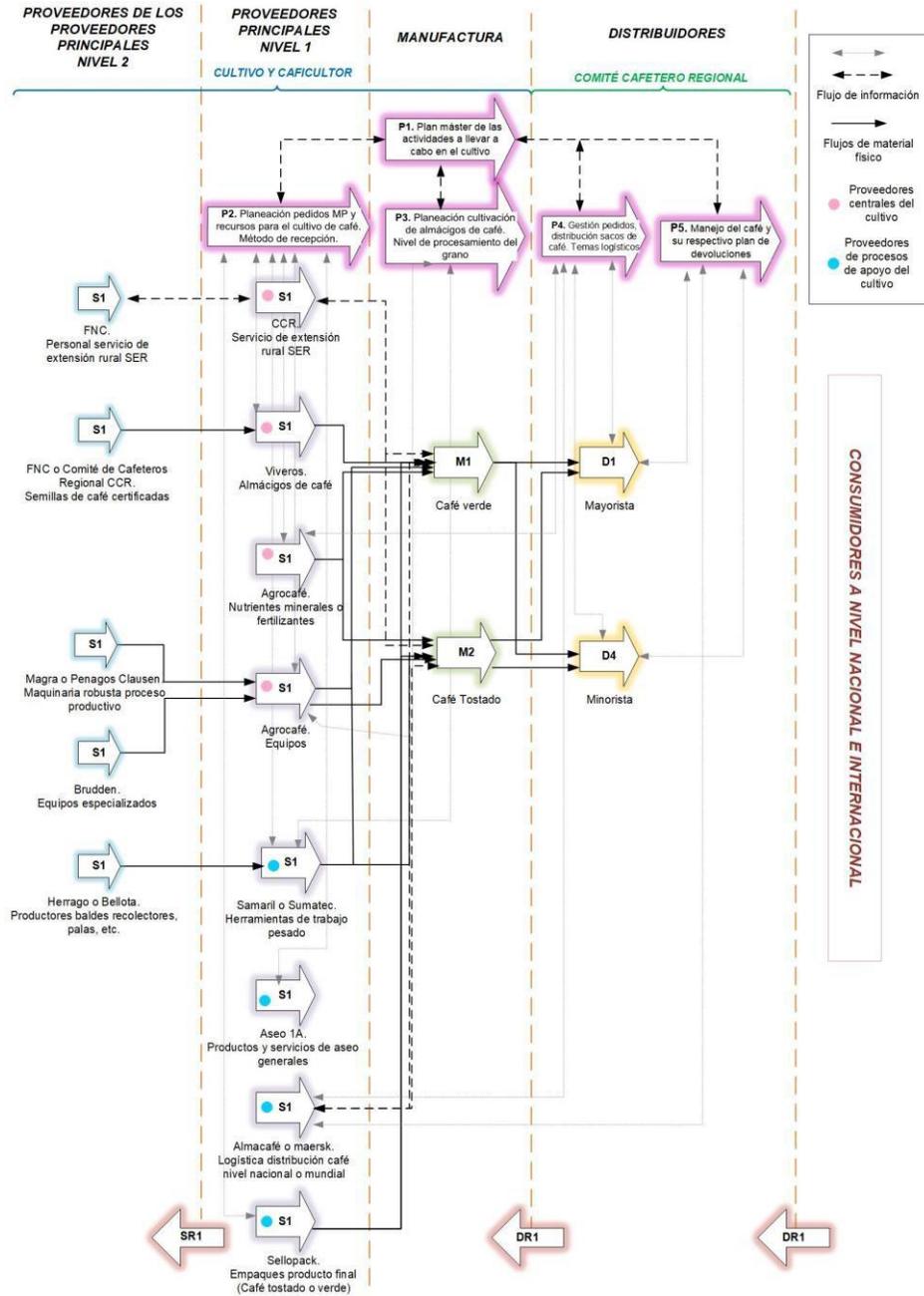
En adición, el rol que asumen los comités cafeteros regionales CCR es fundamental para los agricultores ya que estas entidades a nivel municipal y regional garantizan 5 puntos cruciales para el caficultor. El primero de ellos trata sobre la garantía de compra, esta permite garantizarle al caficultor la compra de los bultos de café al precio del día publicado por la Federación. Como se ha mencionado anteriormente, bajo el acompañamiento de la FNC se brinda el servicio de investigación científica por medio de los profesionales expertos de *cenicafé*.

Por otra parte, los CCR colaboran en el proceso de posicionamiento de la marca del caficultor a nivel nacional con el objeto de incrementar el consumo de café y mejorar la

calidad de vida de los productores a través del incremento en sus ingresos. Y, por último, pero no menos importante, realizan desarrollos sociales que consisten en el apoyo total a las comunidades de familias cafeteras en cuanto a infraestructura, creación de escuelas rurales, dictado de conferencias de seguridad alimentaria, renovación de cafetales y la fertilización periódica de sus cultivos. Como se puede observar, las funciones de los comités definitivamente agregan bastante valor a la CS y a su vez son coherentes en el mejoramiento de la calidad de vida de cada caficultor colombiano [53]. Complementando la realización del *modelo SCOR* de una forma más visual, a continuación, se observa el diagrama de hilos de la CS cafetera.

Figura 9.

Representación gráfica del modelo SCOR



Nota. Diagrama de hilos que representa el flujo de información y producto físico entre proveedores, distribuidores y consumidores pertenecientes a la CS del café.

Debido a la naturaleza de los cultivos, la planeación de las CS más común abarca la adquisición de los recursos sin requerimientos específicos de clientes, esto se debe al tiempo y adversidades que se pueden presentar en los cafetos durante la etapa de crecimiento y maduración. Puede ser aproximadamente de año y medio a dos años hasta la obtención de granos madurados, por lo tanto, la comercialización del producto inicia una vez haya sido procesado y esté disponible para la venta final.

Únicamente los insumos que requieren de almacenamiento por tiempo prolongado son las herramientas de trabajo de pesado y algunos fertilizantes. Por parte de los productos finales obtenidos, se almacenan tanto los granos verdes como los tostados. El nivel de procesamiento de los frutos dependerá de las especificaciones de los clientes, ya que, para efectos de exportación el café debe estar verde.

Al ser un producto perecedero y alimentario, no requiere de un mantenimiento puntual del grano, pero sí de las plantas de café en cuanto a fertilización periódica, tratamiento de prevención de plagas y enfermedades, renovación de cafetos. Razón por la cual no existe ningún tipo de devolución por concepto de revisión o mantenimiento técnico. Las devoluciones se presentan en caso de existencia de granos defectuosos o no aptos para el consumo humano.

2.2.3. Parámetros para clasificar un café colombiano como especial o excelso

Colombia al ser uno de los grandes productores de café a nivel mundial y por sus diversas condiciones geológicas que permiten la obtención de granos con distintas características como forma, sabor y aroma, tiene una gran responsabilidad con el mercado mundial en ofertar y distribuir a cualquier rincón del mundo el sabor del café colombiano.

Con la finalidad de acotar la parametrización de insignias otorgadas al café según sus características, actualmente existen distintas categorías de clasificación del grano teniendo en cuenta su exclusividad. Una de ellas se denomina *café de origen* la cual

evalúa el lugar de procedencia en donde se considera inusual la producción de café. Por otra parte, se encuentra la categoría de certificados expedidos a los cultivos como el *Rainforest que reconoce la agricultura sostenible mediante las buenas prácticas detectadas*. A su vez la categoría de *grano orgánico* exime totalmente la utilización de productos químicos y el más importante para el presente estudio, *rasgos diferenciadores de taza*. Esto se realiza a partir de un perfilamiento de taza que consiste en la descripción de la experiencia al tomarse un café por parte de un barista certificado por la CQI, en donde identifica algunas particularidades que contiene la taza en cuanto a acidez, uniformidad, balance, sabor, aroma, entre otros. Estos factores diferenciadores se originan desde el clima, la geografía y las buenas prácticas del caficultor en el sostenimiento de los cultivos. Adicionalmente, la entidad menciona 4 aspectos fundamentales para un perfil de taza único:

- Aroma: El olor que expide la taza de café en cualquiera de sus preparaciones
- Cuerpo: Es aquella sensación del sabor de la taza y su perdurabilidad en la boca
- Fragancia: Aquella que expide el grano una vez ha sido molido
- Acidez: Intensidad del sabor de la bebida [54].

Como se puede observar, el SC colombiano toma en cuenta parámetros establecidos por el CQI para la calificación del café de alta calidad. Por consiguiente, demuestra totalmente la alineación del sector con el mundial siendo este un factor positivo para la vigencia del café como excelso en el mercado internacional y así aumentar la exigencia en la producción nacional, con el objeto de que Colombia mantenga sus estándares de calidad y pueda acceder a nuevos territorios con nuestro producto estrella, el café.

2.2.3.a. Normatividad de calidad del grano de café verde expedida por el Comité Nacional de Cafeteros CNC. A lo largo de la historia, el café colombiano ha sido reconocido por su variedad de perfiles de taza dependiendo de la región de origen. Al ver el gran auge del café por excelencia, el CNC ha venido monitoreando la calificación del mismo a partir de estándares explícitos que demuestren con precisión lo especial y

único que puede llegar a ser el grano a partir de resoluciones generales aplicables a nivel nacional. Actualmente, bajo la resolución 02 del 2016 se exponen todos aquellos requerimientos y lineamientos de calidad que debe cumplir el grano verde para ser considerado de alta calidad y así mismo ser apto para exportación.

Todo café verde que pretenda ser catalogado como Almendra excelso deberá cumplir lo siguiente:

Figura 10.

Requisitos del café verde en almendra excelso dictados por la FNC



Nota. Descripción de los ítems del Artículo 03 de la *Resolución 02 de 2016*. Tomado de: Comité Nacional de Cafeteros, «Resolución 02 de 2016. Por la cual se unifican y actualizan las normas de calidad del café verde en almendra para exportación,» Federación Nacional de Cafeteros de Colombia, 16 abril 2016. [En línea]. Available: <https://federaciondecafeteros.org/static/files/RESOL04.pdf>.

Además, se establece la métrica para la sectorización de los defectos presentes en el

lote del café. Esta deberá realizarse a través de la tabla de equivalencias entre granos defectuosos y faltas.

Tabla 2.

Equivalencias entre granos defectuosos y faltas

Grano Defectuoso	Falta
Granos Grupo I	
Un(1) Grano Negro	1.0
Un(1) Grano Vinagre	1.0
Un(1) Grano Reposado	1.0
Dos (2) Granos Parcialmente Negro	1.0
Dos 2 Granos Parcialmente Vinagre	1.0
Un(1) Grano Cardenillo	1.0
Dos(2) Granos Parcialmente Cardenillo	1.0
Un(1) Grano Ambar o Mantequilla	1.0
Granos Grupo II	
Cinco(5) Granos Mordido y Cortado	1.0
Cinco(5) Granos Picado por Insecto	1.0
Diez (10) Ligeramente Picado por Broca	1.0
Cinco(5) Granos Malformado Averanado	1.0
Cinco(5) Granos Inmaduro	1.0
Cinco(5) Granos Aplastado	1.0
Cinco(5) Granos Decolorado Veteado	1.0
Cinco(5) Granos Flojos	1.0
Cinco(5) Granos Sobresecado	1.0

Nota. Parametrización de defectos catalogados directamente por la *FNC* y su respectiva penalización numérica. Tomado de: Comité Nacional de Cafeteros, «Resolución 02 de 2016. Por la cual se unifican y actualizan las normas de calidad del café verde en almendra para exportación,» Federación Nacional de Cafeteros de Colombia, 16 abril 2016. [En línea]. Available: <https://federaciondecafeteros.org/static/files/RESOL04.pdf>.

Para aplicar esta evaluación, es necesario tomar muestras de granos verdes de 500 gr, en donde podrá tener hasta 24 faltas de la tabla, si excede este puntaje el café no podrá ser calificado como excelso.

El monitoreo del cumplimiento de los anteriores ítems debe ser ejecutado por la *FNC* con el propósito de incrementar exponencialmente la productividad cafetera y así mismo mantener el café colombiano en el comercio internacional. Sin embargo, la entidad no emite ningún tipo de certificado de calidad, únicamente otorga una constancia para el pago de la contribución cafetera dictada por la *ley 1337 de 2009*.

Tener este tipo de conocimiento sobre las prácticas de valoración de la calidad de los granos de café a nivel Colombia es sustancial y conveniente para el presente estudio ya que aterriza la realidad actual del *SC* en cuanto a estándares de calidad. Y a partir de los mismos, permite reconocer cómo la *FNC* ha venido evolucionando en regulaciones que conduzcan al café colombiano a las categorías más altas de calidad. Por otra parte, es el punto de partida para revisar aquellos factores que inciden con mayor relevancia en la puntuación del café.

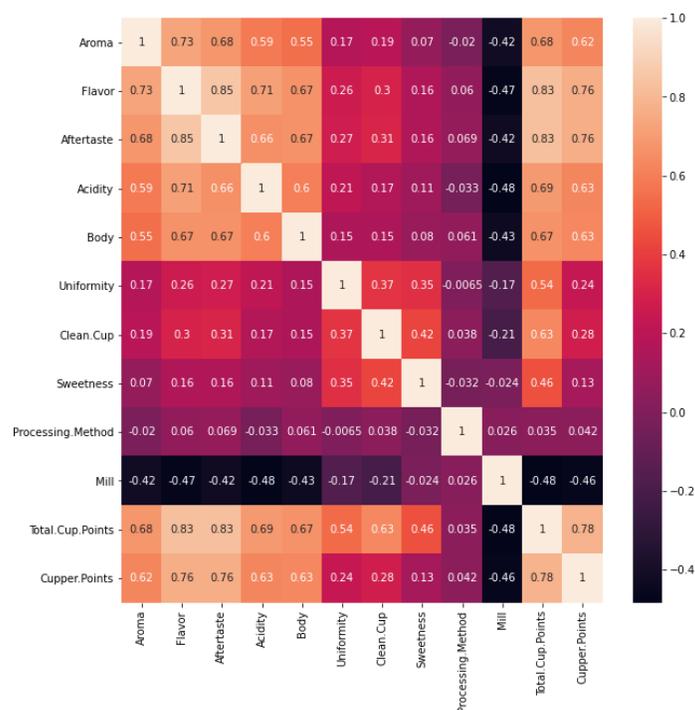
2.2.4. Factores con mayor impacto en la calidad del café según los Total Cup Points

A nivel mundial no sólo se califican los cafés con base en sus condiciones físicas previas a la molienda en donde se considera principalmente la cantidad de defectos visuales presentes por grano de café. Sino que también es necesario identificar, analizar y calificar las tazas una vez preparadas. Es ahí en donde entran a jugar características como el aroma, acidez, cuerpo, sabor, prolongación del sabor en el paladar, dulzor, entre muchos otros. A continuación, se presenta un mapa de calor el cual determina la correlación entre las variables expuestas en ella, donde se identifican las que mayor incidencia tienen al momento de determinar la calificación de una taza de café conocida en el ámbito internacional como *total cup points*.

En el apartado de *Metodología y datos* se detalla con mayor detalle el desarrollo de la matriz de correlación presentada a continuación.

Figura 11.

Matriz de correlación de las variables evaluadas en los Cup Points



Nota. Gráfico matriz de correlación de las variables que intervienen en la puntuación del café *Total Cup Points*.

Como se observa el mapa de calor evalúa un total de 12 variables, de las cuales las que tienen mayor influencia sobre la variable dependiente o en la salida (*total cup points*), fueron el aroma y la prolongación del sabor, mientras que las variables que menos influencia tuvieron fueron el tipo de molienda y el método de procesamiento.

2.3. Metodología y datos

En esta parte del documento, se muestra la metodología y el paso a paso con la que se desarrolla el presente trabajo. En primera medida, se realiza una exploración de las posibles fuentes de las cuales se podría extraer la base de datos y al mismo tiempo se

explora sobre los posibles softwares para desarrollar el proyecto de la forma más eficiente posible.

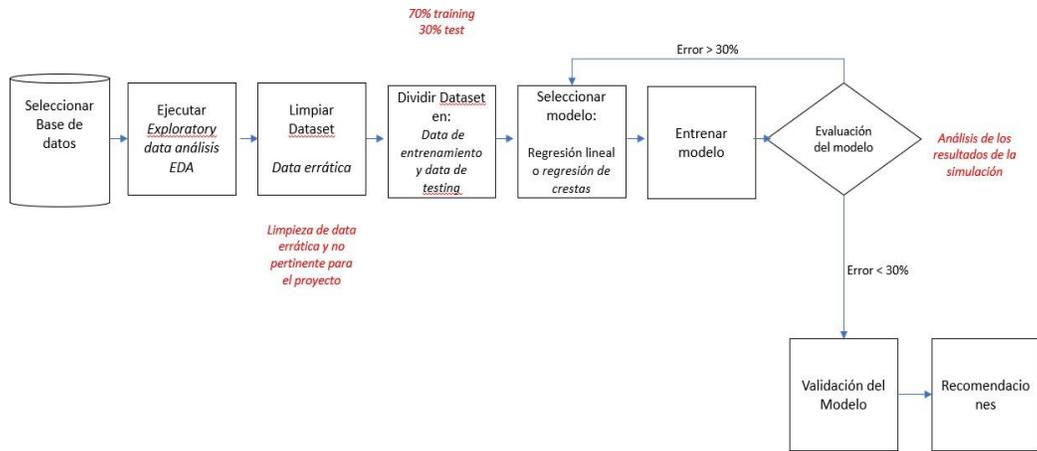
Posteriormente, se procede a buscar y seleccionar el *dataset* que cuente con información tomada a nivel mundial de diferentes fincas y países a lo largo y ancho del planeta tierra. Seguidamente, se importan las librerías dentro del *software* seleccionado que son requeridas para cumplir con el análisis exploratorio de la data al momento de importar y leer el *dataset*. De forma preliminar se identifica el tamaño de la base de datos, la cantidad de registros faltantes, data anómala o errática y tipos de variables; así mismo se observa que la data se encuentre almacenada de forma sistemática permitiendo identificar relaciones preliminares. Dentro del proceso es de vital importancia el manejo que se le dé a los registros erráticos o faltantes.

Para el trabajo en cuestión, se realizó depuración de información que no se considera pertinente, mientras se transforma los datos de tipo texto a tipo numérico con el objetivo de lograr un manejo de la información más cómodo. Seguidamente, se procede a graficar los histogramas de las variables con mayor relevancia para el proyecto, logrando identificar registros que fueron necesarios someter al proceso de limpieza y depuración. Una vez finalizado el proceso de limpieza, se realiza la verificación, para lo cual se volvió a correr histogramas y diagramas de cajas y bigotes donde se observan los resultados del antes y el después. Inmediatamente, se estructura la matriz de calor para representar gráficamente la relación entre variables cuantitativas de forma visual, empleando la intensidad de los colores para representar su nivel de relación. Para este caso, se identifican dos variables principales bajo el criterio de índices altos de correlación con la calificación en términos de calidad de los diferentes tipos de café relacionados en la base de datos.

Todo lo anterior se resume en el siguiente diagrama:

Figura 12.

Diagrama del proceso de tratamiento del dataset para la aplicación del ML



Nota. Secuencia de actividades a ejecutar para la aplicación del ML en unabase de datos

Muestra el procedimiento general de la utilización del ML en el proyecto. Inicialmente consta de la selección, análisis exploratorio, limpieza y división del *dataset*. Posteriormente, la selección de los prototipos para ejecutar entrenamiento de los registros o datos y así lograr la optimización de la información. Finalmente, se evalúa y valida el modelo para generar el planteamiento de recomendaciones en caso de que este haya sido validado de forma correcta.

2.3.1. Selección de Variables

Dentro de las múltiples bases de datos exploradas se selecciona *merged_data_cleaned.csv* de la reconocida subsidiaria de Google y comunidad de científicos de datos norteamericana Kaggle. Para ello se exploraron múltiples comunidades, dentro de las que se evaluaron variables que pudiesen ser de mayor relevancia al momento de determinar la calidad del café. De la misma forma se contemplan más de 57.200 datos (1.338 filas y 43 columnas). Dentro de las principales variables que se trabajaron en el documento, se destacan: aroma, sabor, prolongación del sabor, acidez, cuerpo y uniformidad entre otras variables determinantes. Si bien es

cierto que *la altura a la cual es cultivado y cosechado el café*, en el presente proyecto se determinó que, para la presente base de datos, influye en tan sólo un 25% en la calificación final de calidad una taza de café la altura de la finca en la que se cosecha el café. Por lo que se determinó que no se consideraría para el análisis del presente proyecto.

2.3.2. Lenguaje de programación e importación de módulos y base de datos

El lenguaje de programación seleccionado para realizar el modelamiento fue *Python 3.0*, cuya primera versión fue creada a finales de los 80s. Este lenguaje de programación hoy en día es uno de los posee mayor reputación a nivel mundial, siendo empleado por organizaciones de renombre como *Intel, IBM, NASA, Pixar, Netflix, Facebook, Spotify* que tan sólo hacen parte de los ejemplos más conocidos.

Para programar en *Python* se utilizan comandos, las cuales ejecutan una acción en específico en el ambiente de desarrollo seleccionado. Muchas de estas funciones se encuentran incluidas al interior del lenguaje de programación. Sin embargo, existen funciones, clases y variables incluidas en determinadas librerías, y para ser usadas se requiere su importación a *Python*. A continuación, se detallan las librerías importadas para el modelamiento.

- *Pandas*: Librería utilizada para el manejo de *DataFrames* y análisis de datos de forma tabular, por medio de almacenamiento de datos en filas y columnas.
- *Numpy*: Manejo y adecuación de información para ser implementada en modelos predictivos.
- *Matplotlib*: Es un conjunto de funciones que permiten que *Matplotlib* funcione de manera similar a *MATLAB*. Cada función de *pyplot* permite realizar cambios como, crear una figura, crear áreas de trazado en una figura, trazar algunas líneas en un área de trazado, etc.
- *Seaborn*: Es una biblioteca de visualización de datos construida sobre *Matplotlib* y se encuentra fuertemente integrada con las estructuras de datos que ofrece *pandas* en *Python*.

- *Warnings*: Es una biblioteca que genera advertencias para prevenir al programador de situaciones que no son necesariamente unas excepciones. Generalmente, se produce una advertencia cuando existen algunos elementos de programación obsoletos o erróneos, como palabras clave, función o clase, etc.
- *Sklearn*: *Scikit-learn-Sklearn* es una de las biblioteca más útiles y robustas para el aprendizaje en *Python*. Esta biblioteca proporciona una selección de herramientas eficientes para el ML y el modelado estadístico que incluyen funciones como clasificación, regresión, agrupación y reducción de dimensionalidad a través de una interfaz de consistencia en *Python*.
- *Sklearn.model_selection*: Es una biblioteca de *Python* que ofrece varias características para el procesamiento de datos que pueden ser utilizadas para clasificación, agrupamiento y selección de modelos.
 - *Sklearn.metrics_selection*: El módulo de métricas de *Sklearn* implementa varias funciones de pérdida, puntuación y utilidad para medir el rendimiento de la clasificación.
 - *Sklearn.linear_selection*: La regresión lineal es un algoritmo de ML basado en el aprendizaje supervisado en *Python*, permite realizar tareas de regresión y es un módulo utilizado principalmente para conocer la relación entre las variables y la previsión.

2.3.3. Análisis exploratorio de datos

Una vez importadas las librerías mencionadas anteriormente, es posible trabajar la información que se encuentra en *Dataframes* en el ambiente de desarrollo seleccionado (*Google Collaboratory Research*), por medio de *Python*. En primera medida, se realiza un análisis exploratorio de la información disponible con el propósito de comprender la organización de la data de forma general. Para ello, se ejecutó la función *df.info()* como se ve a continuación.

Figura 13.

Resultado de la línea de código df.info()

```
df.info()
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1339 entries, 0 to 1338
Data columns (total 44 columns):
#   Column              Non-Null Count  Dtype
---  -
0   Unnamed: 0          1339 non-null  int64
1   Species             1339 non-null  object
2   Owner               1332 non-null  object
3   Country.of.Origin  1338 non-null  object
4   Farm.Name           980 non-null   object
5   Lot.Number          276 non-null   object
6   Mill                1021 non-null  object
7   ICO.Number          1182 non-null  object
8   Company             1130 non-null  object
9   Altitude            1113 non-null  object
10  Region              1280 non-null  object
11  Producer            1107 non-null  object
12  Number.of.Bags      1339 non-null  int64
13  Bag.Weight          1339 non-null  object
14  In.Country.Partner  1339 non-null  object
15  Harvest.Year        1292 non-null  object
16  Grading.Date        1339 non-null  object
17  Owner.1             1332 non-null  object
18  Variety             1113 non-null  object
19  Processing.Method    1169 non-null  object
20  Aroma               1339 non-null  float64
21  Flavor              1339 non-null  float64
22  Aftertaste          1339 non-null  float64
23  Acidity             1339 non-null  float64
24  Body                1339 non-null  float64
25  Balance             1339 non-null  float64
26  Uniformity          1339 non-null  float64
27  Clean.Cup           1339 non-null  float64
28  Sweetness           1339 non-null  float64
29  Cupper.Points       1339 non-null  float64
30  Total.Cup.Points    1339 non-null  float64
31  Moisture            1339 non-null  float64
32  Category.One.Defects 1339 non-null  int64
33  Quakers             1338 non-null  float64
34  Color               1121 non-null  object
35  Category.Two.Defects 1339 non-null  int64
36  Expiration          1339 non-null  object
37  Certification.Body  1339 non-null  object
38  Certification.Address 1339 non-null  object
39  Certification.Contact 1339 non-null  object
40  unit_of_measurement 1339 non-null  object
41  altitude_low_meters 1109 non-null  float64
42  altitude_high_meters 1109 non-null  float64
43  altitude_mean_meters 1109 non-null  float64
dtypes: float64(16), int64(4), object(24)
memory usage: 460.4+ KB
```

Nota. Visualización de los datos no nulos y descripción del tipo dedato que se encuentra en cada variable de la base de datos.

Por medio de esta función, se obtiene una vista general de las variables y tipo de información que se encuentra almacenada, permitiendo maximizar el entendimiento, observar la estructura y la tipología de los datos. Además, permite identificar todas las variables a fin de determinar cuáles son críticas o no y consecuentemente plantear posibles supuestos (los cuales serán soportados por medio del modelo predictivo), entre muchos otros análisis. Para el presente caso, se identifican 43 columnas o variables de las cuales: 4 poseen data numérica entera, 16 son de carácter numérico flotante y 24 tipo objetos.

2.3.4. Limpieza de datos

Con la concepción general clara de cómo luce el *dataset*, se procede a retirar aquellos registros que no serán contemplados a lo largo del análisis. Para este caso se emplea la función *df2.drop* logrando obtener como resultado 1.339 filas por 16 columnas, para un total de 21.424 datos, lo que representa una reducción de aproximadamente del 50% de los datos inicialmente disponibles. Dentro de las variables que fueron retiradas en primera medida fueron aquellas que no permitían ejecutar un análisis según lo requerido

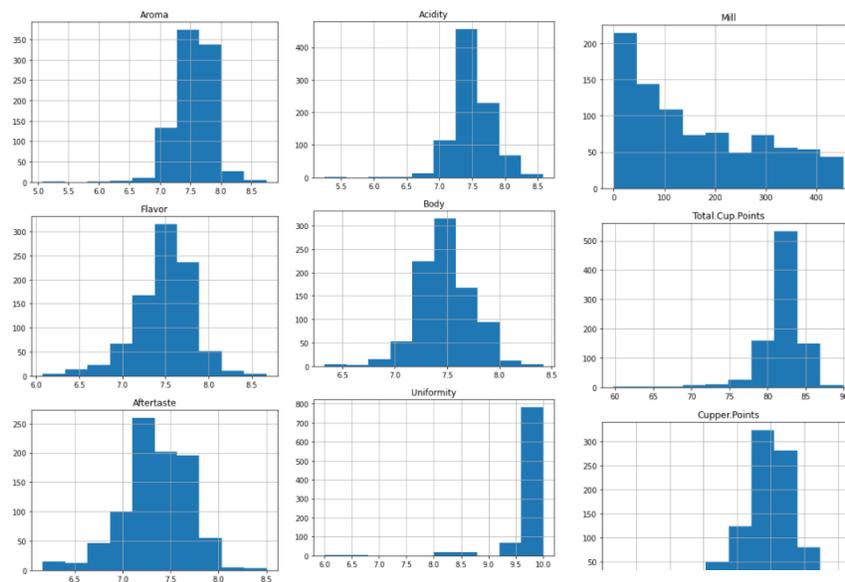
para el presente trabajo como lo es: peso de la bolsa de café, número de lote, nombre de la finca, nombre de la compañía, propietario del cultivo, entre otros.

2.3.5. Diagramas de cajas y bigotes e histogramas

Una vez con la data depurada en cuanto a la cantidad de variables, se procede a graficar por medio de histogramas aquellas de tipo flotante y relevantes bajo el criterio de prioridad en la investigación. Esto con el objetivo de identificar con qué frecuencia se repite un valor determinado en cada una de las variables (propiedades sensoriales del café, etc.) y visualizar si el comportamiento de los datos en cada una de ellas para finalmente definir si presenta una distribución normal y acorde con el tipo de información consignada.

Figura 14.

Histogramas de las variables de interés



Nota. Representación gráfica de los histogramas individuales de las variables influyentes en la calificación del café *Total cup points*.

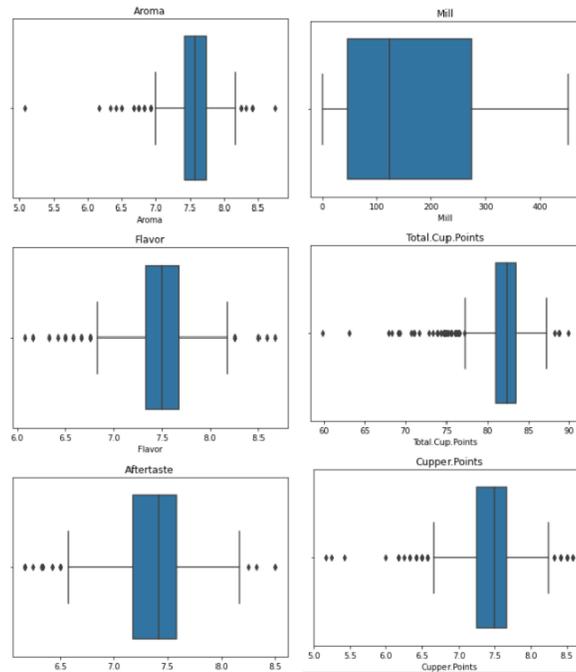
Dentro de las principales variables graficadas se destacan: Sabor, acidez, prolongación del sabor, uniformidad. Se observa que todos los *parámetros*

mencionados anteriormente tienen una distribución normal desde una perspectiva visual, por lo tanto, se considera la geometría del polígono de densidad generado por el histograma *perse*. Dicha distribución normal también es conocida como la campana de *Gauss*, la cual es de vital importancia en el mundo de la estadística porque gran variedad de fenómenos tiende a comportarse de esta forma. Sin embargo, en el caso de la uniformidad y la molienda se considera atípico su comportamiento debido a que esta última presenta un sesgo derecho indicando mayor cantidad de valores en el costado izquierdo y una menor en el costado derecho.

También, en búsqueda de identificar data errática, se hace casi obligatorio analizar la información por medio de los diagramas de cajas y bigotes. En donde a pesar de haber depurado casi el 50% de la data aún existe gran cantidad de registros en cada una de las variables que se encuentra en el exterior de los límites de los cuartiles e inclusive por fuera de los límites máximos y mínimos, como es el caso de Aroma, sabor, preservación del sabor, calificación del café, por mencionar los más representativos. A continuación, se presentan los diagramas de cajas y bigotes de las variables más representativas en el presente trabajo, en donde se evidencia el fenómeno mencionado líneas arriba.

Figura 15.

Diagramas de cajas y bigotes de las variables más influyentes en el estudio

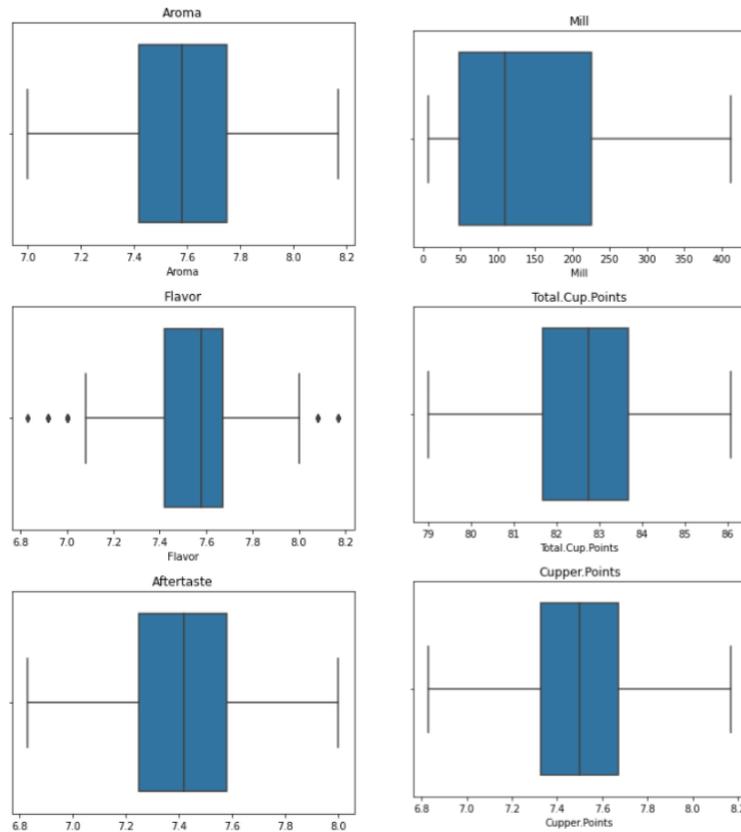


Nota. Diagramas cajas y bigotes de las variables individualmente.

Para suavizar el impacto de estos datos erráticos, se emplea el método de *Tukey* ya que este remueve todos los registros anómalos. También, muestra aquellos datos que se encuentran 1.5 veces por fuera del límite intercuartil superior o inferior. Consecuencia de esto, se obtiene la información consignada en cada variable con una distribución normal o en forma de campana, como se visualiza a continuación.

Figura 16.

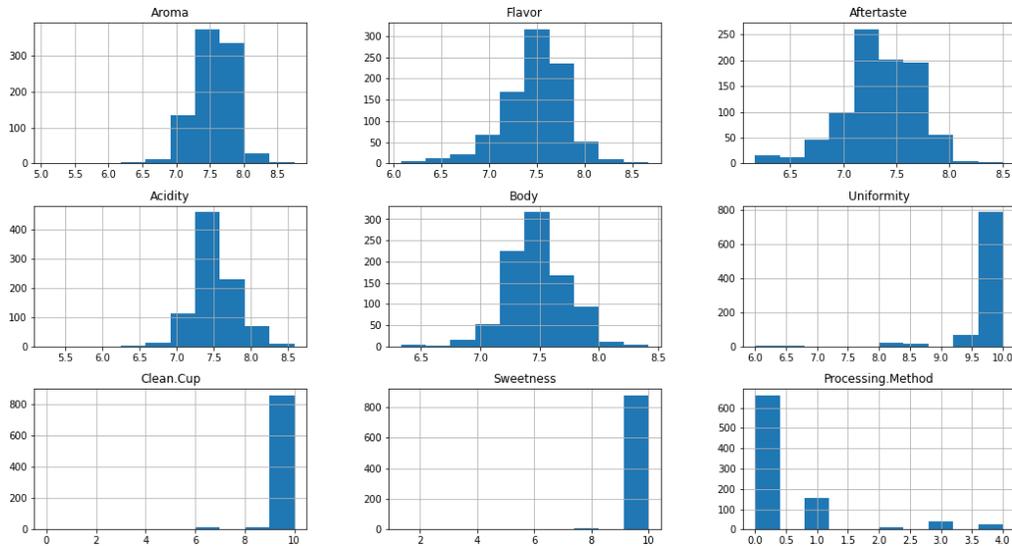
Diagrama de cajas y bigotes con el método de ajuste de Tukey



Nota. Representación gráfica de la inexistencia de data atípica en las variables estudiadas.

Figura 17.

Histogramas consolidados una vez implementado el método Tukey



Nota. Representación gráfica de los histogramas de las variables objetode estudio una vez eliminados los datos anómalos.

2.3.6. Diagrama de Violín

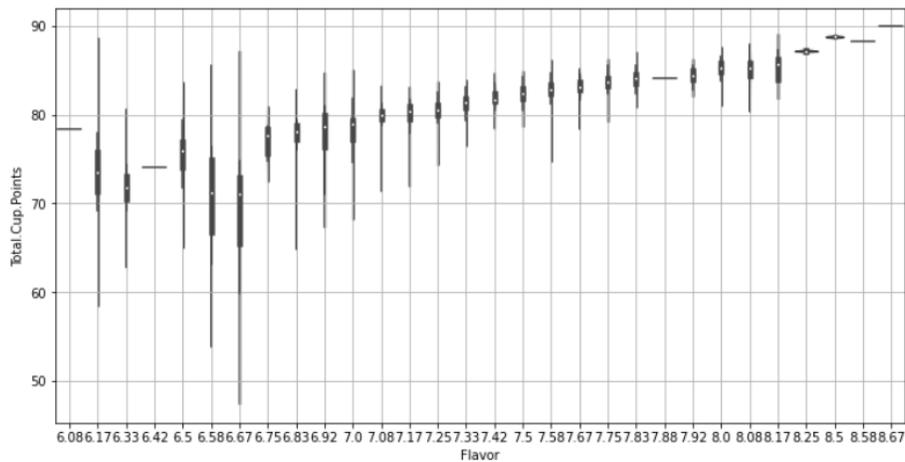
La implementación de los diagramas de violines en el proyecto se debe a que esta herramienta permite comprender de manera más detallada la distribución y variabilidad de los parámetros. En este caso puntual, se evalúa, en primera instancia, la puntuación de calidad consolidada vs. El sabor y, en segunda instancia el dulzor vs. Puntaje final de calidad del café. Esta observación se realizó con el objeto de obtener resultados más al detalle en cuanto a cómo el puntaje obtenido en la categoría de sabor y dulzor puede afectar e impactar directamente el total cup *points*.

A continuación, se presentan dos diagramas de violines, los cuales representan la distribución de los datos y la densidad de probabilidad de estos. En el primer diagrama de violín se expone en el eje x la variable Sabor, la cual según el mapa de calor está directamente correlacionada con la calificación final de la calidad del café. En el eje Y, se observa la variable *total cup points*. En este gráfico se observa en el costado superior

derecho que al tener un valor de 8.67 en el sabor de la taza de café, es casi mandatorio que se tendrá una calificación de 90 puntos en la variable *total cup points*, mientras que si se tiene una calificación de 6.67 en el sabor, el resultado al momento de determinar el *total cup points* puede variar entre un valor de desde los 45 puntos hasta un máximo de 85 puntos, lo que representa una incertidumbre muy amplia. Siendo claro que *si se obtiene un puntaje alto en el sabor se garantiza una buena calificación en el total cup points*.

Figura 18.

Diagrama de violín. Total cup points vs. Flavor (Sabor)

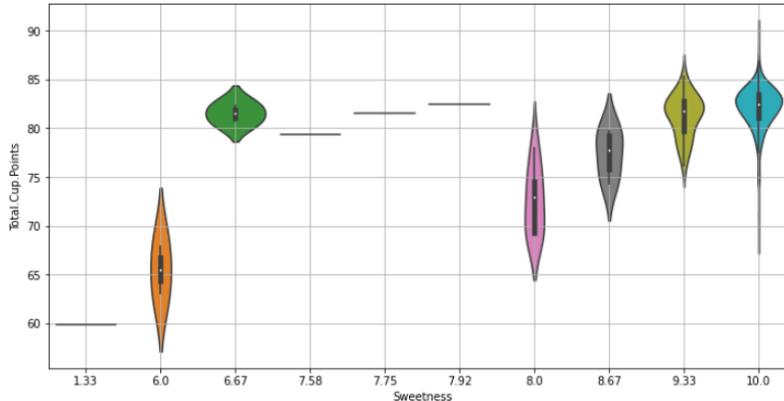


Nota. Comportamiento de la variabilidad que puede surgir en el *total cup points* por medio de la puntuación del sabor

En contraparte al gráfico de violín analizado anteriormente, en el diagrama a continuación se observa que la relación entre dulzor (eje X) y el *total cup points* (eje Y) aunque se obtenga una alta calificación en el dulzor, no se garantiza una alta calificación en el *total cup points*, indicando una baja correlación entre las dos variables numéricas evaluadas.

Figura 19.

Diagrama de violín. Total cup points vs. Sweetness (Dulzor)



Nota. Variabilidad que puede surgir en el *total cup points* por medio de la puntuación del dulzor

2.3.7. Data de entrenamiento y data de prueba

Una vez determinadas las variables de mayor importancia al momento de evaluar la calidad el café, se procede a determinar cuál representa óptimamente el comportamiento de las variables seleccionadas. Esto separando la base de datos en 2 partes, la sección de entrenamiento a la cual se le asigna el 70% de la data (ya limpia), mientras que el 30% de data restante se empleará en la evaluación. En palabras prácticas, ese 30% restante indicará el *score* que será el indicador determinante al momento de catalogar el nivel de precisión del modelo en evaluación.

2.3.8. Modelo Predictivo

El prototipo predictivo seleccionado para el desarrollo del estudio fue regresión lineal, que corresponde a un método estadístico que busca modelar la relación lineal entre una variable continua, denominadas también *respuesta o dependiente (X)* y una o más variables independientes denominadas también *regresores o predictores (Y)*.

En este caso se determina que las variables predictoras serían todas las consideradas en la matriz de correlación, mientras que la variable *total cup points*, se comportaría como la dependiente, con el objetivo de determinar qué tanto influye en la predicción de estas variables independientes al momento de pronosticar el valor o calificación que se le dará a una taza de café según la base de datos en estudio. La siguiente tabla contiene las variables puntuales definidas para el algoritmo.

Tabla 3.

Variables del modelo de regresión lineal

Variables dependientes X	Variable independiente Y
a) Aroma b) Flavor c) Aftertaste d) Acidity e) Body f) Uniformity g) Sweetness h) Balance	<i>Total cup points</i>

Nota. Exposición de las propiedades sensoriales del café como variables de respuesta y puntuación total como independiente.

2.3.8.a. Explicación del algoritmo de regresión lineal en Python. La librería utilizada para el funcionamiento del modelo *scikit learn* establece las instrucciones para el algoritmo de regresión lineal. A continuación, se muestran las líneas de código que conforman el algoritmo implementado para la investigación.

```

Modelo_Resgresion_Lineal=LinearRegression(normalize= True)
Modelo_Resgresion_Lineal.fit(x_train,y_train)
Prediccion_Regresion_Lineal= Modelo_Resgresion_Lineal.predict(x_test)
Print("score:",r2_score(y_test,Prediccion_Regresion_Lineal)*100,"%")

Score: 91,94262063979495 %

Modelo_Resgresion_Lineal.coef_

Array([[ 0.89882659, 1.71433179, 1.68890289, 0.65839615, 0.8857672,
1.36013834, 1.53649471, 1.58404212]])

Modelo_Resgresion_Lineal.intercept_

Array([-2.18017044])

Modelo_Resgresion_Lineal.predict([[8.75, 8.67, 8.5, 8.58, 8.42, 10, 10, 8.42]])

Array([[90.31465736]])

```

Garantiza que las variables receptoras, es decir, las X se normalizarán antes de ejecutar la regresión.

Arroja los coeficientes estimados de la regresión.

Término independiente del modelo lineal.

Predice Y en el modelo lineal.

En el recuadro anterior lo que se encuentra en color verde corresponde a la solución de *Python* al correr el comando dictado. Para comprender un poco más sobre la función de los bloques de instrucciones, el primero permite al programa dividir la data de entrenamiento y prueba, y adicionalmente se obtiene la exactitud del modelo, que para este caso es del 91.94%. La segunda línea determina todos aquellos coeficientes de cada una de las propiedades sensoriales del café evaluadas. La tercera muestra la intercepción en el eje de la puntuación final de calidad en el café, es decir, el *eje Y*. Finalmente, la cuarta línea evalúa el modelo matemático a partir de la asignación de puntuaciones obtenidas en la catación del café por cada propiedad sensorial definida en la tabla 3. Estos valores fueron extraídos de la siguiente tabla.

Tabla 4.

Registros evaluación calidad del café

	Aroma	Flavor	Aftertaste	Acidity	Body	Uniformity	Clean.Cup	Sweetness	Processing.Method	Mill	Total.Cup.Points	Cupper.Points	Balance
1	8.75	8.67	8.50	8.58	8.42	10.0	10.0	10.0	0	0.0	89.92	8.58	8.42
4	8.25	8.50	8.25	8.50	8.42	10.0	10.0	10.0	0	0.0	88.83	8.58	8.33
6	8.42	8.50	8.33	8.50	8.25	10.0	10.0	10.0	0	2.0	88.75	8.50	8.25
9	8.08	8.58	8.50	8.50	7.67	10.0	10.0	10.0	1	4.0	88.25	8.50	8.42
18	8.42	8.25	8.08	8.17	7.92	10.0	10.0	10.0	0	5.0	87.25	8.42	8.00

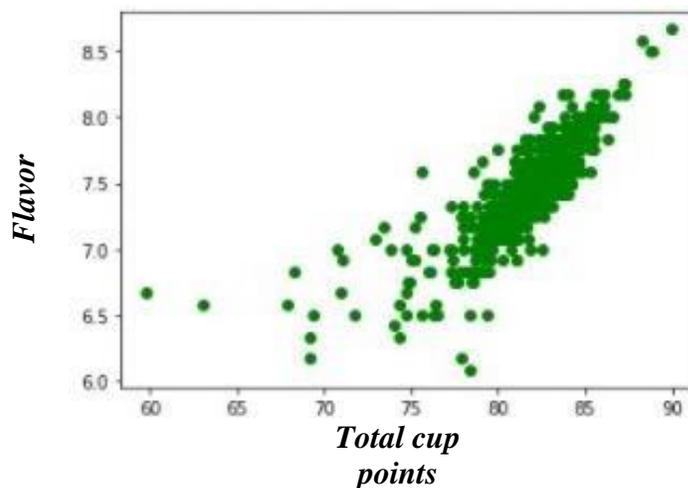
Nota. Información extraída del *dataset* trabajado a lo largo del proyecto.

Como se puede visualizar, en el registro #1 la calificación final en la calidad del café evaluado fue de 89.92 puntos, mientras que el algoritmo con una precisión mayor al 90% predijo una puntuación de 90.31 puntos. Por lo cual, se afirma que el modelo en cuestión es bastante acertado y preciso pues indiscutiblemente se acerca bastante al valor real.

En adición, se graficó la tendencia de los datos de la variable *Flavor* vs. *Total cup points* con el objeto de verificar que efectivamente las valoraciones presentan un comportamiento lineal.

Figura 20.

Gráfica Flavor vs. total cup points



Nota. Distribución grafica con tendencia lineal de las puntuaciones *flavor*

Desde la perspectiva matemática, así luce la función obtenida.

$$f(Xn) = 0.8988 X_1 + 1.7143X_2 + 1.6889X_3 + 0.6584X_4 + 0.8858X_5 \\ + 1.3601X_6 + 1.5365X_7 + 1.5840X_8$$

Donde:

$X_1 \rightarrow$
Aroma X_2
 \rightarrow *Flavor*
 $X_3 \rightarrow$
Aftertaste X_4
 \rightarrow *Acidity*
 $X_5 \rightarrow$ *Body*
 $X_6 \rightarrow$
Uniformity X_7
 \rightarrow *Sweetness*
 $X_8 \rightarrow$ *Balance*

A partir de los efectos generados por la regresión lineal al ser probado con datos experimentales sobre puntuaciones de calificación de calidad del café con base en las propiedades sensoriales, se concluye que el modelo se ajustó eficientemente a la data disponible puesto que tiene una precisión del 91.94% en pronosticar la calificación final a partir de registros históricos. Aun así, cabe aclarar que cada vez que el código se prueba presentará ciertas variaciones debido a la cantidad de datos y las diversas iteraciones que realiza para el éxito del modelamiento. De igual manera, la fluctuación de los valores es mínima y esto proporciona confiabilidad en los resultados.

Toda la información mencionada en este apartado del documento, fue tomada del modelamiento en Python (Ver anexos).

2.3.9. Resultados y validación del modelo predictivo

Una vez el algoritmo fue probado y analizado, es claro que tanto *el sabor del café como la cualidad de perdurar en el paladar son las dos variables o parámetros que tienen mayor importancia al momento de determinar la calidad de una taza*, por lo que sin lugar a duda son los dos principales factores en los que deben enfocarse los caficultores colombianos para adquirir altos puntajes en la calidad del café que cosechan. Para ello se emplearon dos algoritmos: la regresión lineal obtuvo un score de 91,94% y la regresión de crestas un 84.45%. Por lo tanto, el algoritmo que proporciona mejores resultados en predicción es el de regresión lineal.

2.3.10. Factores de alto impacto en la LA a partir de los lineamientos de calidad del café según el total cup points de la data en estudio

A partir de lo obtenido por medio de la distinción de correlación de aquellas medidas sensoriales influyentes en la categorización del café con alta calidad (*Total Cup points*), se pretende enlazar dichos factores de mayor influencia, los cuales son sabor y *after taste* con los eslabones que componen la LA en el SC colombiano. Para ello, a través de la revisión de diversas investigaciones a nivel nacional y global, se expondrán a continuación los componentes que pertenezcan a la LA que interfieren en la obtención de un excelente sabor y *after taste*.

En búsqueda de respuestas concretas que conlleven a la relación de los métodos de fabricación con el sabor y aroma del café, la *universidad A&M* especialista en análisis de alimentos a través de la tecnología menciona diferentes causas que afectan tanto el aroma como el sabor del café. Sin duda alguna la escogencia de la tierra o lote en donde se instalará el cultivo de los cafetos es trascendental ya que el estudio arroja un aumento en la calidad del grano cuando las matas de café se encuentran a gran altura, explicando que, al tener mayor altura se prolonga aún más el tiempo de maduración y así se obtendrá un café con un mejor sabor. Del mismo modo, *la neutralización del suelo en cuanto al pH balanceado y la tierra rica en nutrientes desencadena la producción de un café de alta calidad*. Ya que el objetivo principal de las cerezas de café es alcanzar finalmente su estado de maduración, es vital contar con recursos naturales como el agua y la luz natural. Por otra parte, afirman que la definición de un plan preciso sobre los métodos de procesamiento ayudará al caficultor a comprender y visualizar los resultados que obtendrá en la cosecha, y de manera anticipada, puede pronosticar cuál será la calidad del grano. Con respecto al método de recolección de las bayas maduras, el manual se considera más efectivo que el automático para conservar la calidad del café ya que permite seleccionar puntualmente aquellos granos en condiciones de 100% maduración. En conclusión, las condiciones ideales para obtener café de alta calidad son la exposición del cultivo a una temperatura que se encuentre en el rango de 16 a 20°C, humedad relativa del lugar entre 75 al 95% y la precipitación natural del suelo esté de 2.000 mm a 3.000 mm [55].

Ahora bien, desde una perspectiva más técnica de la agricultura, un estudio realizado en Etiopía sobre *los granos de café Arábica halla la fuerte correlación del nitrógeno con la calidad del fruto ya que aporta una excelente conductividad del agua hacia las raíces de los cafetos, mejorando de cierta forma la eficiencia del agua dentro de la planta y a su vez la incorporación de los nutrientes en las raíces*. Por medio de la observación del comportamiento de los frutos expuestos a temperaturas entre los 12 y 29° C, y una precipitación anual de 1.300 a 1.700 mm se demuestra a través de la puntuación de los parámetros sensoriales (Balance y sabor) y la altitud, cómo el sabor presenta una tendencia positiva al incrementar la altitud a la cual fue cosechado [56].

En cuanto al puntaje de cata del café, el cual es el eje principal del presente estudio, se menciona como factor fundamental la escogencia de la semilla teniendo en cuenta el origen de la misma ya que se demostró que la calidad depende de este factor y no de la variedad como tal. *Así como tampoco interfiere significativamente el tamaño del grano con los parámetros evaluados en la calidad de una taza de café* [57].

Según la enciclopedia *Food Chemistry*, en el proceso de tostado de los frutos, estos sufren alteraciones en su estructura química y física ya que extrae la humedad en cierta proporción [58]. Siguiendo esta premisa, para evitar el riesgo de extracción descontrolada de la humedad de los granos y desde la perspectiva de la LA, es vital contar con una máquina tostadora en excelentes condiciones de calibración y mantenimiento para no sobrepasar la extracción de agua y así no afectar negativamente la calidad.

Con base en esta revisión de resultados de estudios anteriores, es posible reafirmar cómo se logra enlazar la calidad de los frutos cafeteros y la influencia de algunos elementos en la LA.

2.3.11. Recomendaciones para la selección y evaluación de los proveedores

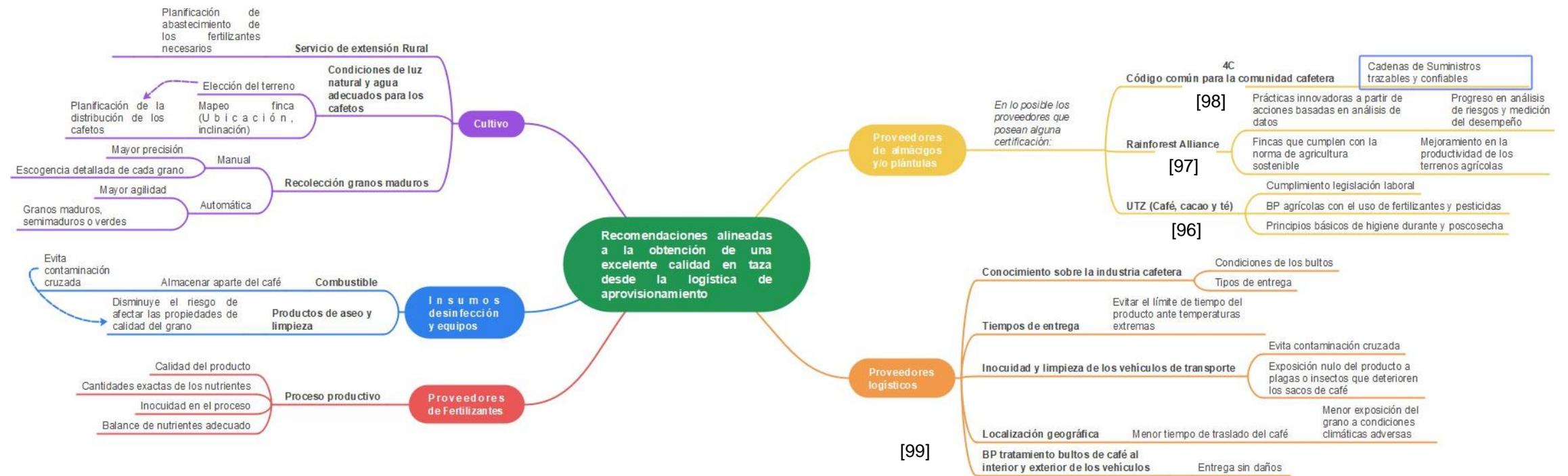
En este punto es donde finalmente se consolida todo lo ejecutado durante el presente

proyecto ya que es el resultado final que generará valor a la realidad de los caficultores en miras que incrementar su productividad sin perder de vista la calidad. Teniendo en cuenta que, para obtener calificaciones altas en la evaluación de las propiedades sensoriales presentes en el café hay que centrarse en todo aquello que influye en la prolongación en boca y notas diferenciales en el sabor. Por lo tanto, se analizaron distintos parámetros a considerar en los proveedores de materias primas, recursos y servicios que hacen parte de la *LA*.

A continuación, en el siguiente diagrama se aterrizan las recomendaciones.

Figura 21.

Diagrama recomendaciones para la LA desde la prolongación y percepción del sabor



Nota. Recopilación de aquellas recomendaciones alineadas a la LA desde la prolongación y detección del sabor de una taza de café para disminuir totalmente el riesgo de generar acciones que repercutan negativamente en la calidad del café.

En el diagrama se puede observar cómo se dividen las recomendaciones dictadas a través de todo el abordaje de información y análisis realizado en el proyecto. Las ramas son cinco. La primera hace referencia al cultivo, en donde se exponen 3 recomendaciones principales: contratación indispensable del servicio de extensión rural, La tierra que cuente con luz natural y agua, y finalmente en cuanto al método de recolección de granos maduros para gestionar aquellas herramientas y ejecutar este proceso.

Por parte de los insumos de desinfección y equipos, se contempla el adecuado almacenamiento tanto del combustible como de los productos de limpieza para evitar la contaminación cruzada que se puede dar con los bultos de café, en palabras prácticas, destinar un sitio específico para ubicar estos insumos.

Entrando en materia de proveedores, se tienen 3 categorías: fertilizantes, almácigos y plántulas de café, y logísticos. En los proveedores de fertilizantes únicamente se recomienda realizar el seguimiento al proceso productivo ya que permitirá identificar si las proporciones que se están utilizando para fabricar ciertos nutrientes compuestos son las correctas y así no utilizar fertilizantes que no se encuentren totalmente balanceados. En cuanto a los proveedores de almácigos y/o plántulas, se hace énfasis en la obtención de certificaciones que en conjunto validan que las cadenas de suministro son confiables y trazables, también la capacidad de los proveedores al incluir prácticas innovadoras e implementación de los principios básicos de inocuidad y desinfección; pero fundamentalmente se identifique si se realiza una agricultura sostenible.

Finalmente, para los proveedores logísticos se plantean 5 recomendaciones todas estas enfocadas al excelente tratamiento de los bultos de café, creación de condiciones seguras que no afecten la integridad del grano y no se generen daños en el transporte.

3. MARCO REFERENCIAL

3.1. Marco conceptual

En este apartado del documento, se mencionan los conceptos y términos clave que abarca el proyecto con el fin de tener claridad sobre la información que será abordada a lo largo del desarrollo de la investigación.

3.1.1. *Calidad de Café*

A lo largo del tiempo se le han dado diferentes connotaciones al término calidad. Según el renombrado ingeniero y profesor *W. Edwards* la calidad hace referencia a traducir las necesidades futuras de los usuarios en características medibles. Solo así, un producto puede ser diseñado y fabricado para dar satisfacción a un precio que el cliente pagará [59]. En palabras prácticas, lo que no se puede medir no se puede mejorar, de tal forma que, si no es posible medir lo que busca el cliente, no se logrará llegar a ello, lo que aplicaría no solo a la industria del café, sino a muchas más.

Al momento de hablar de calidad en el marco del gusto comúnmente se suele pensar de forma muy subjetiva, y es que, al momento de experimentar placer, la subjetividad juega un papel muy importante. Es por ello, que surge la *Coffee Quality Institute CQI*, organización encargada de evaluar y clasificar la calidad del café de forma objetiva e imparcial, por medio de un sistema estandarizado que permite a profesionales de diferentes países discutir en un lenguaje común qué café es merecedor a la insignia de café especial y cual no. A continuación, una breve explicación del proceso. Los jueces o profesionales acreditados por la *CQI* brindan una calificación de 0 a 100 puntos en donde 100 es la máxima calificación y 0 es la mínima. Todo café que obtenga una calificación de más de 80 puntos tiene el título de café especial, mientras que los que tengan una calificación menor, obtienen una calificación de café tradicional. Continuando con este lineamiento de ámbito internacional, se procede a catalogar la muestra de data estudiada en el presente proyecto conforme a lo expuesto por la *CQI*.

3.1.2. Algoritmo

En el contexto de la tecnología, la definición de algoritmo se convierte en la base fundamental para la creación de conjuntos de códigos que adopten nuevas formas de cubrir las necesidades para las cuales fueron programadas. A lo largo de todo el desarrollo de la presente investigación, es clave comprender sencillamente este concepto ya que todo girará en torno a los algoritmos. A continuación, se exponen algunas definiciones.

Harsh Bhasin de la universidad de *Oxford* menciona que los algoritmos conllevan una serie de pasos consecutivos para realizar una actividad o tarea específica. Adicional a ello se componen de una entrada, el proceso de interacción y por último la salida. A través del proceso se alimenta el algoritmo y se obtiene el resultado objetivo [60]. Complementando un poco lo expuesto por el autor anterior, *N. Umoja* parte de la idea de visualizar los algoritmos como un conjunto de pasos precisos que permiten ejecutar el mismo en un computador. Además, se da solución óptima a la problemática identificada en primera instancia por medio de las herramientas computacionales [61]. Con base en las anteriores definiciones de algoritmo, hasta este punto se visualiza este concepto desde la construcción de procedimientos consecutivos para cumplir un objetivo final.

Pero en definitiva, *A.Maldonado* y *E.Sestari* lograron puntualizar claramente los algoritmos como “Secuencia lógica de procesamiento de datos inmateriales que buscan lograr un objetivo” [62].

Al estudiar las definiciones anteriormente mencionadas, se define como concepto líder de la presente investigación el expuesta por *Harsh Bhasin* debido a que trata minuciosamente los elementos clave de un algoritmo y a su vez, va muy ligado al *ML*, el cual es el tema central del proyecto.

3.1.3. Inteligencia artificial

El mundo ha venido revolucionando la forma de ejecutar las cosas, y en busca de la competitividad y desarrollar nuevas maneras de crear, surge el concepto de *IA*. Para

nadie es un secreto que cada día más organizaciones y entidades a nivel mundial encaminan todos sus procesos hacia la innovación. Por esta razón, es preciso mencionar varios autores que definen la *IA* desde distintas perspectivas.

Dentro de la indagación exhausta de varios términos de la *IA* se encontraron varias definiciones de expertos, en donde cada uno habla desde su experticia en la temática. Una de las interpretaciones relevantes para la presente investigación es hablar de la *IA* como la referencia principal que pretende imitar muy de cerca la inteligencia humana. Desde el punto de vista técnico, se visualiza la *IA* como el desarrollo consecutivo de interacciones informáticas que conllevan al descubrimiento de estructuras semejantes por medio del entendimiento del comportamiento del cerebro humano [63]. En efecto la academia ha tocado bastante la *IA* ya que es el hito que ha desencadenado el acrecentamiento de las industrias y al mismo tiempo de las sociedades. Por tal razón las instituciones abarcan este concepto como una vertiente de la informática que contempla los comportamientos inteligentes desde la automatización. Ahora bien, hay que observar la *IA* como el sistema especializado en analizar, interpretar e intuir los seres inteligentes para tomar decisiones acertadas y eficientes. Adicionalmente, es una tecnología diseñada especialmente para ser desarrollada en computadores que contengan estructuras similares al pensamiento humano y pueda ser replicable en robots y softwares inteligentes [64] .

Observando cómo el concepto de *IA* ha venido evolucionado al transcurso de los años por medio de las exigencias del mundo globalizado y enmarca una *IA* aplicada a las necesidades del mismo, se puede concluir que la *IA* es una tecnología que abarca la utilización de algoritmos matemáticos con el fin de realizar una simulación similar a las habilidades del cerebro humano.

3.1.3.a. Machine Learning. Considerando que el ML comprende un panorama amplio de aplicaciones, este ha sido definido desde varios puntos de vista. En el año 2009, E. Olivas y J. Guerrero definen este concepto como la tecnología que enlaza en tiempo real la creación, diseño y posterior desenvolvimiento de algoritmos estructurados que garantizan el aprendizaje de los dispositivos tecnológicos, especialmente de los

computadores. Adicionalmente, afirman que el ML es un subcampo de la IA [65].

Por otra parte, una de las compañías más importantes a nivel mundial, la empresa multinacional estadounidense de tecnología y consultoría IBM, habla del *ML* como una expresión de la *IA* que fomenta el aprendizaje de los sistemas a partir de data documentada en lugar de la programación explícita, con el objeto de alimentar un algoritmo que resulte más preciso. Para aplicar realmente el *ML*, se maneja el término de modelo como la salida de información generada a partir de la alimentación del algoritmo con datos históricos. Un ejemplo muy puntual es cuando se pretende realizar un algoritmo que arroje como resultado final un pronóstico, este se podrá hacer bajo un modelamiento predictivo, cuando reciba data se obtendrá el resultado entrenado. [66]

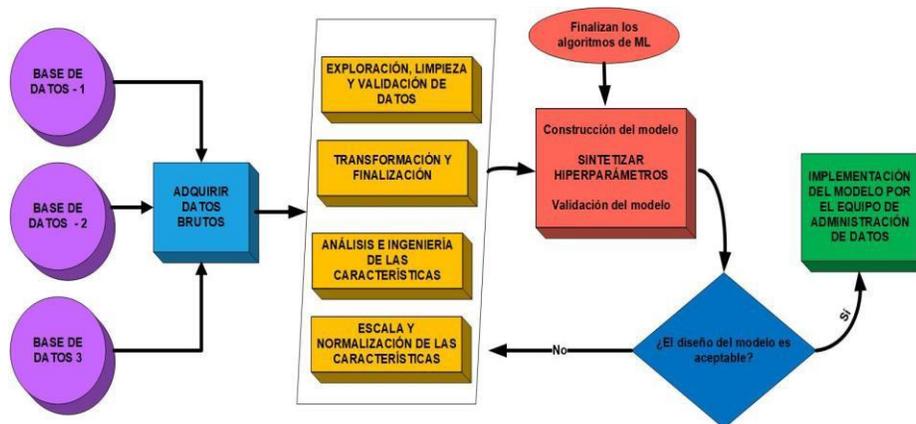
También es importante resaltar el aporte que realizó el investigador científico del centro de MIT de logística y transporte, Daniel Merchán dentro de la convención del año 2018 sobre los fundamentos del aprendizaje automático, donde define el *ML* como “máquinas capaces de realizar funciones cognitivas asociadas con la mente humana” [67]. El *ML* enmarca: el reconocimiento de voz, procesamiento del lenguaje natural, visión por computadora y la robótica, como un subcampo de la *IA*, pero aclara que el campo más dominante es el *ML*, argumentando que la usabilidad de datos pasados promueve y construye modelos idóneos que permiten a futuro realizar pronósticos de data no existente. Es por esto que se menciona el uso de prototipos probabilísticos y algoritmos que requieren evaluación estadística para analizar su grado de confiabilidad a partir de los siguientes pasos de obtención de una aplicación de aprendizaje automático:

- Identificación del problema
- Limpieza de los datos
- Implementación del modelo
- Entrenamiento y pruebas
- Evaluación y Despliegue
- Actualización constante del modelo

Con el objetivo de visualizar el *ML* de manera más sencilla y concreta, V. Reddy y S. Kadre [68], especialistas en el análisis de datos a través de la tecnología, exponen el siguiente procedimiento para la aplicación del *ML*.

Figura 22.

Procedimiento para la aplicación del ML



Nota. Diagrama explicativo para analizar los parámetros de implementación del ML. Tomado de: *A typical machine learning project in actual practice*. V. Reddy Konasani and S. Kadre, *Machine Learning and Deep Learning Using Python and TensorFlow*, 2021.

Analizando e interpretando todas las definiciones anteriormente mencionadas sobre el ML, para el presente proyecto se habla del ML como el campo de la IA que permite obtener datos futuros a través de algoritmos entrenados con data histórica.

3.1.4. Cadena de suministros

Cuando se habla de la cadena de suministros CS casi siempre se llega a la conclusión de que se encuentra en todo el entorno productivo y económico, pero ¿Qué es realmente la cadena de suministros? Para responder a esta pregunta, a continuación, se exponen algunas definiciones propuestas por varios autores especializados en la temática.

La OEM (*Original Equipment Manufacturer* de Estados Unidos) habla sobre la CS desde

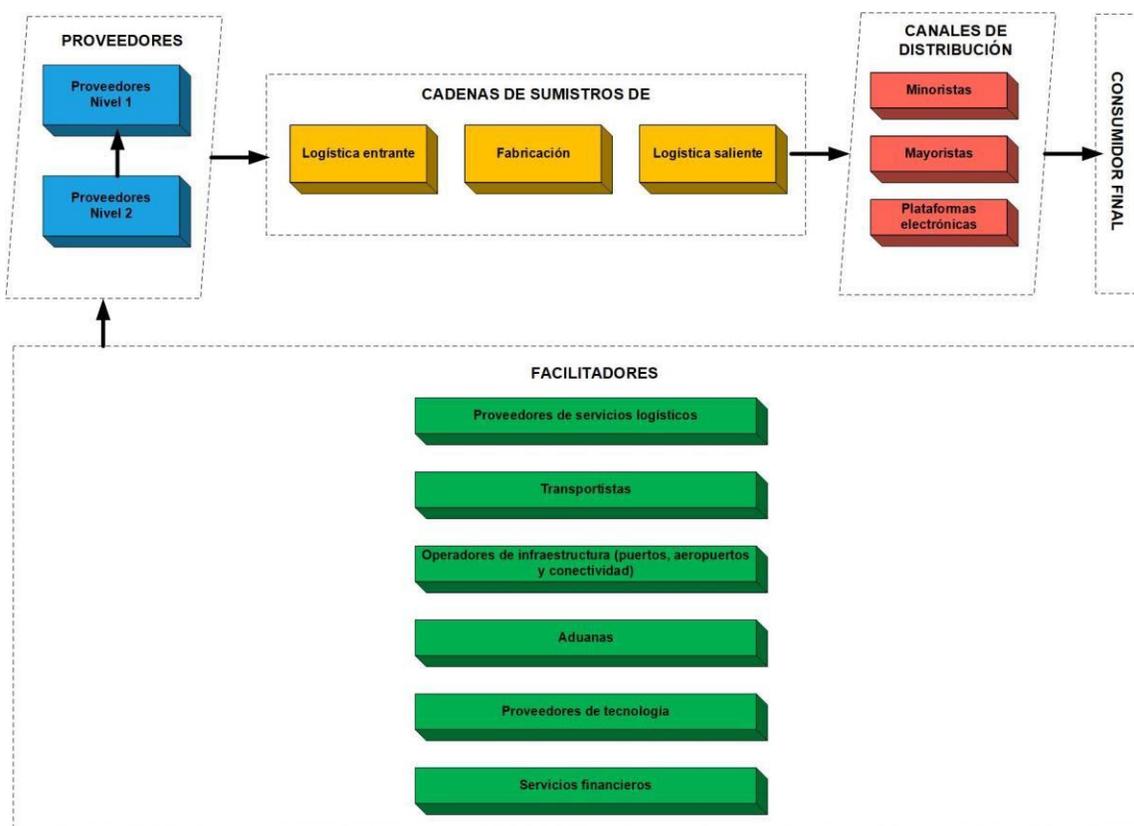
el enfoque de construcción de asociaciones ya que los consumidores y proveedores realizan un trabajo en conjunto para suplir sus propios objetivos, lo que desencadena la ejecución de actividades que conlleven a ofrecer bienes y servicios que suplan la demanda [69]. Así mismo, Ronald Ballou especialista en Logística explica la CS como la relación de actividades dentro del proceso productivo del bien o servicio que permiten la transformación de los recursos y adicionalmente añaden valor a los demandantes [70].

Para visualizar la CS desde un panorama más puntual y confiable, es preciso mencionar el concepto expuesto en el Manual Práctico de Logística, donde se habla sobre cómo la CS directamente involucra en su ejecución talento humano, desarrollo de nuevas tecnologías, procesos de negociación e instalaciones que hacen posible la fabricación de bienes desde la manufactura de la materia prima y posteriormente son vendidos a los consumidores [71]. De modo que se ha hablado sobre la CS desde un enfoque procedimental e ingenieril, para abarcar y tomar varias definiciones que enriquezcan el entendimiento de la naturaleza de la CS, el Banco Interamericano de Desarrollo *BID* contextualiza y afirma cómo las economías actuales indirectamente se transforman en CS ya que, para llevar a cabo los objetivos, se estructuran las actividades a ejecutar desde el prototipo de un producto o servicio hasta que se realizan los despachos a los clientes [72].

Además de los conceptos anteriormente mencionados, el *BID* aborda algunos actores principales de la CS. Estos son:

Figura 23.

Actores principales de la cadena de suministros



Nota. Esquema sobre la estructura de una cadena de suministros general. Tomado de: BID. *Cadena de suministro 4.0: Mejores prácticas internacionales y hoja de ruta para América Latina. 2019.*

Estos factores permiten el flujo de la información en el desarrollo de la CS. Analizando cada concepto que a lo largo del tiempo han surgido con base en las exigencias del mundo, se asume como definición principal de la cadena de suministros la adoptada por el BID porque en pocas palabras define perfectamente el alcance del concepto y a su vez, es la que posee mayor afinidad con la presente investigación.

Teniendo claro los actores que influyen en CS, para el desarrollo del proyecto sólo se tiene en cuenta la logística de aprovisionamiento o entrada ya que es la que directamente impacta al objetivo general planteado.

3.1.4.a. Logística de aprovisionamiento. La mayoría de las empresas a nivel mundial, actualmente se encuentran en la búsqueda de generar valor y así mismo ser competitivos en un mercado que cada vez más se vuelve más exigente. Este objetivo lo cubre totalmente la logística de entrada o aprovisionamiento ya que es el pilar que permite la excelente ejecución de la CS. Por ser un término tan conocido en la industria, existen varios enfoques para definirlo.

A continuación, se exponen los más relevantes para el proyecto:

Tabla 5.

Definiciones sobre logística de aprovisionamiento

AUTOR	DEFINICIÓN
<p><i>California's Manufacturing Networks</i> CMTC 2018</p>	<p>"Los actos de adquisición de material, manipulación, almacenamiento y transporte." [73]</p>
<p><i>Chartered Institute of Procurement & Supply</i> CIPS 2021</p>	<p>"La logística es el flujo de fondos, bienes e información entre origen y consumo. La logística involucra información, manejo de materiales, producción, empaque, inventario, transporte, almacenamiento y, a menudo, seguridad." [74]</p>
<p><i>Council of Supply Chain Management Professionals</i> CSCMP 2021</p>	<p>"El movimiento de materiales de proveedores y vendedores hacia los procesos de producción o almacenamiento. Instalaciones." [75]</p>

Nota. Tabla definiciones sobre el concepto de logística de aprovisionamiento desde la perspectiva de varios autores.

Adicional a las anteriores definiciones planteadas por entidades internacionales reconocidas en el campo de la CS, *Massachusetts Institute of Technology MIT* menciona que la LA abarca todo el proceso de abastecimiento de materias primas hasta que se dispone de la misma en el área de producción. También, tiene en cuenta aspectos como: plan de materiales y recursos que permiten cumplir con un rango específico de unidades a producir y además se encuentre en la capacidad de cubrir las alteraciones de la demanda; sistema de gestión de la entrega y el transporte del material [76]. En el contexto de la investigación, se tendrá en cuenta la anterior definición debido a que menciona los aspectos que se impactarán a través de la creación del algoritmo.

3.1.5. Productividad

En el entorno cotidiano y empresarial es bastante común escuchar el término de productividad, pero, realmente ¿Qué es productividad? En esta sección se expone la variedad de definiciones dispuestas por autores conocedores del tema.

La universidad del CEMA ubicada en Buenos Aires, Argentina define la productividad como un indicador que tiene como objetivo manifestar el grado de utilización correcta de los recursos en el contexto de la producción de bienes tangibles e intangibles. Adicional a esto, parte de la relación entre la producción obtenida dentro de un periodo determinado y los medios utilizados [77].

Así mismo, el Banco Mundial habla de la productividad como la capacidad que posee la economía para transformar eficientemente los factores de producción tales como el capital, mano de obra, entre otros. Además, enmarca dos conceptos fundamentales: El valor agregado por trabajador y productividad total de los factores. El primer concepto se encuentra a partir de la división entre ingresos recibidos por concepto de la producción y el número de empleados totales, convirtiéndose en un indicador de la productividad laboral [78] mientras que el segundo concepto abarca más de 2 factores que pueden influir en el resultado de la productividad, estos son definidos por las empresas de acuerdo con los objetivos planteados.

Otro aporte importante para la productividad lo realizó el especialista en Administración de Negocios y Magíster en Investigación y Docencia Universitaria de la Universidad

Sergio Arboleda, Jorge Medina en la definición de un modelo integral de productividad donde nombra este término como el método para utilizar “los factores de producción en la generación de bienes y servicios para la sociedad” con la finalidad de aumentar la eficiencia y eficacia de utilización de los recursos. Consecuentemente, indica que este concepto se puede observar como un objetivo estratégico de las organizaciones porque sin esta medición es imposible lograr cuantificar el nivel necesario para mantenerse en un mundo tan competitivo y globalizado [79].

Todos los aportes de autores que enmarcaron el concepto de productividad son de gran relevancia para la presente investigación, pero, en definitiva, la guía será la expuesta por el Banco Mundial porque considera factores bastante importantes en el marco de la productividad y son la base para desarrollar el objetivo del proyecto.

3.2. Marco Teórico

Con el objeto de profundizar un poco más en la teoría fundamental y que aporta valor a la presente investigación, en el siguiente apartado se describen las temáticas centrales desde un punto de vista más específico.

3.2.1. Producción y calificación del café

El proceso de producción de café es largo, delicado y cauteloso, inclusive llegando a ser catalogado por expertos del sector como un arte y por otros tantos como un ritual. El viaje para obtener una deliciosa taza de café inicia con la obtención de la semilla de café, para posteriormente ser situada en su respectivo germinador; actualmente también existe la posibilidad de realizar la siembra por medio de plántulas, las cuales pueden ser adquiridas en tiendas especializadas del sector. Una vez obtenida la semilla de café o plántula, la siguiente etapa es la siembra la cual es considerada de gran complejidad por la gran cantidad de variables dentro de las cuales se contemplan posibles distribuciones y diseños que pueden ser empleados a lo largo del proceso de siembra.

Una vez sembradas las semillas de café; y previo a la cosecha se encuentra el proceso del cultivo, el cual usualmente viene incluido con la adición de plaguicidas, fertilizantes y abonos los cuales cumplen funciones específicas durante esta etapa. Continuando con

el proceso de producción y una vez finalizada la etapa de cultivo inicia el periodo de cosecha, el cual abarca la etapa de recolección, en donde entra a tomar importancia la selección del grano con el objetivo obtener solamente los frutos maduros o en palabras prácticas, los granos de café que se encuentren sanos y listos para ser recolectados. Una vez seleccionados los granos maduros, se procede con la fase de despulpado el cual coloquialmente se conoce como el proceso por medio del cual se le retira la cáscara al fruto, permitiendo extraer la semilla que después se emplea en la preparación de la bebida de café que se conoce popularmente. Seguidamente, se da inicio a la etapa de lavado y secado, procesos críticos para la determinación de la calidad de la taza de café. Inmediatamente después se da luz verde para trillarlo y finalmente tostarlo. Cada uno de estos procesos puede variar a partir de la característica del café que se le quiera dar prelación y en definitiva dependerá de las condiciones de cada caficultor y la cadena logística que se tenga mapeada.

Teniendo claro a grandes rasgos el proceso de producción de café, se hace necesario profundizar en la calificación y jerarquización de cafés de mayor renombre a nivel internacional en torno al SC. La *CQI*, organización internacional que desde 1996 trabaja alineado con diferentes entidades a nivel mundial y creando sinergias para tener una estandarización y un lenguaje globalizado en marco de la calidad de café, es en definitiva un punto de referencia, y es por ello que se adopta en el presente proyecto. Si bien la *CQI* evalúa tazas de café origen robusta y arábigo, en el presente trabajo solo se evaluará data de café arábigo, el cual es considerado a nivel mundial como el de mejor calidad por su aroma, su dulzor, su acidez y menor cantidad de cafeína que se traduce en sabores menos amargos.

En el proceso de calificación de *CQI*, se evalúan 10 factores durante la cata de la taza de café. A continuación, se resume la interpretación de los 10 factores según *Robert McKeon Aloe*, científico de datos:

Tabla 6.

Resumen de factores evaluados por profesionales de CQI durante la puntuación de un café especial

Q-score Metrics

Label	Description	Extra Notes
Aroma	Smell of dry coffee, wet coffee at crust break, and wet coffee as it steeps	
Flavor	Mid-range notes in between first impression and aftertaste	Intensity, quality, and complexity
Aftertaste	Length and quality of enjoyable flavor after the coffee is swallowed	
Acidity	Brightness (Good) Sour (Bad)	
Body	Tactile feeling between tongue and mouth	
Balance	Balance of Flavor, Aftertaste, Acidity, and Body	
Uniformity	Consistency of flavor across cups	Scored 2 points over 5 cups
Clean.Cup	Lacking negative tastes from beginning to end of taste	Scored 2 points over 5 cups
Sweetness	Sugary flavor (good) Astringency flavors (bad)	
Cupper.Points	Holistic score by Cupper	
Total	Total of the previous 10 metrics	

Nota. Tabulación de los 10 parámetros evaluados por los Calificadores de la CQI, detallando: parámetros, descripción y notas adicionales. Tomado de: *R. McKeon, "Grading the Usefulness of Q-Scores for Coffee Scores"*, 2020.

<https://towardsdatascience.com/grading-the-usefulness-of-q-scores-for-coffee-grading-8ff2e8439572>

Como se observa en la tabla 6, se evalúan factores del café tales como el aroma, el sabor (notas percibidas), prolongación del sabor (duración y calidad de sabor), acidez (viva o amarga), cuerpo (detectado tanto en el paladar como en lengua), balance (de las 5 características mencionadas anteriormente), uniformidad (consistencia de sabor a lo largo de varias copas del mismo lote), limpieza de copa (presencia de sabores negativos

desde el inicio hasta el final de la evaluación), dulzor (evitando sabores astringentes), puntuación holística y gran total [80].

Finalmente, es así como después de considerar estos 10 parámetros, que el *CQI* determina si un café es especial o no a nivel mundial, dando fe de imparcialidad, profesionalismo y objetividad, siendo en definitiva un punto de partida tanto para compradores como para vendedores al momento de comercializar el grano teniendo en consideración que dentro de los cafés especiales los de mayor puntaje, serán de una mayor calidad que aquellos que estén cercanos a una puntuación total de 80 puntos.

3.2.2. Métodos de aplicación del ML

Para obtener modelamientos precisos y acordes con las problemáticas identificadas, el *ML* desprende varias ramas de acuerdo con el tipo y volumen de datos que se utilicen para la construcción de algoritmos. Por ende, define las siguientes categorías:

- Aprendizaje supervisado
- Aprendizaje no supervisado
- Aprendizaje reforzado
- *Deep learning*

3.2.2.a. Aprendizaje supervisado. Resultado de la variabilidad de las necesidades y aplicaciones del *ML*, surge el aprendizaje supervisado *AS*, el cual tiene como objeto establecer patrones dentro de un conjunto de datos que permitan su posterior análisis. Según Judith Hurwitz y Daniel Kirsch mencionan que todo el proceso del *AS* empieza por un conjunto específico de datos, en el que se conoce la forma de clasificación de los mismos. Por ejemplo, en el mundo del internet se pueden encontrar miles de imágenes de animales y su respectiva descripción. A partir de este supuesto, es factible crear un algoritmo de *ML* que permita distinguir un animal de otro. Al definir el nombre de la especie del animal, cuando se realiza la búsqueda este puede estar categorizado en infinidad de vertientes. Debido a que se han identificado los atributos y el significado de los datos, los usuarios que están entrenando los datos modelados lo entienden bien para que se ajusten a los detalles de las etiquetas [66].

Por ser una de las aplicaciones más representativas del *ML*, se encuentra en el mundo de la ciencia y la investigación varias definiciones. El profesor *M.Gopal de IIT Delpi*, autor de más de 150 artículos científicos sobre el *ML* aborda el *AS* desde el supuesto de que la máquina se encuentra totalmente diseñada para procesar información sin modificaciones y es capaz de convertirla en valores observados del sistema, es decir, en vectores de entrada y salida para cada uno de los estados previamente definidos [21]. Adicional a esto, todo el proceso se lleva a cabo a través del conjunto de datos que se utiliza para inferir el modelo en el sistema.

Los profesores expertos en ciberseguridad y tecnología de Australia e India, autores del libro *Blockchain and Machine Learning for e-Healthcare Systems* interpretan el *AS* como la integración de variables independientes que permiten pronosticar el valor futuro de la variable objetivo definida. Por ende, surge la interacción de las variables dependientes como salida y las independientes como entrantes al algoritmo [81].

Bajo la observación de la información consolidada sobre el *AS* se concluye que este método de adaptación del *ML* se basa en la clasificación de data intacta a partir de la creación y desarrollo de un algoritmo que logre agrupar dichos datos en un determinado ítem, previamente programado.

En el marco del *AS*, existen varios tipos de vertientes, entre ellas cabe destacar: i) Clasificación y ii) Regresión, las cuales se detallan a continuación.

3.2.2.a.i. Clasificación. En consecuencia, de que las aplicaciones del *ML* varían según la naturaleza de los datos y el objetivo, se habla de clasificación cuando la data observada proviene de un conjunto finito de valores [66]. Para comprender de manera más práctica este concepto, *Michael Beyeler* en su libro *ML for OpenCV* lo define como la aplicación que se utiliza para predecir categorías a través de los datos. Un ejemplo claro es el siguiente:

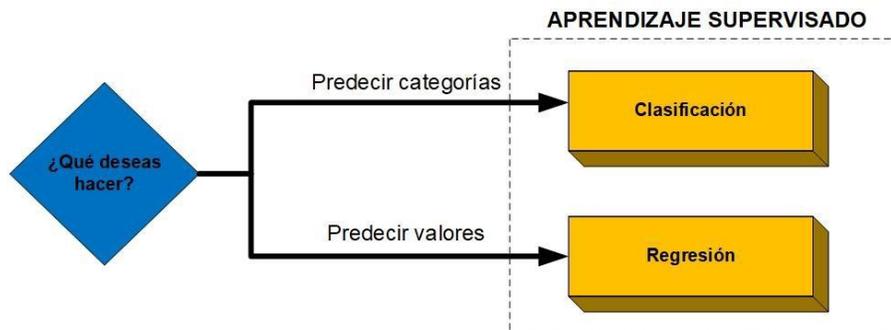
...Cuando intentamos deducir si una imagen contiene un gato o un perro. Aquí, las etiquetas de los datos son categóricas, ya sea una u otra, pero nunca una mezcla de las mismas. Por ejemplo, una imagen contiene un gato o un perro, nunca un 50% de gato y un 50% de perro, y nuestro trabajo es simplemente decir cuál es.

Debido a que no siempre se requerirá que la salida del modelo se encuentre en 2 categorías, se contemplan varias. Por esta razón, cuando los datos sólo tienen 2 opciones de clasificación, se denomina clasificación binaria o de dos clases. Sí, por el contrario, los datos pueden categorizarse en más de 2 categorías, se denomina clasificación de clases múltiples [33].

3.2.2.a.ii. Regresión. El AS va muy ligado a la estadística, es por esto que se habla de regresión en este tipo de desarrollo del ML. Ethem Alpaydin, profesor de ingeniería informática de la Universidad Ozyegin ubicada en Estambul, nombra que cada modelo obtenido a partir de una entrada de datos y se espere una salida, debe partir desde ciertos supuestos de dependencia y a su vez, ajustar los distintos parámetros con el fin de obtener predicciones precisas sobre los datos. En regresión, el desempeño depende de cuán cerca esté el modelo [26]. Para comprender un poco más el concepto de Regresión, es preciso mencionar la definición que le da Michael Beyeler, especialista en la ciencia de la computación de UC Irvine. Cuando se usa data para pronosticar valores reales se habla de un algoritmo de regresión. Un caso puntual en la economía, es donde se pretende hallar el valor exacto de las acciones. Este método permite definir un valor que se pretende alcanzar y como resultado se obtendrá el que arroje el puntaje más alto de precisión. Por lo tanto, el resultado para este caso puede ser la obtención de valor de las acciones con la menor puntuación de error posible [33].

Figura 24.

Diferenciación entre Clasificación-Regresión según el algoritmo esperado



Nota: Tipos de decisiones al implementar cualquiera de las 2 tipologías del aprendizaje supervisado. Tomado de: *M. Beyeler, ML for OpenCV. Differentiating between classification and regression problems.*

3.2.2.b. Aprendizaje no supervisado. El aprendizaje no supervisado ANS es ideal para manejar data de grandes dimensiones que no contengan etiquetas para realizar la respectiva clasificación. Esta aplicación es usada por *Facebook, Instagram*, entre otras redes sociales. Según *IBM*, para la comprensión y manejo adecuado de los datos, este método requiere de varios algoritmos que sean capaces de entender el significado de la información para poder ser clasificada, basada en patrones o grupos definidos. Después de la segmentación de la información, los valores que se encuentren sin ningún tipo de etiqueta, generan los parámetros y a su vez su respectiva clasificación. Es decir que la naturaleza de este método es definir las etiquetas de la data para que pueda ser supervisada. Con el propósito de obtener mejor entendimiento sobre la aplicación de este método, *J.Hurwitz* menciona el siguiente ejemplo:

...En salud, la recolección enormes cantidades de datos sobre una enfermedad específica pueden ayudar a los médicos a adquirir información sobre los patrones de los síntomas y relacionarlos con los resultados de los pacientes. Llevaría demasiado tiempo etiquetar todas las fuentes de datos asociadas con una enfermedad como la diabetes. Por lo tanto, un enfoque de ANS puede ser útil para hallar resultados más rápidamente que un

enfoque de AS. Este método se puede implementar como un primer paso para procesar la data por medio del AS [66].

Este concepto también fue abordado por analistas de la plataforma *EBay*, los cuales indican el *ANS* como el método que proporciona información minuciosa de los datos que no contienen etiquetas. Frecuentemente, en el manejo de data de grandes proporciones en cuanto a características y registros no se encuentra estructurada, lo que da pie a que no se obtenga una observación a simple vista de la información necesaria para posterior análisis, por lo tanto, esta técnica usualmente se emplea para agrupar los datos mediante la exploración de estructuras ocultas y también disminuir el nivel de complejidad de la data sin suprimir contenido valioso de la misma [82].

También, uno de los exponentes más importantes de todo el mundo del *ML*, *M.Gopal* lo define como el método utilizado cuando los datos de construcción del algoritmo no tienen un objetivo final o salida definida. Cuando se presenta esta inconsistente, se parte del conocimiento de las características del grupo de vectores para descifrar las similitudes subyacentes [21].

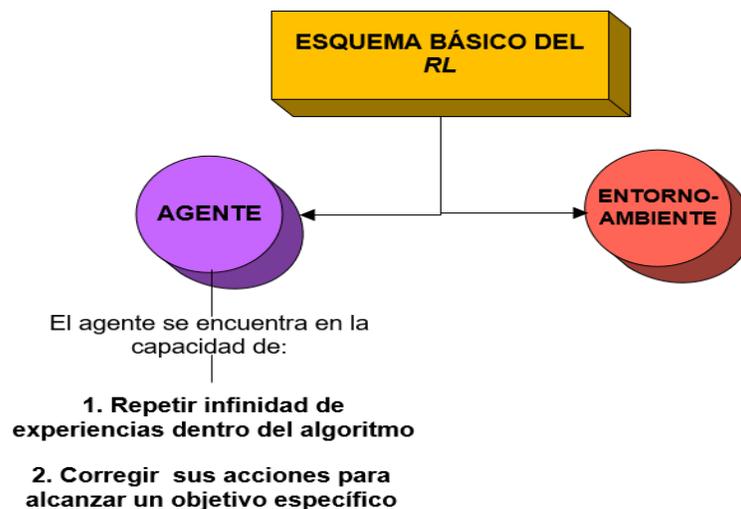
Finalmente, para complementar todo lo que abarca el *ANS*, *Ahindra Nag*, especialista en la aplicación de la tecnología en el área de la química, lo define como “El *ANS* descubre patrones o características en los datos de entrada sin la ayuda de un maestro, esencialmente realizando un agrupamiento del espacio de entrada” [83]. Se considera que la anterior definición es la que más afinidad posee con la presente investigación debido a que menciona el concepto de forma concreta y acertada.

3.2.2.c. Aprendizaje reforzado. El mundo del manejo de los datos es tan amplio que se considera necesario adoptar varias vertientes para poder manejar, analizar y acotar la información de tal forma que se obtenga éxito en los prototipos que se planteen. Por consiguiente, el *ML* aborda el aprendizaje reforzado *RL*. *Giuseppe Bonaccorso* experto en *ML*, *IA* y *Data Science*, miembro de *Institute of Electrical and Electronics Engineers IEEE* contextualiza el *RL* como un “conjunto de enfoques que permite a un agente aprender a comportarse en un entorno desconocido gracias a las recompensas que se otorgan después de cada posible acción” [84]. Adicional a esto, afirma que el *RL* es uno de los conceptos que más ha tomado fuerza en los últimos años, consecuencia de la creación

de modelos de aprendizaje profundo a través de algoritmos estándar simples que son capaces de resolver problemas bastante complejos, como, por ejemplo, aprender a jugar un juego de Atari perfectamente. Para poder emplear y entender el funcionamiento del *RL*, se mencionan los siguientes conceptos:

Figura 25.

Esquema básico de los componentes del RL



Nota. Componentes principales del aprendizaje reforzado. Tomado de: *Basic RL schema. Mastering Machine Learning Algorithms (2nd edition). 2020.*

M.Gopal, especialista en *ML*, aborda este concepto hablando de la interacción que posee el sistema con el ámbito, el cual le permite mejorar el desarrollo para la obtención de conclusiones e ideas escalares. En segunda instancia, el método arroja una alerta de refuerzo que se convierte en un excelente indicador de la capacidad de contestación del modelo [21]. Por otra parte, IBM indica que el *RL* es un prototipo de aprendizaje conductual. En primer lugar, se realiza la retroalimentación de los datos de tal forma que el algoritmo los reciba y pueda guiar al usuario directo al resultado óptimo. Cabe aclarar que el *RL* no se entrena a partir de datos de muestra porque este tipo de *ML* aprende a través de prueba y error. "Por lo tanto, una secuencia de decisiones acertadas resultará en que el proceso sea 'reforzado' porque es el más adecuado para

solucionar el problema en cuestión” [66].

Se considera muy acorde con la investigación la definición de *RL* dada por la compañía *IBM* puesto que expone la esencia y objetivo principal de este tipo de *ML* y a su vez, menciona claramente los pasos que abarca todo el proceso de desarrollo del aprendizaje reforzado.

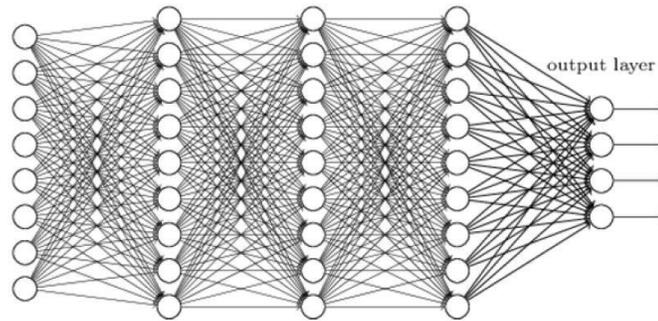
3.2.2.d. Deep Learning. En la búsqueda de herramientas que proporcionen la usabilidad de bastante data y al mismo tiempo se pueda procesar por un mismo hilo conductor que obtenga como efecto una salida única, nace el concepto de *Deep Learning DL* o también conocido como *Neural Networks NN*. Dentro de la infinidad de descripciones encontradas sobre el *DL*, se considera de mayor peso y concordancia con la presente investigación las siguientes:

En el año 2018, el especialista en economía, negocios, *Data Science* y *ML* *Matt Taddy* estructura el *DL* como la construcción de las redes neuronales como grupos de código simples. Por lo tanto, se convierte como uno de los métodos más fuertes debido a la utilización de patrones esenciales que optimizan el tiempo de cálculo y simplifican el entrenamiento. Este tiene combinaciones lineales de entradas que pasan a través de funciones de activación no lineales llamadas nodos [85].

Con fin de comprender más a fondo la filosofía del *DL*, se realiza el siguiente diagrama:

Figura 26.

Red de cinco capas adaptada por Nielsen



Nota. Red de cinco capas, adaptada de Nielsen. Tomado de: M.Taddy, “The Technological Elements of Artificial Intelligence”, NBER,2018.

Antes de explicar la anterior figura, es necesario tener claro que “un conjunto de nodos que toma diferentes sumas ponderadas de las mismas entradas se denomina 'capa', y la salida de los nodos de una capa se convierte en entrada para la siguiente capa”. Los círculos que se observan en la figura 5 son los nodos, la primera hilera de nodos, de izquierda a derecha, usualmente vienen siendo datos sin ningún tipo de modificación o data procesada previamente a través de un conjunto externo de capas. Después de todo el proceso de iteraciones y procesamiento de la información, la capa de salida *output layer* arroja las predicciones.

Visto el *DL* de forma más específica, es vital comprender este concepto desde un punto más práctico, como lo aterriza la empresa IBM definiéndolo como un método puntual del *ML* que abarca las redes neuronales a través de capas consecutivas con el fin de realizar iteraciones de los datos que permitan el aprendizaje. Esta rama del *ML* es bastante útil cuando se pretende aprender a través de data no estructurada. El *DL* está especialmente diseñado para imitar el funcionamiento del cerebro humano, para partir

del entendimiento del cerebro y poder realizar patrones que entrenen a los dispositivos tecnológicos. Para que, de esta manera, se logre la resolución de problemas que se encuentran definidos incorrectamente [66]. Este concepto es frecuentemente usado para aplicaciones que involucren el reconocimiento de imágenes, voz y visión por computadora.

Para concretar el aprendizaje profundo, *V. Konasani* y *S. Kadre* lo definen como una vertiente del *ML* que permite la entrada de data neta, es decir, sin ningún tipo de procesamiento o análisis, y a medida de las interacciones que realizan las capas sustrae información relevante, como por ejemplo las características de los datos [68] . Teniendo en cuenta la información anteriormente expuesta lo largo de este ítem, el *DL* es la iteración de varios algoritmos que contiene data netamente sin modificar y a través de los nodos obtiene un resultado preciso.

3.2.3. Cadena de suministro del sector cafetero colombiano

El SC colombiano tiene un amplio panorama en cuanto a la cadena de suministro CSya que debido a la variedad de granos de café y las diversas condiciones ambientales como la ubicación del cultivo, se requiere de diferentes recursos para mantener la calidad y el excelente estado de este. Además, es importante reconocer todos los factores que influyen en la cadena de valor de la producción del café para esclarecer las vertientes que impactan directamente a la actual investigación.

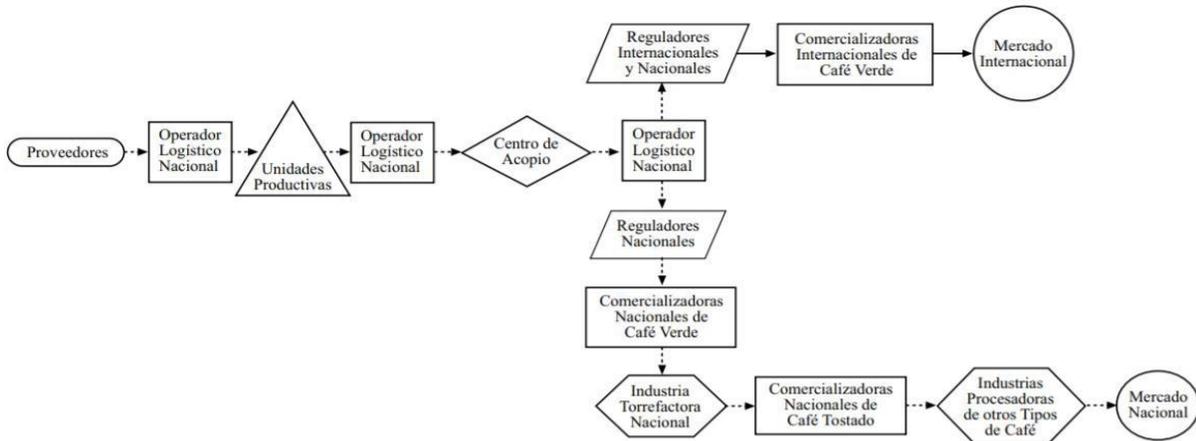
Para identificar los factores de la CS del SC colombiano, se debe conocer el sistema de producción del café, el cual incluye los siguientes conceptos: límites del sistema, entradas, componentes, interacciones y salidas. Se profundizará netamente en las entradas ya que es el pilar de la construcción del algoritmo. Las entradas dentro del sistema de la producción de café las define la entidad nacional Cenicafé como los flujos provenientes del entorno externo del sistema de producción, precisando que los más usuales para el sector agrícola son: agua, sol, insumos, asistencia técnica, créditos, mano de obra, entre otros [86]. Estos recursos necesarios para los cultivos cafeteros van estrechamente relacionados con la planeación de la cadena de suministro ya que es esta la que permitirá que estén disponibles los recursos tangibles (excepto los naturales) justo

cuando sea necesario su utilización.

Adicionalmente a lo expuesto anteriormente, Rafael García y Érika Olaya [87], magíster en Ingeniería Industrial, caracterizan la cadena de abastecimiento local de café, así:

Figura 27.

Caracterización de la cadena de abastecimiento local de café



Nota. Diagrama sobre los actores influyentes en la cadena de abastecimiento local de café en Colombia. Tomado de: R. García, E. Olaya, Caracterización de las cadenas de valor y abastecimiento del sector agroindustrial del café, 2006. <http://infocafes.com/portal/wp-content/uploads/2016/06/v19n31a08.pdf>.

En la anterior figura, se puede observar cómo la CS puede variar de acuerdo con el mercado que se dirija el café. Adicional a esto, son distintos los factores influyentes dentro de toda la cadena, teniendo cada uno su grado y responsabilidad para que todos los eslabones funcionen correctamente con la finalidad de ofrecer un producto con altos estándares de calidad y ganar territorio dentro del mercado cafetero nacional y posteriormente mundial.

Por otra parte, es fundamental recalcar el papel clave de los agricultores y personal porque son el primer eslabón de la CS y a su vez, el filtro para garantizar el bajo porcentaje en cuanto a la pérdida de desperdicios, controlar la calidad de los insumos utilizados para los cultivos, cuidado y preservación del excelente estado de las plantas

del café, entre otras vertientes que surgen dentro del proceso.

3.3. Marco Histórico

El objetivo del marco histórico es realizar un recorrido a lo largo de la evolución de la temática abarcada en el proyecto desde su surgimiento hasta la actualidad. A continuación, se detalla el proceso de desarrollo del *ML* y sus efectos en la sociedad como se conoce hoy en día.

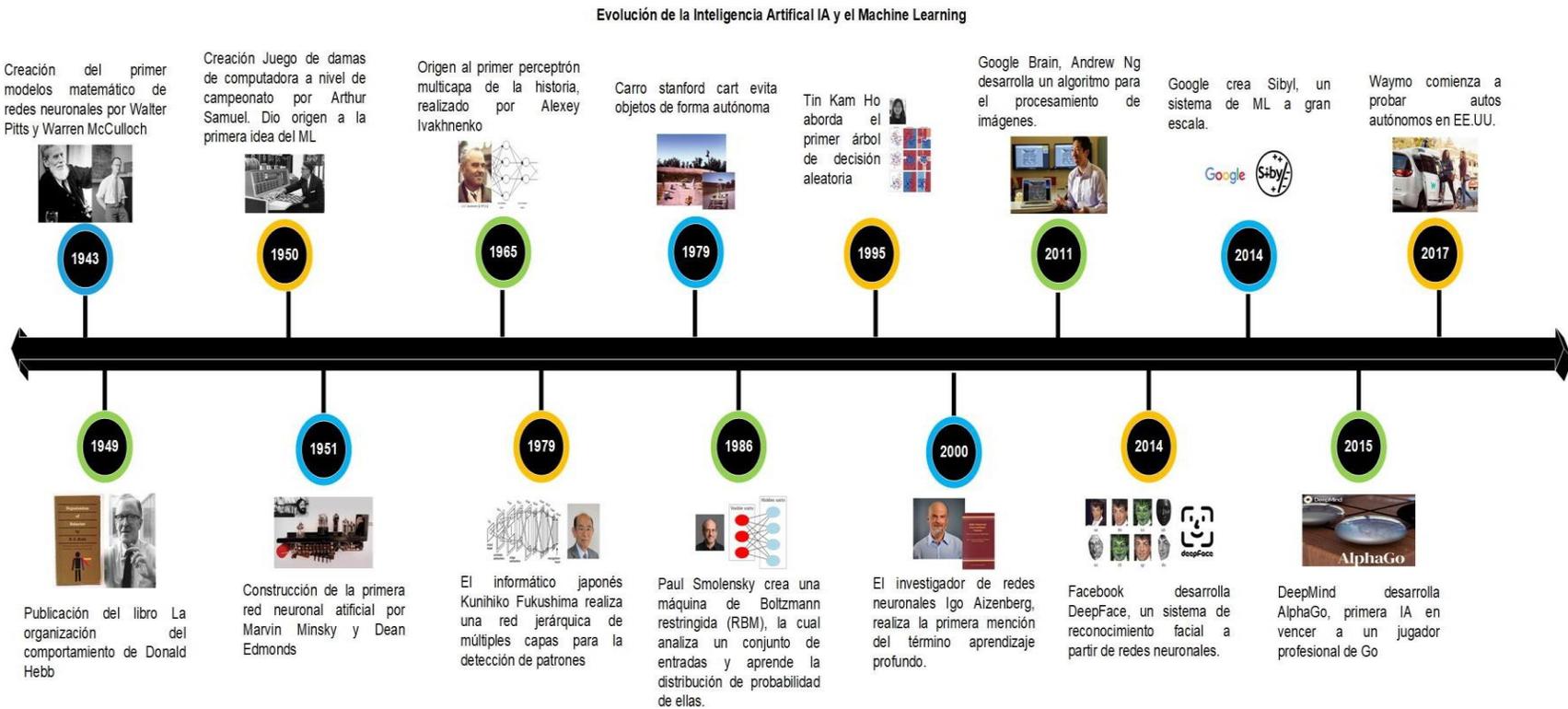
3.3.1. Evolución de la IA y el ML

Con el paso del tiempo y surgimiento de nuevas necesidades para entender el mundo de manera más clara, varios pensadores han desarrollado teorías, conceptos y métodos que conduzcan a la mejora continua, optimización de tiempo, facilidad en tareas complejas convencionales, creación de innovadores formas de realizar tareas repetitivas, entre muchas otras; todas con el propósito de hacer la vida humana un poco más austera y práctica. Aunque en este apartado no se detalla cada avance obtenido en cuanto a la innovación de la tecnología, sí se mencionan los hitos que han marcado la evolución de la IA y el ML.

Para entender el mundo tecnológico al cual la sociedad se enfrenta cada día es necesario detallar el proceso del *ML*, para ello se acude a la publicación realizada en 2020 por la compañía ucraniana *Label Your Data* [88].

Figura 28.

Hitos del ML a nivel mundial



Nota. Se describen los principales hitos del ML y se detalla sus efectos en la actualidad de esta tecnología. Tomado de: Gladchuk, V., 2021. *A history of Machine Learning from 1940s to present days.* <https://labeledyourdata.com/articles/history-of-machine-learning-how-did-it-all-start>

En la figura 28. se observa cómo el ser humano desde 1943 buscó alternativas para cambiar totalmente la percepción de lo creado y así mismo, transformar y desarrollar el pensamiento de la mente humana para plasmarlo de manera similar en máquinas automatizadas. Cabe mencionar que todo el movimiento tecnológico está estrechamente relacionado con las revoluciones industriales de la historia ya que, debido a estas, el ser humano se vio en la necesidad de regenerar todo lo creado y empezar a tomar estudios para basarse en ellos y plantear nuevos métodos.

Algunos de los hitos más representativos de la *IA* y el *ML* son: En 1951, con el surgimiento de la primera red neuronal artificial realizada por los científicos estadounidenses *Marvin Minsky* y *Dean Edmonds* , la cual constó de 40 neuronas interconectadas con memoria a corto y largo plazo. El informático japonés *Kunihiko Fukushima* a raíz de la necesidad de detectar patrones y originar redes neuronales convolucionales, crea la arquitectura de *Neocognitron* en 1979, basada en una red jerárquica de múltiples capas; este principio es usualmente empleado hasta el día de hoy para prototipos de análisis de imágenes.

Siguiendo el mismo propósito de *K. Fukushima*, de generar nuevas alternativas de aplicación de la *IA* y el *ML*, en 1995 *Tin Kam Ho* crea el primer árbol de decisión aleatoria permitiendo la creación y fusión de múltiples decisiones de *IA* para mejorar significativamente la precisión y toma de decisiones. El 2014 fue un año crucial para evidenciar la aplicabilidad de los conceptos y métodos definidos por los grandes pensadores de la *IA*, puesto que el equipo de investigación de Facebook desarrolló una red neuronal compuesta por 9 capas, detectando con una precisión del 97,35% el rostro humano comparado con imágenes. También, *Google* desarrolló un grupo de algoritmos autónomos con su propio sistema.

Todos los hitos de la *IA* y el *ML* tienen repercusiones en el mundo y la sociedad, es por esto que es fundamental contemplar, desde el impacto al ser humano tanto negativo y positivo, el desarrollo de nuevas tecnologías que realmente aporten y contribuyan al crecimiento de la sociedad en general.

3.4. Marco Normativo

Con naturalidad en el marco normativo se exhibe y detalla el conjunto de leyes, normas y reglamentos que atañen o enmarcan el tema desarrollado en el proyecto. Sin embargo, actualmente no se cuenta con una legislación oficial que reglamente y regule la IA, ni a nivel nacional o internacional, y es que se presentan gran cantidad de preocupaciones o dificultades frente a esta tecnología que está siendo implementada tanto en el sector privado como en el público, a gran y pequeña escala.

A pesar de que no se cuenta con una regulación clara con respecto al manejo y directrices en el marco del ML puntualmente, a nivel nacional se comienza a percibir matices con la intención de generarla lo antes posible. Es por ello, que a continuación se menciona lo publicado en el CONPES económico 3975, para la política nacional y la transición digital e inteligencia artificial, en donde se plasman 14 principios para el desarrollo de la IA en Colombia. Dentro de estos se resaltan dos de ellos: i) *políticas basadas en evidencias y métricas de impacto para la regulación* y ii) *Experimentación regulatoria*: En donde el primero propone que toda legislación en el marco de la IA debes sustentarse con evidencias con el objetivo de no limitar el desarrollo de la misma, sino que, por el contrario, proponer un entorno de mercado competitivo que estimule el crecimiento y el desarrollo de la IA. Por otra parte, el segundo principio mencionado propone una regulación novedosa que permita a los empresarios y usuarios de la IA experimentar con la tecnología, arriesgarse y tomar nuevos retos considerando siempre la gobernanza del estado colombiano [89].

Además de lo mencionado por el CONPES, el gobierno nacional ha presentado algunos lineamientos para construir una guía en cuanto al manejo de las nuevas tecnologías en entidades públicas. Estas nacen a raíz del surgimiento de la contingencia que ha provocado el COVID-19 en Colombia. A continuación, se evidencian algunos puntos que enmarca la norma.

Tabla 7.

Directrices y lineamientos nacionales con respecto a la IA

Fuente	Área	Tipo	Descripción	Alcance
Presidencia de la República	Derecho Informático – Comunicaciones	Directiva 3 de 2021 <i>Lineamientos para el uso de servicio en la nube, inteligencia artificial, seguridad digital y gestión de datos.</i> [76]	<ol style="list-style-type: none"> 1. Inteligencia Artificial <ol style="list-style-type: none"> 1.1. El uso de sistemas de IA deberá propender por el crecimiento inclusivo, el desarrollo sostenible y el bienestar de los ciudadanos. La IA debe mejorar la calidad de vida de los colombianos. 1.2. En caso de implementar proyectos de IA, deberán informar sobre sus avances, dentro de los respectivos informes anuales de rendición de cuentas 1.3. Formular y desarrollar proyectos de IA, dando cumplimiento a las recomendaciones y principios éticos en la materia. 1.4. Desplegar los sistemas de IA sobre la base de sus funciones, el contexto y en consonancia con el estado de la técnica. Para este fin implementarán medidas que permitan demostrar de forma activa su diligencia y cuidado en la implementación de estos sistemas. 1.5. Según sus roles, el contexto y su capacidad para actuar, aplicar un enfoque sistemático de gestión de riesgos en cada fase del ciclo de vida del sistema de IA para abordar los peligros relacionados con los sistemas de IA y su implementación, reduciendo cualquier riesgo de discriminación, entre otras posibles afectaciones. 1.6. Documentar los procesos y las decisiones adoptadas durante el ciclo de vida del sistema de IA, para permitir el análisis de sus resultados, teniendo en cuenta el contexto y siendo coherente con el estado de la técnica. 	Entidades públicas de la rama ejecutiva del orden nacional

Nota. Especificaciones de las directrices en cuanto a la regulación de la inteligencia artificial en Colombia. Tomado de: Directiva 3 de 2021, «Lineamientos para el uso de servicio en la nube, inteligencia artificial, seguridad digital y gestión de datos.,» Presidencia de la República, Colombia. <https://uamericaleyexinfo.ezproxy.uamerica.edu.co/normativa/detalle/directiva-3-de-2021-67802/txt> [90].

Se puede observar el compromiso que está adoptando el Gobierno Nacional al dictar algunas directrices que conlleven a las buenas prácticas de la utilización de IA en las entidades públicas. Aunque aún se quedan cortas en cuanto al alcance, es importante resaltar que Colombia ya está contemplando lineamientos para marcar el buen camino de las nuevas tecnologías con el fin de impactar positivamente la sociedad.

Bajo la carencia de una regulación formal en el marco de la IA, se procede a mencionar algunas de las principales razones por las cuales no se han regulado esta tecnología que sin duda alguna cobra más y más fuerza con el pasar de los días. Según el *OpenMind* del banco BBVA, uno de los principales factores que ha impedido la regulación de la IA es su rápido ritmo de cambio, sumado a que hoy en día se puede legislar a un ritmo determinado y siguiendo un debido proceso el cual suele ser largo y con múltiples fases o etapas. Es por esto que se expone la oportunidad de que las empresas o sectores económicos que explotan la IA, tomen responsabilidad de los efectos y consecuencias de las acciones ejecutadas por sus mecanismos de IA, ya que solo así estas organizaciones se verán obligadas a innovar en la transparencia de sus procesos o mecanismos en el marco de la IA [91].

Si bien el tema de responsabilidad es delicado, este se vuelve aún más complejo cuando se agregan variables de suma importancia, por ejemplo, los agentes de la cadena de valor de la IA: creador, programador, propietario y usuario. ¿Quién debería asumir la responsabilidad en caso de un evento no planeado? preguntas como estas son las que deberán ser resueltas por los altos tribunales sorteando soluciones para cada situación en específico. Un ejemplo claro es el expuesto por el diario Cinco Días, quien expone su preocupación frente a la IA en el ámbito comercial puntualmente; en donde por ejemplo, por medio de la IA una máquina logra aprender una práctica o ruta para obtener mayores beneficios económicos acudiendo a prácticas no éticas o anticompetitivas; esto en la medida de que, en el caso de la IA de aprendizaje por refuerzo, la máquina aprende a partir de estímulos, los cuales conllevan a la percepción de mayores beneficios económicos. Situaciones como esta se pueden presentar y deben ser reguladas tanto a nivel nacional como a nivel internacional. La capacidad de generar optimizaciones, datos,

y nuevas rutas de negocio por medio de la IA es real, en algunas ocasiones se puede presentar el escenario de que una creación sea desarrollada íntegramente por una máquina. En esta situación ¿Quién debe realizar la protección de los derechos de autor? Este tipo de patentes vale la pena dedicarle mayor detalle y claridad en términos legales y jurídicos y como ya se comentó anteriormente no se cuenta con la claridad de cómo proceder en dichos escenarios [92].

Ya que se han mencionado varias razones por las que no se cuenta con una regulación de la IA bien estructurada; a nivel internacional en el año 2020, el Comité Económico y Social Europeo CESE realizó una serie de observaciones y recomendaciones sobre el ideal de regulación de la IA desde la excelencia y confianza. Algunos puntos que expone el CESE son: la importancia del análisis y diferentes perspectivas que debe tener la IA para mantener un enfoque sociotécnico constante y no se sesgue a un solo tipo evaluativo, con el fin de brindar un estudio amplio de los riesgos elevados que puede originar la IA. También, la entidad habla sobre la óptica limitada que tiene la IA al sólo basarse en datos, originando limitaciones para posibles desarrollos en el futuro. Por lo cual, es importante promover e incentivar el surgimiento de nuevos sistemas de IA que tengan en cuenta el conocimiento y razonamiento, pero al mismo tiempo sean defensores de los valores y principios humanos. Contemplando la situación actual debido a la pandemia, el CESE expone que “los métodos de la IA y los enfoques empleados en la lucha contra la contingencia actual debido al COVID-19 deben ser sólidos, eficaces, transparentes y explicables. Al mismo tiempo, deben velar por los derechos humanos, los principios éticos y la legislación existente, además de ser justos, inclusivos y voluntarios” [93].

Finalmente se concluye que existe un gran vacío en lo que se refiere al marco legal y normativo de cara a la revolución 4.0 y en especial en el marco de la IA, la cual cada vez toma más posicionamiento en la economía como se conoce hoy en día. En definitiva, los gobiernos se encuentran en una carrera contrarreloj y está en las manos de cada país ir un paso adelante o esperar al tradicional ensayo y error.

4. RESULTADOS

Desde el año 2019 hasta el día de hoy, se observaron distintos estudios con utilización del *ML* en solucionar problemáticas en el sector agricultor y caficultor global. Demostrando así la gran influencia de las nuevas tecnologías en panoramas realmente inciertos para los agricultores y que, a partir de estas herramientas pueden encaminar y visualizar claramente las soluciones a los inconvenientes que se enfrenta en el día a día cada sector.

Con respecto al modelo SCOR, específicamente en el diagrama de hilos se pudo identificar cómo los proveedores de los proveedores principales pueden afectar crucialmente la calidad de las cerezas de café. Por ejemplo, las semillas de café con la certificación de la variedad son indispensables para los viveros que alimentan los cultivos. Por otro lado, se tiene el servicio de extensión rural el cual se convierte en el centro de construcción del cultivo y finalmente, los nutrientes minerales ya que son los que preservan y potencializan la productividad de los cafetos, además mantienen aquellas condiciones que la tierra por sí misma no le puede brindar al cultivo.

A lo largo de la exploración de información varios autores afirmaban la relación de la calidad del café con la altura a la cual se encuentra el cultivo. Esta premisa fue analizada y, para los registros observados dentro de la base de datos estudiada en *Python*, no existe correlación entre la calificación total y la altura en m. Además, en el mapa de calor reafirma que influye en menos del 20%.

Se evaluó la probabilidad de obtener cierta calificación total vs. Las propiedades sensoriales como el dulzor y sabor detectado en una taza de café, por medio de la construcción de diagramas de violines. Los cuales, permitieron visualizar la variabilidad que puede incidir en la calificación final del café y así observar un panorama más aterrizado en cuanto a puntuaciones específicas. El más interesante por su comportamiento fue el del sabor, pues si se obtiene un puntaje de 6,67 o 8,17 puntos la probabilidad de obtener una puntuación mayor a 80 puntos en la evaluación final es igual. Mientras que en el dulzor si se obtiene un puntaje de 6, la probabilidad de puntaje final

puede oscilar entre los 60 a 75 puntos y si, por el contrario, resulta de 10 la calificación final puede estar entre 70 a 90 puntos.

En definitiva, los algoritmos del *ML* probados (regresión lineal y de crestas) bajo la base de datos disponible cumplen totalmente con las expectativas esperadas, ya que su porcentaje de confiabilidad superó el 77% en los dos casos y es un excelente punto de partida para empezar a incorporar este tipo de *IA* en los procesos productivos del sector cafetero colombiano que encaminen sin duda alguna al aumento de la productividad del sector bajo supuestos confiables.

5. CONCLUSIONES

Los algoritmos de *ML* que se encuentran en furor para el aprendizaje de programas computacionales a través de data son el *K-means*, máquina de vectores de soporte, redes neuronales artificiales y bosques aleatorios. En más del 60% de los estudios abordados, se evidencia la combinación de diversas herramientas para complementar eficientemente los resultados esperados. Adicionalmente, los estudios de caso comúnmente abordados en la industria agrícola son la detección y control temprana de propagación de enfermedades y plagas, y composición de propiedades físico – químicas de los alimentos. En la industria cafetera se aborda el control de cultivos en cuanto a la categorización de las etapas de maduración del grano, dispersión de fertilizantes y la calidad a través de la presencia de proteínas en cada fruto observado. Finalmente, las herramientas del *ML* en todos los casos arrojaron un porcentaje de precisión mayor al 90%.

A través de la exploración de información por medio de entidades de peso en el sector cafetero y lo analizado a lo largo del presente proyecto, se establecen 6 aspectos indispensables para la conservación de la calidad del café colombiano: primordial la selección de la semilla, almácigos o plántulas de café que cuenten con certificación de la originalidad de la variedad (Castillo – arábica) emitida por la *FNC*, fertilización y renovación de los cafetos bajo el previo análisis del suelo, cumplimiento de las buenas prácticas enfocadas a las condiciones de almacenamiento de los bultos de café, adquisición de secadoras eficientes y finalmente, ubicación del cultivo que cuente con una excelente disponibilidad de agua y luz natural.

El modelo *SCOR* se acopló exitosamente al estudio porque permitió identificar aquellos puntos clave de la *CS* total que engloba a la caficultura. Aunque el análisis del *SCOR* es bastante extenso, en el presente proyecto se enfocó en los proveedores ya que son considerados el eje principal de la y son aquellos influyentes en la obtención de cultivos eficientes y en la generación de frutos de café con alta calidad, lo cual es el propósito del presente estudio.

Todo proveedor que intervenga de manera directa e indirecta con la estructuración y cuidado del cultivo, fortalecimiento de los cafetos y manipulación del grano es de vital importancia que el caficultor cuente con requisitos establecidos para realizar la selección y evaluación de los mismos para tener un control mayor ante las adversidades o acciones que pueda cometer el proveedor, y estas repercutan en el mal sabor y prolongación del mismo.

No se optó por escoger ninguno de los algoritmos de *ML* en auge debido que algunos hacen parte del *Deep Learning* y el aprendizaje no supervisado, por lo cual estos tipos no se ajustaban a la data en estudio porque cada dato tenía su respectiva etiqueta, es decir, se tenía certeza de que cada puntaje correspondía a una propiedad sensorial en específico. Y, por otro lado, su alcance es demasiado amplio para el estudio. Por tales razones, se modeló el algoritmo de regresión lineal, perteneciente a la categoría de aprendizaje supervisado, el cual trabaja data con etiquetas definidas previamente.

En definitiva, los algoritmos del *ML* son aplicables a casos puntuales de la *LA* del SC colombiano siempre y cuando se identifique el problema, se analice qué información hay disponible para correr el modelo y se puntualice en la solución tentativa que puede arrojar el algoritmo para impactar positivamente la *LA* de cada caficultor. Teniendo claro el objetivo y lo anterior, el *ML* es supremamente versátil y adaptable a cualquier problemática que se pretenda abordar.

BIBLIOGRAFÍA

- [1] ICONTEC, «Buenas prácticas de manufactura de la industria del café,» ICONTEC, Norma Técnica Colombiana, 2003. [En línea]. Available: <https://vdocuments.es/ntc-5181.html>. [Último acceso: 10 Septiembre 2021].
- [2] Comité Nacional de Cafeteros, «Resolución 02 de 2016. Por la cual se unifican y actualizan las normas de calidad del café verde en almendra para exportación,» Federación Nacional de Cafeteros de Colombia, 16 abril 2016. [En línea]. Available: <https://federaciondecafeteros.org/static/files/RESOL04.pdf>. [Último acceso: 22 Agosto 2021].
- [3] «Informe del Gerente al 88 Congreso Nacional de Cafeteros Virtual,» Flipsnack, 2021. [En línea]. Available: <https://www.flipsnack.com/federaciondecafeteros/informe-del-gerente-al-88-congreso-nacional-de-cafeteros-virtual.html>. [Último acceso: 08 Marzo 2021].
- [4] P.Yoshida, «Índice de competitividad regional cafetero,» Federación de cafeteros, 2021. [En línea]. Available: https://federaciondecafeteros.org/static/files/Indice_competitividad_cafetero.pdf. [Último acceso: 11 Marzo 2021].
- [5] I. Admin, «Los algoritmos al servicio del agro - Conexión Intal,» Conexión intal, 2021. [En línea]. Available: <https://conexionintal.iadb.org/2018/07/27/ideas1-2/>. [Último acceso: 26 Mayo 2021].
- [6] «TIC para el agro - SAC - Sociedad de Agricultores de Colombia,» SAC – Sociedad de Agricultores de Colombia, 2021. [En línea]. Available: <https://sac.org.co/tic-para-el-agro/>. [Último acceso: 26 Mayo 2021].
- [7] «La innovación llega al sector cafetero,» Siglodata.co, 2021. [En línea]. Available: <http://www.siglodata.co/s/observatorio/20181123LATINPYMEINNPULSA.pdf>. [Último acceso: 12 Marzo 2021].

- [8] O. Cerquera, «El acompañamiento institucional en el desarrollo del sector cafetero colombiano,» Universidad Surcolombiana, Scielo, 2014. [En línea]. Available: <http://www.scielo.org.co/pdf/fype/v7n1/v7n1a09.pdf> . [Último acceso: 11 Marzo 2021].
- [9] «FNC se reúne con cafeteros de Antioquia, Caldas, Quindío y Risaralda para analizar efectos de covid-19 y retos del sector - Federación Nacional de Cafeteros,» Federación Nacional de Cafeteros, 2021. [En línea]. Available: <https://federaciondefcafeteros.org/wp/listado-noticias/fnc-se-reune-con-cafeteros-de-antioquia-caldas-quindio-y-risaralda-para-analizar-efectos-de-covid-19-y-retos-del-sector/>. [Último acceso: 12 Marzo 2021].
- [10] H. Ali, «Digital transformation, development and productivity in developing countries: is artificial intelligence a curse or a blessing? » Emerald.com, 2021. [En línea]. Available: <https://www.emerald.com/insight/content/doi/10.1108/REPS-11-2019-0145/full/pdf?title=digital-transformation-development-and-productivity-in-developing-countries-is-artificial-intelligence-a-curse-or-a-blessing>. [Último acceso: 13 Marzo 2021].
- [11] H. Nguyen a. F. Mwesigye, «Coffee value chain analysis Opportunities for youth employment in Uganda., » 2021, [En línea]. Available: <http://www.fao.org/3/cb0413en/cb0413en.pdf>. [Último acceso: 13 Marzo 2021].
- [12] W. Eustasio a. J. Dioses, «Classification of Immature and Mature Coffee Beans Using RGB Values and Machine Learning Algorithms, » Warse, 2021. [En línea]. Available: <http://www.warse.org/IJETER/static/pdf/file/ijeter22872020.pdf>. [Último acceso: 22 Marzo 2021].
- [13] W.Gomes de Costa. et al., «Machine learning and statistics to quality environments through multi-traits in Coffea arabica,» Wroclaw University of Science and Technology, 2021. [En línea]. Available: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0245298>.

[Último acceso: 22 Marzo 2021].

- [14] Organisation for Economic Co-operation and Development, «Modelling Adaptation to Climate Change in Agriculture, » OECD, 2021. [En línea]. Available: https://www.oecd-ilibrary.org/agriculture-and-food/modelling-adaptation-to-climate-change-in-agriculture_5jxrclljnxbxq-en. [Último acceso: 16 Agosto 2021].
- [15] Kashyap, Ramgopal Kumar, A. V. Senthil. «Challenges and Applications for Implementing Machine Learning in Computer Vision - 8.6 Unsupervised Learning,» IGI Global, 2020. [En línea]. Available: <https://app.knovel.com/hotlink/pdf/id:kt01263FR2/challenges-applications/artificial-unsupervised>. [Último acceso: 16 Agosto 2021].
- [16] J. Gao. e. al., «Pixel-level aflatoxin detecting in maize based on feature selection and hyperspectral imaging», » Science Direct, 2021. [En línea]. Available: <https://www.sciencedirect.com/science/article/abs/pii/S138614252030247X?via%3Dihub>. [Último acceso: 16 Agosto 2021].
- [17] H. Jung Min. e. al., «Development of a smartphone-based lateral-flow imaging system using machine-learning classifiers for detection of Salmonella spp, » Science Direct, 2021. [En línea]. Available: <https://www.sciencedirect.com/science/article/abs/pii/S0167701221001561?via%3Dihub>. [Último acceso: 17 Agosto 2021].
- [18] P. Benavides a. L. Constantino, «El barrenador del tallo y la raíz del café, Plagiohammus colombiensis,» Cenicafé, 2015. [En línea]. Available: <https://www.cenicafe.org/es/publications/2.Barrenador.pdf>. [Último acceso: 16 Agosto 2021].
- [19] L. de Oliveira Aparecido. e. al., «Machine learning algorithms for forecasting the incidence of Coffea arabica pests and diseases, » Springer link, 2020. [En línea]. Available: <https://doi.org/10.1007/s00484-019-01856-1>. [Último acceso: 18 Agosto 2021].

- [20] Yi Peng. e. al., "Characterization of Brazilian coffee based on isotope ratio mass spectrometry ($\delta^{13}\text{C}$, $\delta^{18}\text{O}$, $\delta^2\text{H}$, and $\delta^{15}\text{N}$) and supervised chemometrics," Food Chemistry Anhui Agricultural University, 2019. [Online]. Available: https://app.dimensions.ai/details/publication/pub.1116998595?search_mode=content&search_text=KNN%20AND%20coffee&search_type=kws&search_field=full_search. [Accessed 17 Agosto 2021].
- [21] M. Gopal, Applied Machine Learning, McGraw-Hill, 2019.
- [22] W. Z. e. al, «Applications of Support Vector Machine in Genomic Prediction in Pig and Maize Populations, » Dimensions, 2021. [En línea]. Available: https://app.dimensions.ai/details/publication/pub.1133073755?search_mode=content&search_text=SVM%20and%20agriculture&search_type=kws&search_field=full_search. [Último acceso: 17 Agosto 2021].
- [23] L. Huang. e. al, «Potential of Visible and Near-Infrared Hyperspectral Imaging for Detection of Diaphania pyloalis Larvae and Damage on Mulberry Leaves, » Dimensions, 2018. [En línea]. Available: https://app.dimensions.ai/details/publication/pub.1105192856?search_mode=content&search_text=SVM%20and%20agriculture&search_type=kws&search_field=full_search. [Último acceso: 18 Agosto 2021].
- [24] M. Baudelet, «Laser Spectroscopy for Sensing - Fundamentals, Techniques and Applications - 12.1 Introduction: The Historical 'Terahertz Gap, » Elsevier, 2014. [En línea]. Available: <https://app.knovel.com/hotlink/pdf/id:kt00U8D57J/laser-spectroscopy-sensing/introduction-historical>. [Último acceso: 17 Agosto 2021].
- [25] S. Y. e. al, «Determination of the Geographical Origin of Coffee Beans Using Terahertz Spectroscopy Combined With Machine Learning Methods,» Frontiers, 2021. [En línea]. Available: <https://www.frontiersin.org/articles/10.3389/fnut.2021.680627/full>. [Último acceso: 19 Agosto 2021].
- [26] E. Alpaydin, ML: The new IA, Cambridge, MA: The MIT Press, 2016.

- [27] M. T. e. al, «Crop mapping from image time series: Deep learning with multi-scale label hierarchies, » Science Direct, 2021. [En línea]. Available: <https://www-sciencedirect-com.ezproxy.uamerica.edu.co/science/article/abs/pii/S0168169921001149?via%3Dihub>. [Último acceso: 18 Agosto 2021].
- [28] Keras Team, « Keras documentation: Adagrad,» Keras.io, 2021. [En línea]. Available: <https://keras.io/api/optimizers/adagrad/>. [Último acceso: 18 Agosto 2021].
- [29] S. Saedi, M.Rezaei, a. H. Khosravi, «Real-time recognition of on-branch olive ripening stages by a deep convolutional neural network,» Science Direct, 2021. [En línea]. Available: <https://www-scopus-com.ezproxy.uamerica.edu.co/record/display.uri?eid=2-s2.0-85106990377&origin=resultslist&sort=plff&src=s&sid=6504ac3392ce42352167a31cfb2cdf21&sot=b&sdt=b&sl=45&s=TITLE-ABSKEY%28neural+network+AND+agriculture%29&relpos=17&citeCnt=0&search>. [Último acceso: 18 Agosto 2021].
- [30] M. de Oliveira, A. dos Santos, E. Kazama, G. de SouzaRolim and R. Pereira daSilva, «Determination of application volume for coffee plantations using artificial neural networks and remote sensing,» Science Direct, 2021. [En línea]. Available: <https://www-sciencedirect-com.ezproxy.uamerica.edu.co/science/article/abs/pii/S0168169921001149?via%3Dihub>. [Último acceso: 18 Agosto 2021].
- [31] J. T. Ballinger and G. J. Shugar, «Chemical Technicians' Ready Reference Handbook,» Access Engineering Library, 2011. [En línea]. Available: <https://www-accessengineeringlibrary-com.ezproxy.uamerica.edu.co/content/book/9780071745925/chapter/chapter27#/ch27table8>. [Último acceso: 18 Agosto 2021].
- [32] W. Aguirre, «Espectroscopia infrarrojo y técnicas de machine learning y deep learning para la detección y clasificación de arándanos,» Journal.upao.edu.pe, 2019. [En línea]. Available:

- <http://journal.upao.edu.pe/PuebloContinente/article/view/1363/1171>. [Último acceso: 18 Agosto 2021].
- [33] M. Beyeler, *ML for OpenCV*, Birmingham, UK: Packt Publishing, 2027, p. 45.
- [34] A. Dantas de Medeiros, L. Junio da Silva and J. Oliveira Ribeiro, «Machine Learning for Seed Quality Classification: An Advanced Approach Using Merger Data from FT-NIR Spectroscopy and X-ray Imaging,» *Dimensions*, 2020. [En línea]. Available: https://app.dimensions.ai/details/publication/pub.1129864031?search_mode=content&search_text=NIR%20AND%20machine%20learning%20AND%20agriculture&search_type=kws&search_field=full_search. [Último acceso: 18 Agosto 2021].
- [35] M. Zhu et al, «Fast determination of lipid and protein content in green coffee beans from different origins using NIR spectroscopy and chemometrics,» *Dimensions*, 2021. [En línea]. Available: https://app.dimensions.ai/details/publication/pub.1139065470?search_mode=content&search_text=coffee%20AND%20k-means&search_type=kws&search_field=full_search&or_facet_year=2019&or_facet_year=2020&or_facet_year=2021. [Último acceso: 18 Agosto 2021].
- [36] International Labour Organization ILO, «Sustainable supply chains to build back better,» ILO, 2020. [En línea]. Available: https://www.ilo.org/wcmsp5/groups/public/---ed_dialogue/---sector/documents/genericdocument/wcms_777245.pdf. [Último acceso: 4 Septiembre 2021].
- [37] Organización internacional del café, , «Efectos de la COVID-19 en el sector mundial del café: Encuesta de los miembros exportadores de la OIC,» *Serie coffee break N° 3*, Junio 2020. [En línea]. Available: <https://www.ico.org/documents/cy2019-20/coffee-break-series-3c.pdf> . [Último acceso: 5 Septiembre 2021].
- [38] Semana, «Estos son los países que más exportan café,» *Revista Semana*, 2019.

- [39] Federación Nacional de Cafeteros, «Servicio de Extensión rural - Federación Nacional de Cafeteros,» Federación Nacional de Cafeteros.
- [40] «Cosecha asistida de café,» Canal FNC Youtube.com, 2016.
- [41] L. Quintero, «Productos químicos y agroquímicos de los cultivos de café,» Federación Nacional de Cafeteros Youtube.com, 2020.
- [42] S. Sadeghian, «Nuevos Grados de Fertilizante,» Tips del Profesor Yarumo Canal FNC Youtube.com, 2020.
- [43] J. Chalaca, «Tecnologías de la información y comunicación en la Cadena productiva de cafés especiales,» Universidad de Nariño, Trabajo de Grado pregrado, 2021. [En línea]. Available: <http://sired.udenar.edu.co/5971/1/TECNOLOG%20C3%8DAS%20DE%20LA%20INFORMACI%20Y%20COMUNICACI%20EN%20LA%20CADENA%20PR.pdf>. [Último acceso: 10 Septiembre 2021].
- [44] «DECRETO 410 DE 1971,» Suin-juriscal.gov.co, 2021. [En línea]. Available: http://www.suin-juriscal.gov.co/viewDocument.asp?id=1833376#ver_1833529. [Último acceso: 10 Septiembre 2021].
- [45] «Empaque de café,» Federación Nacional de Cafeteros, 2021. [En línea]. Available: <https://federaciondefcafeteros.org/wp/?search-type=all&s=empaque>. [Último acceso: 10 Septiembre 2021].
- [46] J. Ramos, «Herramientas para la Comercialización del Café, Tips del Profesor Yarumo,» Youtube.com, 2019. [En línea]. Available: <https://www.youtube.com/watch?v=lwNEdO-WICY>. [Último acceso: 10 Septiembre].
- [47] Guerra, G, «Empaque Del Café Tostado: Tipos de Bolsas Para Elegir,» *Perfect Daily Grind español.*, 2021.

- [48] «Café de Colombia, “Proceso de elaboración del Café en Colombia”,» FNC, youtube.com, 2019. [En línea]. Available: <https://www.youtube.com/watch?v=Q7o49APAwHM>. [Último acceso: 10 Septiembre 2021].
- [49] Cardona, A, « Almacafé cuenta con el mayor complejo logístico e industrial de café a nivel mundial en Soacha.,» Agronegocios.co., 2018. [En línea]. Available: <https://www.agronegocios.co/agricultura/almacafe-cuenta-con-el-mayor-complejo-logistico-e-industrial-de-cafe-a-nivel-mundial-en-soacha-2738666>. [Último acceso: 18 Septiembre 2021].
- [50] Alma café operador logístico, «Gestión del Transporte,» Canal de Youtube Alma café, 2019. [En línea]. Available: <https://www.youtube.com/watch?v=t5CB5gUyqgo>. [Último acceso: 18 Septiembre 2021].
- [51] «Chatarrización y peajes, dos problemas del transporte de carga,» *El universal*, 2018.
- [52] Staff, F., «Infraestructura: las nuevas vías y proyectos claves de la reactivación,» *Forbes*, 2021.
- [53] «Información institucional global sobre las funciones principales de los comités a nivel Colombia,» Comité de cafeteros de Antioquia, [En línea]. Available: <https://fncantioquia.org/infografia-institucional>. [Último acceso: 06 Octubre 2021].
- [54] FNC, « Real Academia del Café – FNC. Curso express: Conociendo el mundo delcafé colombiano,» Realacademiadelcafe.com, 2021. [En línea]. Available: <https://realacademiadelcafe.com/>. [Último acceso: 07 Octubre 2021].
- [55] YAO, T., « Brew methods effect on coffee flavor and aroma,» Oaktrust.library.tamu.edu. , 2017. [En línea]. Available: <https://oaktrust.library.tamu.edu/bitstream/handle/1969.1/174893/YAO-THESIS-2017.pdf?sequence=1&isAllowed=y>. [Último acceso: 08 Octubre 2021].

- [56] Debela Bote, A., «Tree management and environmental conditions affect coffee (Coffea arabica L.) bean quality,» Taylor & Francis. , 2021. [En línea]. Available: <https://www.tandfonline.com/doi/full/10.1016/j.njas.2017.09.002>. [Último acceso: 08 Octubre 2021].
- [57] Luna González, A., Macías, A., Taboada, O. and Morales, V., , «Cup quality attributes of Catimors as affected by size and shape of coffee bean (Coffea arabica L.),» Taylor & Francis, 2019. [En línea]. Available: <https://www.tandfonline.com/doi/full/10.1080/10942912.2019.1603997>. [Último acceso: 10 Octubre 2021].
- [58] Tilahun Getachew, A. and SooChuna, B , «Coffee Flavor.,» Science Direct, 2021.[En línea]. Available: <https://www-sciencedirect-com.ezproxy.uamerica.edu.co/science/article/pii/B9780081005965216582>. [Último acceso: 10 Octubre 2021].
- [59] University Cambridge, «Quality and Reliability in Engineering. Cambridge,» 2009. [En línea]. Available: https://assets.cambridge.org/97805215/15221/excerpt/9780521515221_excerpt.pdf.
- [60] H.Bhasin, « Algorithms - Design and Analysis,» Oxford University Press, 2015. [En línea]. Available: <https://app.knovel.com/hotlink/toc/id:kpADA00011/algorithms-design-analysis/algorithms-design-analysis>. [Último acceso: 21 Abril 2021].
- [61] N. Safiya Umoja, «Algorithms of Oppression: How Search Engines Reinforce Racism,» New York: NYU Pres, 2018. [En línea]. Available: https://preview.aer.io/Algorithms_of_Oppression-Mjc1OTQ3?social=1&retail=1&emailcap=0.
- [62] E. Sestari and A. Marconatto, «Vista de Algorithm and Language,» Revistas.ucr.ac.cr, 2019. [En línea]. Available: <https://revistas.ucr.ac.cr/index.php/filosofia/article/view/43158/43422>. [Último acceso: 13 Abril 2021].

- [63] Y. Sugomori, B. Kaluza, F. Soares and A. Souza, Deep learning: Practical Neural Networks with Java, Birmingham: Packt Publishing, 2015.
- [64] G. Bekda, S. Melih Nigdeli and M. Yucel, Artificial Intelligence and Machine Learning Applications in Civil, Mechanical, and Industrial Engineering, 1 ed., IGI Global, p. 1.
- [65] E. Soria Olivas, J. Martin Guerrero, M. Martinez Sober, J. Magdalena Benedito and A. Serrano López, Research on Machine Learning Applications and Trends: Algorithms, Methods and Techniques, IGI Global, 2019.
- [66] Machine Learning for Dummies, IBM, 2021.
- [67] S. Caballero and J. Brice, Artificial Intelligence/ML + Supply Chain Planning, MIT Center for Transportation & Logistics, 2021.
- [68] V. Reddy Konasani and S. Kadre , Machine Learning and Deep Learning Using Python and TensorFlow, McGraw Hill, 2021.
- [69] Surviving Supply Chain Integration: Strategies for Small Manufacturers, The National Academics of Sciences Engineering Medicine, 2000.
- [70] Ballou, Ronald, Business Logistics management, Prentice Hall, 2004.
- [71] Pricewaterhousecoopers, Manual Práctico de Logística.
- [72] A. Calatayud and R. Katz, CADENA DE SUMINISTRO 4.0. Mejores Prácticas Internacionales y Hoja de Ruta para América Latina, Banco Interamericano de Desarrollo BID, 2019.
- [73] S. Brand, «Supply Chain Management and Logistics: What's the Difference?» Cmtc.com, 2018. [En línea]. Available: <https://www.cmtc.com/blog/supply-chain-management-and-logistics-whats-the-difference>. [Último acceso: 11 Abril 2021].
- [74] «Inbound Logistics,» Cips.org, 2021. [En línea]. Available: <https://www.cips.org/knowledge/procurement-topics-and-skills/logistics-management/logistics1/>. . [Último acceso: 11 Abril 2021].

- [75] «CSCMP Supply Chain Management Definitions and Glossary,» Council of Supply Chain Management Professionals CSCMP, 2021. [En línea]. Available: https://cscmp.org/CSCMP/Educate/SCM_Definitions_and_Glossary_of_Terms.aspx. [Último acceso: 11 Abril 2021].
- [76] H.Xuefang, W.Ezra, *Inbound Logistics Optimization*, Massachusetts: Massachusetts Institute of Technology, 2020.
- [77] E. Felsing and P. Runza, «Productividad: Un Estudio de Caso en un Departamento de Siniestros. Maestría,» Universidad del CEMA, 2002. [En línea]. Available: https://ucema.edu.ar/posgrado-download/tesinas2002/Felsinger_MADE.pdf. [Último acceso: 21 Abril 2021].
- [78] World Bank Group, «Productivity and innovation in the czech republic: a firm-level perspective, » European Commission, 2019. [En línea]. Available: <https://openknowledge.worldbank.org/bitstream/handle/10986/33084/Background-Note-Productivity-and-Innovation-in-the-Czech-Republic-A-Firm-Level-Perspective.pdf?sequence=1&isAllowed=y> . [Último acceso: 14 Abril 2021].
- [79] J.Medina, «Modelo integral de productividad, aspectos importantes para su implementación,» *Revista EAN*, 2010.
- [80] R. McKeon, «Grading the Usefulness of Q-Scores for Coffee Scores,» 2020. [En línea]. Available: <https://towardsdatascience.com/grading-the-usefulness-of-q-scores-for-coffee-grading-8ff2e8439572>. [Último acceso: 21 Abril 2021].
- [81] B. Balusamy, N. Chilamkurti and B. Lucia Agnes, *Blockchain and Machine Learning for e-Healthcare Systems*, Institution of Engineering and Technology.
- [82] A. Isoni and D. Cervellin , *Machine Learning for the Web*, Packt Publishing,, 2016.
- [83] A. Nag, *Biosystems Engineering*, The McGraw-Hill Companies, Inc, 2010.
- [84] G. Bonaccorso, *Mastering Machine Learning Algorithms*, Packt Publishing, 2020.
- [85] M.Taddy, «The Technological Elements of Artificial Intelligence,» National Bureau of Economic Research NBER, Feb 2018. [En línea]. Available:

- <https://www.nber.org/papers/w24301>. [Último acceso: 11 Abril 2021].
- [86] Centro Nacional de Investigaciones de Café, «Manejo Agronómico de los Sistemas de Producción de Café,» Cenicafé, 2020. [En línea]. Available: <https://doi.org/10.38141/cenbook-0002>. [Último acceso: 18 Abril 2021].
- [87] R. García, E. Olaya, «Caracterización de las cadenas de valor y abastecimiento del sector agroindustrial del café,» 2006. [En línea]. Available: <http://infocafes.com/portal/wp-content/uploads/2016/06/v19n31a08.pdf>. [Último acceso: 21 Abril 2021].
- [88] Gladchuk, V, «A history of Machine Learning from 1940s to present days,» Labeledyourdata.com, 2020. [En línea]. Available: <https://labeledyourdata.com/articles/history-of-machine-learning-how-did-it-all-start/>. [Último acceso: 21 Abril 2021].
- [89] Documento conpes económico 3975, «Política Nacional para la Transformación Digital e Inteligencia Artificial,» Consejo Nacional de Política Económica y Social –CONPES, [En línea]. Available: <https://uamerica-leyex-info.ezproxy.uamerica.edu.co/normativa/detalle/documento-conpes-economico-politica-nacional-para-la-transformacion-digital-e-inteligencia-arti-51678/pdf>. 2019. Pág. 22.
- [90] Directiva 3 de 2021, «Lineamientos para el uso de servicio en la nube, inteligencia artificial, seguridad digital y gestión de datos.,» Presidencia de la República, Colombia., 15 Marzo 2021. [En línea]. Available: <https://uamerica-leyex-info.ezproxy.uamerica.edu.co/normativa/detalle/directiva-3-de-2021-67802/txt> . [Último acceso: 28 Abril 2021].
- [91] OpenMind BBVA., «La última década y el futuro del impacto de la IA en la sociedad,» 2019. [En línea]. Available: <https://www.bbvaopenmind.com/articulos/la-ultima-decada-y-el-futuro-del-impacto-de-la-ia-en-la-sociedad/>.

- [92] cincodias.elpais., 2020. [En línea]. Available: https://cincodias.elpais.com/cincodias/2020/02/06/legal/1580972637_539834.html.
- [93] Comité Económico y Social Europeo, «Dictamen del Comité Económico y Social Europeo sobre el “Libro blanco sobre la inteligencia artificial: un enfoque europeo orientado a la excelencia y la confianza”, de 16 julio 2020,» CESE, Unión europea, 16 Julio 2020. [En línea]. Available: <https://uamerica-leyex-info.ezproxy.uamerica.edu.co/normativa/detalle/dictamen-del-comite-economico-y-social-europeo-sobre-el-libro-blanco-sobre-la-inteligencia-artif-64623/pdf>. [Último acceso: 28 Abril 2021].
- [94] S. Asmar, «La geografía y el clima, claves para la producción con buen sabor del café especial,» Agronegocios, 2020. [En línea]. Available: <https://www.agronegocios.co/agricultura/la-geografia-y-el-clima-claves-para-la-produccion-con-buen-sabor-del-cafe-especial-3044031>
- [95] ICONTEC, «Buenas prácticas de manufactura de la industria del café,» 2003. [En línea]. Available: <https://vdocuments.es/ntc-5181.html>.
- [96] Fundación UTZ, «UTZ certified,» CERES Colombia, 2015. [En línea]. Available: https://cerescolombia-cert.com/wp-content/uploads/2020/06/3.2.39_es_Breve-Info-UTZ_Inf_15-12-01.pdf. [Último acceso: 15 Octubre 2021].
- [97] Rainforest Alliance, «Estándar de Agricultura Sostenible rainforest alliance. Requisitos de la cadena de suministro,» CERES Colombia, 2020. [En línea]. Available: <https://cerescolombia-cert.com/wp-content/uploads/2021/10/Estandar-de-Agricultura-Sostenible-2020-Requisitos-de-la-cadena-de-suministro.pdf>. [Último acceso: Octubre 14 2021].
- [98] ISEALL ALLIANCE, «Certificación 4C.,» CERES Colombia, 2021. [En línea]. Available: https://cerescolombia-cert.com/wp-content/uploads/2021/07/210218_4C-Unit-Certification_Step-by-Step_ES.pdf.

[Último acceso: 14 Octubre 2021].

- [99] V. Nadya Regina Galo. e. al, «Selección de proveedor de servicios logísticos: alineación entre criterios e indicadores,» Facultad de Ciencias Económicas. Universidad Nacional de Colombia., 2018. [En línea]. Available: <https://www.redalyc.org/journal/818/81856556005/html/>. [Último acceso: 14 Octubre 2021].

