

Optimización sostenible de la cadena de suministro alimentaria: revisión bibliográfica de aplicaciones de machine learning en logística y reducción de desperdicios

Jhonny Enrique Amaya Rodriguez

Asesora

Edith Johana Morales Liberato

Universidad Nacional Abierta y a Distancia UNAD

Escuela de Ciencias Básicas Tecnología e Ingeniería

Especialización en Ciencia de Datos y Analítica

2025

Resumen

La presente investigación aborda el impacto del *machine learning* en la optimización de la cadena de suministro sostenible en la industria alimentaria a partir de una revisión exhaustiva de referencias bibliográficas. Dada la creciente demanda de productos sostenibles y las exigencias normativas internacionales, las organizaciones del sector alimentario enfrentan el reto de transformar sus procesos logísticos y operativos hacia modelos más eficientes y responsables. A través del análisis de contenido y la revisión de estudios de caso, se identificaron los beneficios del uso de algoritmos de aprendizaje automático en la predicción de la demanda, la gestión de inventarios y la reducción de desperdicios. En la revisión bibliográfica realizada, se identificó que los modelos de inteligencia artificial más recurrentemente aplicados en la optimización sostenible de cadenas de suministro alimentarias fueron los algoritmos de Random Forest y las redes neuronales recurrentes (RNN). Diversos estudios (Amado Mateus, 2024; Dhiman et al., 2024; Ochoa-Barragán et al., 2024) reportaron que estos modelos alcanzaron precisiones superiores al 90 % en la predicción de la demanda, lo cual permitió una optimización significativa en el uso de recursos y en la planificación de la distribución. Además, se observaron mejoras concretas como un incremento del 12 % en eficiencia energética, una reducción del 18 % en los tiempos de entrega, y un nivel de trazabilidad superior al 95 %, gracias a la implementación de sistemas inteligentes integrados a lo largo de la cadena logística (Juárez, 2024; Hong y Xiao, 2024; Rodríguez, 2022). La investigación concluye que el *machine learning* no solo mejora el desempeño operativo, sino que representa una herramienta clave para avanzar hacia una industria alimentaria más sostenible, innovadora y competitiva.

Palabras clave: Machine learning, sostenibilidad, cadena de suministro, Big Data y trazabilidad.

Abstract

The present research addresses the impact of machine learning on the optimization of sustainable supply chains in the food industry through an exhaustive review of bibliographic references. Given the growing demand for sustainable products and international regulatory requirements, organizations in the food sector face the challenge of transforming their logistical and operational processes toward more efficient and responsible models. Through content analysis and the review of case studies, the benefits of using machine learning algorithms in demand forecasting, inventory management, and waste reduction were identified. The literature review revealed that the artificial intelligence models most frequently applied in the sustainable optimization of food supply chains were Random Forest algorithms and recurrent neural networks (RNN). Several studies (Amado Mateus, 2024; Dhiman et al., 2024; Ochoa-Barragán et al., 2024) reported that these models achieved accuracies greater than 90% in demand prediction, which enabled significant optimization in resource use and distribution planning. Furthermore, concrete improvements were observed, such as a 12% increase in energy efficiency, an 18% reduction in delivery times, and a traceability level above 95%, thanks to the implementation of intelligent systems integrated throughout the logistics chain (Juárez, 2024; Hong y Xiao, 2024; Rodríguez, 2022). The theoretical framework developed in this study offers a replicable model that articulates sustainability with regulatory compliance and digital transformation, which can be applied by organizations of different sizes. The research concludes that machine learning not only improves operational performance, but also represents a key tool for advancing towards a more sustainable, innovative, and competitive food industry.

Keywords: Machine learning, sustainability, supply chain, Big Data, and traceability.

Tabla de Contenido

Introducción	8
Planteamiento del Problema	10
Justificación	14
Objetivos	16
Objetivo General.....	16
Objetivos Específicos.....	16
Marco Conceptual	17
Machine Learning Aplicados a Cadenas de Suministro Sostenibles	17
Machine Learning (ML)	17
Aprendizaje Supervisado	19
Regresión: Lineal, Múltiple, Logística	20
Clasificación: Árboles de Decisión, Random Forest, SVM, ANN	20
Ensemble: XGBoost, Gradient Boosting	22
Aprendizaje No Supervisado	22
Clustering: K-Means, Clustering Jerárquico, DBSCAN	23
Reducción de Dimensionalidad: PCA, t-SNE	25
Asociación: A priori	26
Aprendizaje por Refuerzo	27
Q-Learning	27
Deep Q-Networks (DQN)	28
Policy Gradients	29
Aplicaciones de Aprendizaje Supervisado: Predicción de Precios, Detección de Spam.....	30

	5
Limitaciones	35
Aplicaciones de Aprendizaje no Supervisado: Segmentación, Anomalías, Recomendación, Reducción	35
Ventajas	36
Limitaciones	36
Aplicaciones Aprendizaje por Refuerzo: Robots, Rutas, Juegos	37
Ventajas	38
Limitaciones	38
Metodología	40
Criterios de Selección y Alcance de la Revisión Bibliográfica.....	40
Impacto Económico, Social y Ambiental del <i>Machine Learning</i> en la Industria Alimentaria..	42
Marco Teórico Aplicado: Machine Learning y Sostenibilidad en la Gestión de la Cadena de Suministro Alimentaria	51
Tecnología: Aplicaciones de Machine Learning Como Herramienta Analítica y Predictiva ...	52
Gestión de la Cadena de Suministro: Procesos Logísticos, Trazabilidad, Inventario, Transporte y Calidad.....	53
Sostenibilidad: Enfoque Integral Desde lo Económico, Social y Ambiental	54
Conclusiones	56
Recomendaciones	58
Referencias Bibliográficas	60

Lista de Tablas

Tabla 1 <i>Casos de Uso en Aprendizaje Supervisado</i>	34
Tabla 2 <i>Comparativa de Tipos de Aprendizaje en Machine Learning</i>	39

Lista de Figuras

Figura 1 <i>Estructura básica del ML</i>	18
Figura 2 <i>Resumen de Técnicas y Aplicaciones en Aprendizaje Supervisado</i>	19
Figura 3 <i>Tipos de Regresión y Ejemplos de Aplicación en Cadenas de Suministro</i>	21
Figura 4. <i>Agrupación en el aprendizaje automático</i>	23
Figura 5 <i>Clustering Jerárquico</i>	24
Figura 6. <i>Frecuencia de Uso de Algoritmos de Clustering en Cadenas de Suministro</i>	25
Figura 7 <i>PCA, t-SNE</i>	26
Figura 8 <i>Q-Learning Y Deep Q-Learning</i>	28

Introducción

Las cadenas de suministro en la industria alimentaria desempeñan un papel fundamental en la conexión entre productores, procesadores, distribuidores y consumidores. Su eficiencia y sostenibilidad influyen directamente en la seguridad alimentaria, la calidad del producto final y el impacto ambiental de todo el sistema agroalimentario. Sin embargo, estos sistemas complejos enfrentan numerosos desafíos relacionados con la gestión de recursos, la reducción de desperdicios, la trazabilidad de productos y la adaptación a normativas ambientales cada vez más estrictas. En este contexto, las nuevas tecnologías, en particular el *machine learning*, surgen como herramientas clave para transformar la manera en que se gestionan las cadenas de suministro.

El *machine learning*, como rama de la inteligencia artificial, permite analizar grandes volúmenes de datos históricos y en tiempo real, identificar patrones complejos y generar modelos predictivos que apoyen la toma de decisiones estratégicas. Estas capacidades ofrecen oportunidades significativas para optimizar la cadena de suministro alimentaria, desde la predicción de la demanda y la gestión de inventarios hasta la distribución eficiente y la reducción del desperdicio de alimentos perecederos. En consecuencia, su aplicación no solo mejora la eficiencia operativa, sino que también contribuye a la sostenibilidad ambiental y al cumplimiento de normativas internacionales, como los Objetivos de Desarrollo Sostenible (ODS) y el Pacto Verde Europeo.

A pesar de su potencial, la adopción de tecnologías basadas en *machine learning* en el sector alimentario sigue siendo limitada, especialmente en empresas pequeñas y medianas que enfrentan barreras tecnológicas, económicas y de capacitación. Esta brecha tecnológica limita la capacidad del sector para adaptarse a los cambios del mercado, reducir su impacto ambiental y

aumentar su competitividad global. De ahí la necesidad de investigar cómo el *machine learning* puede integrarse de forma efectiva en la gestión logística de la cadena de suministro alimentaria.

Esta monografía tiene como propósito central realizar una revisión bibliográfica sistemática que permita describir y analizar cómo las técnicas de *machine learning* han sido aplicadas en la optimización de procesos logísticos dentro de las cadenas de suministro alimentarias, con énfasis en la reducción sostenible del desperdicio de recursos y el mejoramiento de la eficiencia operativa. A través del estudio de literatura especializada, el análisis de casos documentados y la propuesta de un marco teórico aplicado, se busca aportar una comprensión integral que favorezca la transformación digital y sostenible del sector alimentario. Esta investigación, de carácter teórico y analítico, pretende generar conocimiento útil y transferible, orientado a ofrecer soluciones prácticas frente a los desafíos actuales en términos de sostenibilidad, eficiencia logística y cumplimiento regulatorio.

Planteamiento del Problema

La creciente demanda de productos sostenibles ha transformado la industria alimentaria, obligando a las empresas a optimizar sus cadenas de suministro para cumplir con normativas ambientales y responder a las expectativas del mercado. Sin embargo, la falta de integración de tecnologías avanzadas en los procesos de producción, distribución y logística ha dificultado la transición hacia modelos de suministro sostenibles (Rodríguez, 2018). Esta investigación analiza cómo un modelo basado en machine learning y análisis de datos puede mejorar la eficiencia y sostenibilidad en la cadena de suministro alimentaria, mediante la optimización del uso de recursos, la reducción de desperdicios y el fortalecimiento de la toma de decisiones estratégicas.

Las cadenas de suministro en la industria alimentaria enfrentan múltiples desafíos relacionados con la eficiencia operativa, el impacto ambiental y la competitividad en mercados globalizados. La falta de herramientas predictivas que permitan anticipar fluctuaciones en la demanda, optimizar rutas logísticas y reducir desperdicios representa un obstáculo significativo para la sostenibilidad del sector (Meléndez Mogollón et al., 2022). En este contexto, el machine learning ofrece soluciones innovadoras al permitir el análisis de grandes volúmenes de datos, la identificación de patrones de consumo y la optimización de procesos productivos (Al Zaabi et al., 2020). A pesar de su potencial, la adopción de esta tecnología sigue siendo limitada en la industria alimentaria, lo que impide una gestión eficiente de los recursos y una planificación estratégica más sostenible.

Uno de los principales problemas en la cadena de suministro alimentaria es la falta de trazabilidad y monitoreo en tiempo real, lo que dificulta la detección temprana de ineficiencias y genera pérdidas económicas y ambientales (Juárez, 2024). La ausencia de sistemas avanzados de análisis de datos restringe la capacidad de las empresas para adaptarse a cambios en la oferta y

demanda, aumentando los costos operativos y reduciendo la eficacia en la distribución de productos (González, 2019). El machine learning permite optimizar estos procesos mediante modelos predictivos que anticipan problemas logísticos, optimizan la distribución de productos y reducen el desperdicio de alimentos, mejorando así la eficiencia general de la cadena de suministro (Gómez y Alberto, 2019).

En términos normativos, la industria alimentaria está cada vez más regulada en materia de sostenibilidad. Normas internacionales como la Directiva de Diligencia Debida en Sostenibilidad Corporativa de la Unión Europea, el Pacto Verde Europeo y los Objetivos de Desarrollo Sostenible (ODS) de la ONU exigen a las empresas la reducción de su impacto ambiental y la adopción de prácticas de producción responsables (Quintero-Campos, 2019). No obstante, el cumplimiento de estas regulaciones sigue siendo un desafío, ya que muchas empresas carecen de herramientas adecuadas para evaluar su desempeño ambiental y generar estrategias efectivas de sostenibilidad. La implementación de machine learning en la gestión de la cadena de suministro podría facilitar la adaptación a estas normativas al proporcionar análisis automatizados y datos precisos sobre la eficiencia energética, la reducción de emisiones y el uso de recursos (Morcillo-Bellido, 2018).

Los estudios analizados en esta revisión bibliográfica coinciden en que la no adopción de tecnologías como el machine learning en la industria alimentaria genera consecuencias significativas en términos de costos operativos, sostenibilidad ambiental y competitividad en el mercado. Diversas fuentes documentan que la falta de herramientas predictivas y de optimización en la cadena de suministro conlleva un incremento en el desperdicio de alimentos, ineficiencias logísticas y una mayor huella de carbono, factores que pueden derivar en sanciones regulatorias y pérdida de reputación frente a consumidores cada vez más exigentes (Melamed-

Varela et al., 2018). Además, según Sánchez Jiménez y García Hernández (2019), esta carencia tecnológica limita la capacidad de anticipación y adaptación a las dinámicas del mercado, lo que reduce la rentabilidad a largo plazo. En este contexto, la presente monografía busca sistematizar evidencia empírica y teórica que permita comprender cómo las técnicas de *machine learning* pueden responder de forma efectiva a estos desafíos estructurales dentro de las cadenas de suministro alimentarias.

Por otro lado, la integración de técnicas de *machine learning* en la gestión de las cadenas de suministro alimentarias no solo aporta mejoras operativas, sino que también genera impactos significativos en los ámbitos económico, social y ambiental. Su capacidad para procesar grandes volúmenes de datos en tiempo real permite una estimación más precisa de la demanda, una mejor asignación de recursos y la reducción de costos logísticos (Coronado-Hernández y García-Sabater, 2019), lo que incide directamente en la rentabilidad y sostenibilidad económica de las empresas. Desde una perspectiva social, la optimización de la distribución de productos perecederos contribuye a disminuir el desperdicio alimentario, favoreciendo una mayor disponibilidad de alimentos y reduciendo las pérdidas en sectores vulnerables (Fernández y Bajo, 2018). En cuanto al impacto ambiental, la automatización de decisiones basada en datos facilita el cumplimiento de regulaciones ambientales, la reducción de emisiones asociadas al transporte y el uso eficiente de los recursos, promoviendo una cadena de suministro con menor huella ecológica (Rivera y Caballero, 2020). Así, el *machine learning* se presenta como una herramienta estratégica para avanzar hacia una transformación integral del sector alimentario, enmarcada en los principios de sostenibilidad.

No obstante, los beneficios tecnológicos no se limitan al plano operativo. La incorporación de herramientas basadas en *machine learning* representa una transformación

estructural que va más allá de la eficiencia puntual. Implica la adopción de un enfoque de gestión centrado en datos, que exige reconfigurar los procesos internos y fomentar una cultura organizacional orientada a la toma de decisiones fundamentadas en evidencia cuantificable (Ochoa et al, 2024). Esta transición permite a las empresas anticiparse a escenarios complejos, adaptarse con mayor agilidad a los cambios del entorno y consolidar una ventaja competitiva sostenible. En el sector alimentario, donde la variabilidad de la demanda y la perecibilidad de los productos plantean desafíos constantes, esta capacidad de adaptación se convierte en un factor decisivo para garantizar tanto la viabilidad económica como el compromiso con la sostenibilidad ambiental.

En este contexto, la falta de adopción de tecnologías emergentes como el *machine learning* en las cadenas de suministro alimentarias representa no solo una barrera para la eficiencia logística, sino también una amenaza para la sostenibilidad del sector en sus dimensiones económica, social y ambiental.

Si bien existen múltiples estudios que abordan aplicaciones específicas de estas técnicas, aún persiste la necesidad de consolidar y analizar de forma sistemática el conocimiento disponible. Por ello, esta monografía propone una revisión bibliográfica que permita identificar los enfoques más relevantes, los modelos más utilizados y los impactos documentados en la gestión de cadenas de suministro sostenibles. A partir de esta necesidad, surge la siguiente pregunta de investigación que orienta el desarrollo del presente estudio:

¿Cómo contribuyen las técnicas de *machine learning* a la optimización sostenible de las cadenas de suministro alimentario, en términos económicos, sociales y ambientales, según la evidencia bibliográfica reciente?

Justificación

El crecimiento acelerado de la industria alimentaria, junto con la presión social y regulatoria por adoptar prácticas sostenibles, ha impulsado la necesidad de transformar las cadenas de suministro hacia modelos más eficientes, trazables y ambientalmente responsables. Esta transformación requiere, entre otros factores, el uso de tecnologías avanzadas como el machine learning, el análisis de Big Data y el Internet de las Cosas (IoT), que permiten anticipar la demanda, reducir el desperdicio, optimizar rutas logísticas y garantizar el cumplimiento de normativas ambientales. Sin embargo, aunque se han documentado múltiples casos de aplicación de estas tecnologías, la información se encuentra dispersa y poco sistematizada, dificultando una comprensión integral de su verdadero impacto económico, social y ambiental en el contexto alimentario.

Pese a los avances tecnológicos, existen vacíos importantes en la literatura sobre cómo estas herramientas contribuyen específicamente a la sostenibilidad en la industria alimentaria. Muchos estudios se centran en mejoras operativas o eficiencia logística, sin abordar de forma explícita su incidencia en el cumplimiento de regulaciones, la mitigación del impacto ambiental o la generación de valor social. Además, la ausencia de un marco teórico consolidado sobre la integración de ciencia de datos en cadenas de suministro sostenibles limita el desarrollo de estrategias replicables y adaptables a distintas realidades empresariales y geográficas.

En paralelo, las exigencias normativas internacionales, como la Agenda 2030, el Pacto Verde Europeo o las leyes de Responsabilidad Extendida del Productor, demandan que las empresas implementen sistemas que les permitan medir y mejorar continuamente su desempeño ambiental. Sin embargo, muchas organizaciones aún enfrentan dificultades para incorporar herramientas analíticas que faciliten este proceso, lo que pone en riesgo su competitividad,

acceso a mercados y sostenibilidad financiera. La falta de trazabilidad y de capacidad predictiva también compromete la seguridad alimentaria y reduce la confianza de los consumidores.

Frente a este escenario, esta monografía se propone abordar estos vacíos mediante una revisión bibliográfica sistemática, que permita identificar, clasificar y analizar las aplicaciones más relevantes del machine learning en la optimización sostenible de las cadenas de suministro alimentarias. El propósito es comprender en qué medida estas tecnologías han demostrado contribuir a la sostenibilidad en sus tres dimensiones económica, social y ambiental, y qué modelos, herramientas y enfoques han sido más eficaces y replicables. Esta revisión busca, además, ofrecer una base conceptual robusta que permita orientar futuras investigaciones o implementaciones tecnológicas en el sector.

Objetivos

Objetivo General

Explorar el impacto del *machine learning* en la mejora de las cadenas de suministro sostenibles en la industria alimentaria, considerando su influencia en la eficiencia logística, la reducción del desperdicio y la predicción de la demanda, así como su aporte a la sostenibilidad económica, social y ambiental.

Objetivos Específicos

Identificar los principales enfoques de machine learning aplicados en las cadenas de suministro alimentarias sostenibles.

Analizar los efectos del machine learning en la sostenibilidad económica, ambiental y social del sector alimentario.

Proponer un marco teórico que relacione el uso del machine learning con estrategias sostenibles en la gestión de la cadena de suministro.

Marco Conceptual

Machine Learning Aplicados a Cadenas de Suministro Sostenibles

Machine Learning (ML)

En el ámbito de la inteligencia artificial, *Machine Learning* se define como la disciplina que permite a los sistemas informáticos aprender de forma autónoma a partir de los datos, mejorando progresivamente su desempeño en tareas específicas sin necesidad de ser programados explícitamente para cada una de ellas (Acuña, 2019). Esta capacidad de aprendizaje surge del análisis estadístico y la detección de patrones en grandes volúmenes de datos, lo que convierte al *Machine Learning* en una herramienta fundamental en contextos donde la programación tradicional sería ineficiente o inviable (Russell y Norvig, 2021). Desde una perspectiva funcional, su objetivo principal es construir algoritmos que generalicen comportamientos a partir de experiencias pasadas, optimizando así procesos complejos y dinámicos.

El funcionamiento de *Machine Learning* parte de una estructura en la cual un modelo recibe datos de entrada (también denominados atributos o variables independientes) y aprende a asociarlos con salidas esperadas (resultados o etiquetas), ajustando sus parámetros internos para minimizar el error en las predicciones (Badghish, 2024). Este proceso, que se conoce como entrenamiento, permite al sistema aprender una función de mapeo que será aplicada posteriormente a nuevos datos no vistos. Una vez entrenado, el modelo puede realizar tareas como clasificación, regresión, agrupamiento, detección de anomalías o predicción temporal, dependiendo del tipo de algoritmo y del objetivo del análisis (Alpaydin, 2020).

Actualmente, *Machine Learning* se divide en tres grandes categorías según la naturaleza del aprendizaje: supervisado, no supervisado y por refuerzo. El aprendizaje supervisado utiliza

datos etiquetados para entrenar modelos que predicen resultados específicos, como ocurre en la detección de fraude bancario o en el diagnóstico de enfermedades (Bains, 2024). El aprendizaje no supervisado, en contraste, se orienta a descubrir patrones o agrupaciones dentro de datos no etiquetados, como en la segmentación de clientes en marketing. Finalmente, el aprendizaje por refuerzo se basa en la interacción de un agente con su entorno, donde las acciones se ajustan mediante recompensas o penalizaciones para alcanzar un objetivo a largo plazo, como sucede en la robótica o la conducción autónoma (Sarker, 2021).

Figura 1

Estructura básica del ML



Nota. Hastie (2017)

El crecimiento del *Machine Learning* se ha visto impulsado por el acceso masivo a datos digitales y por el incremento en la capacidad computacional. Este fenómeno ha permitido su implementación en sectores estratégicos como la salud, la industria, las finanzas y el comercio. Por ejemplo, en medicina se emplea para analizar imágenes diagnósticas mediante redes neuronales convolucionales, y en logística para predecir la demanda de productos con modelos de series temporales (Goodfellow, Bengio y Courville, 2016). A su vez, las plataformas de

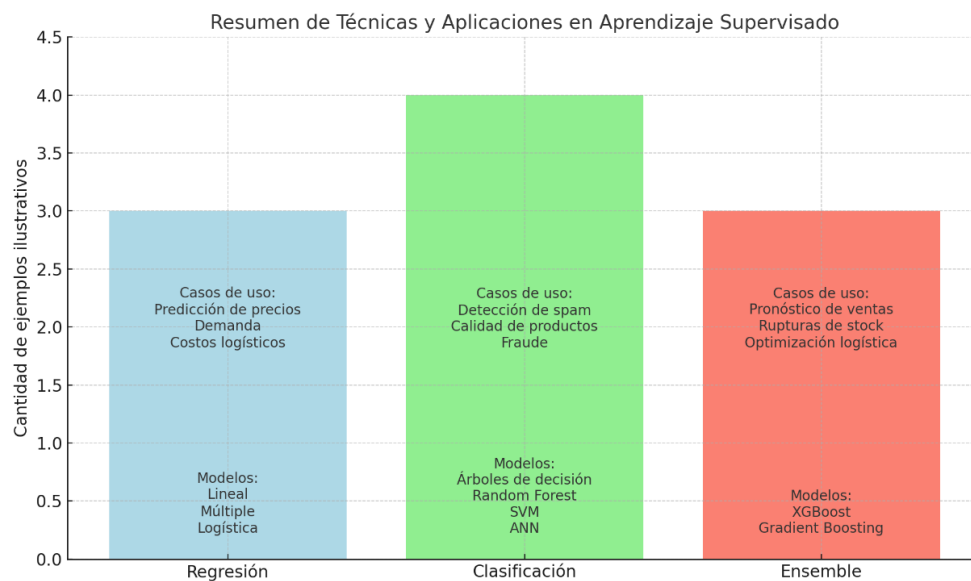
recomendación, como las utilizadas por Netflix o Amazon, se sustentan en modelos de *Machine Learning* que analizan comportamientos de los usuarios para personalizar contenidos.

Aprendizaje Supervisado

En el contexto del *machine learning*, el aprendizaje supervisado representa una de las metodologías más consolidadas y extendidas. Esta técnica consiste en entrenar algoritmos a partir de conjuntos de datos que incluyen tanto las variables independientes como las etiquetas o resultados esperados. A través de este enfoque, el modelo aprende a asociar patrones en los datos con determinadas salidas, permitiendo realizar predicciones sobre nuevos datos no etiquetados (Alpaydin, 2020). Esta relación entre entrada y salida convierte al aprendizaje supervisado en una herramienta esencial para tareas de clasificación, regresión y estimación.

Figura 2

Resumen de Técnicas y Aplicaciones en Aprendizaje Supervisado



Regresión: Lineal, Múltiple, Logística

Uno de los principales campos de aplicación del aprendizaje supervisado es la regresión, cuyo objetivo es predecir valores continuos. La regresión lineal, como forma básica, establece una relación proporcional entre una variable dependiente y una independiente. En contextos más complejos, la regresión múltiple incorpora múltiples variables predictoras para mejorar la precisión de las estimaciones. Por su parte, la regresión logística se orienta a la predicción de variables categóricas binarias, como la probabilidad de ocurrencia de un evento (James et al., 2021). En cadenas de suministro, estos modelos se aplican para prever la demanda, optimizar el inventario y calcular costos operativos.

Estos métodos se caracterizan por su interpretabilidad y eficiencia computacional. La regresión lineal, por ejemplo, permite visualizar el efecto marginal de cada predictor sobre la variable respuesta, lo que resulta útil en entornos donde se requieren explicaciones claras para la toma de decisiones (Hastie, Tibshirani y Friedman, 2017). No obstante, su rendimiento disminuye cuando la relación entre variables no es lineal o cuando existe multicolinealidad entre los predictores.

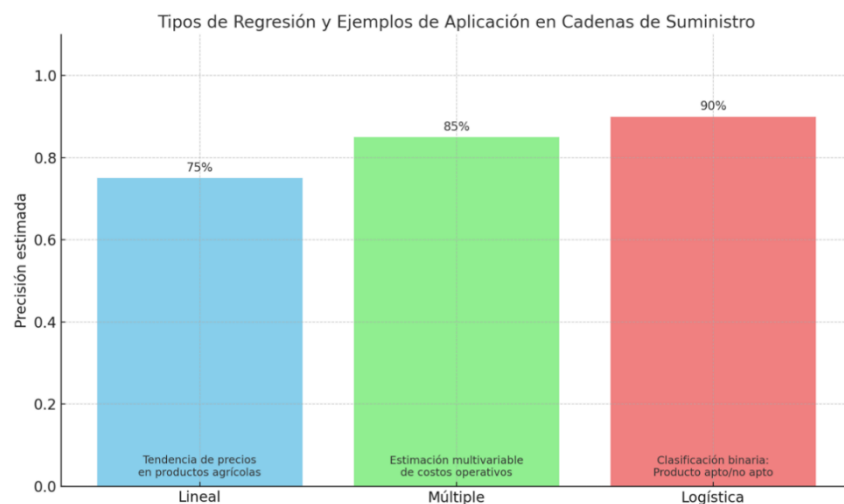
Clasificación: Árboles de Decisión, Random Forest, SVM, ANN

Dentro del aprendizaje supervisado, la clasificación se encarga de asignar etiquetas discretas a nuevas observaciones, basándose en patrones aprendidos durante la fase de entrenamiento. Uno de los métodos más utilizados es el árbol de decisión, que divide el espacio de atributos en regiones homogéneas mediante reglas jerárquicas. Este enfoque es fácilmente interpretable y útil para análisis exploratorios. Sin embargo, es susceptible al sobreajuste, lo que ha motivado el desarrollo de modelos más robustos como el bosque aleatorio (*Random Forest*), que combina múltiples árboles para mejorar la generalización (Breiman, 2001).

Otro algoritmo relevante es la máquina de soporte vectorial (*Support Vector Machine*, SVM), que construye hiperplanos óptimos para separar las clases en espacios de alta dimensión. Este modelo ha demostrado eficacia en conjuntos de datos pequeños con márgenes claros entre categorías. A su vez, las redes neuronales artificiales (*Artificial Neural Networks*, ANN) permiten abordar problemas no lineales complejos mediante la integración de múltiples capas ocultas, lo que ha sido potenciado por la disponibilidad de grandes volúmenes de datos y poder computacional (Goodfellow, Bengio y Courville, 2016). En el ámbito de la industria alimentaria, los modelos de clasificación son útiles para el control de calidad de productos, la identificación de fallos logísticos y la detección de comportamientos atípicos en el consumo. Su implementación puede facilitar una trazabilidad más precisa y contribuir al cumplimiento de normativas sanitarias y de sostenibilidad.

Figura 3

Tipos de Regresión y Ejemplos de Aplicación en Cadenas de Suministro



Ensemble: XGBoost, Gradient Boosting

Los métodos de *ensemble learning* agrupan múltiples modelos débiles para generar un predictor más robusto. Entre los algoritmos más utilizados en esta categoría se encuentran el *Gradient Boosting* y su versión optimizada, *Extreme Gradient Boosting* (XGBoost). Ambos combinan modelos secuenciales, donde cada nuevo modelo corrige los errores del anterior, generando una mejora progresiva en el rendimiento. Estos algoritmos destacan por su alta precisión, tolerancia al ruido y capacidad para manejar datos faltantes, lo que los convierte en una opción preferida en competencias de ciencia de datos y en aplicaciones industriales exigentes (Chen, 2016).

En cadenas de suministro, XGBoost permite predecir rupturas de stock, optimizar rutas de entrega y anticipar la evolución de la demanda en función de múltiples variables. Su capacidad para manejar relaciones no lineales y su flexibilidad ante datos heterogéneos lo hacen ideal para contextos donde la precisión es prioritaria. No obstante, su desventaja principal radica en su complejidad de interpretación, lo cual puede dificultar su adopción en entornos donde la explicabilidad es una necesidad regulatoria.

Aprendizaje No Supervisado

El aprendizaje no supervisado representa una categoría fundamental dentro del campo del *machine learning*, en la cual los algoritmos procesan datos sin etiquetas o salidas previamente definidas. A diferencia del aprendizaje supervisado, que utiliza pares de entrada y salida para entrenar modelos, el aprendizaje no supervisado se enfoca en descubrir patrones subyacentes, estructuras latentes o relaciones significativas dentro de los conjuntos de datos (Goodfellow, Bengio y Courville, 2016). Esta característica lo convierte en una herramienta poderosa para el

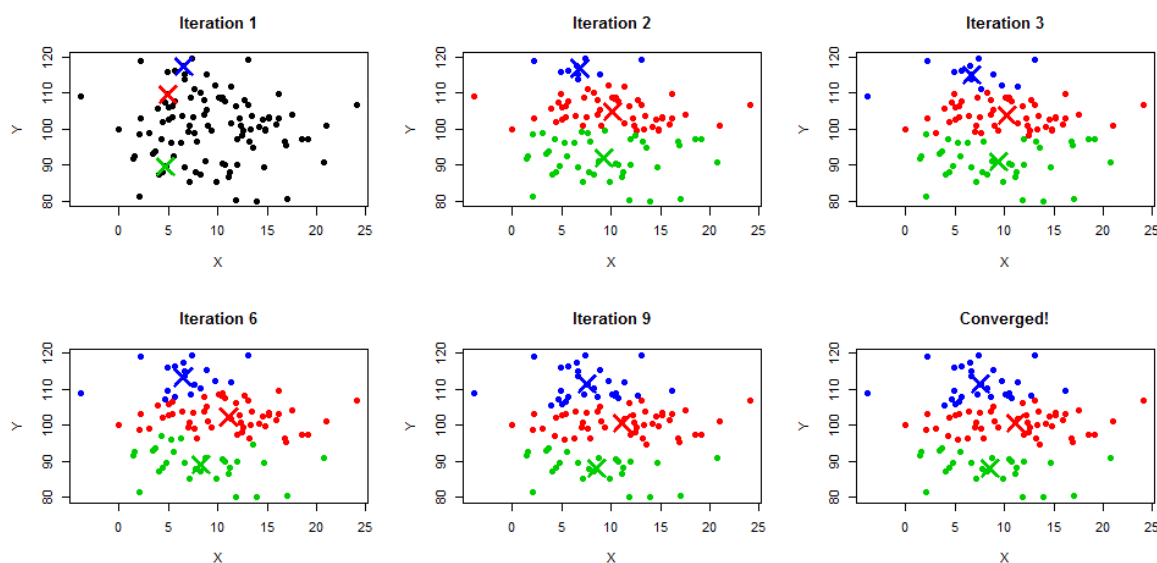
análisis exploratorio y la generación de conocimiento en contextos donde la información etiquetada es limitada o inexistente.

Clustering: K-Means, Clustering Jerárquico, DBSCAN

Uno de los enfoques más representativos del aprendizaje no supervisado es el agrupamiento o *clustering*. Este conjunto de técnicas busca dividir los datos en grupos o clústeres de elementos similares según alguna métrica de distancia. El algoritmo K-Means es uno de los más utilizados debido a su simplicidad y eficiencia. Este método parte de una selección aleatoria de centroides y asigna los datos al clúster más cercano, actualizando iterativamente los centroides hasta alcanzar una convergencia (Jain, 2010)

Figura 4

Agrupación en el aprendizaje automático



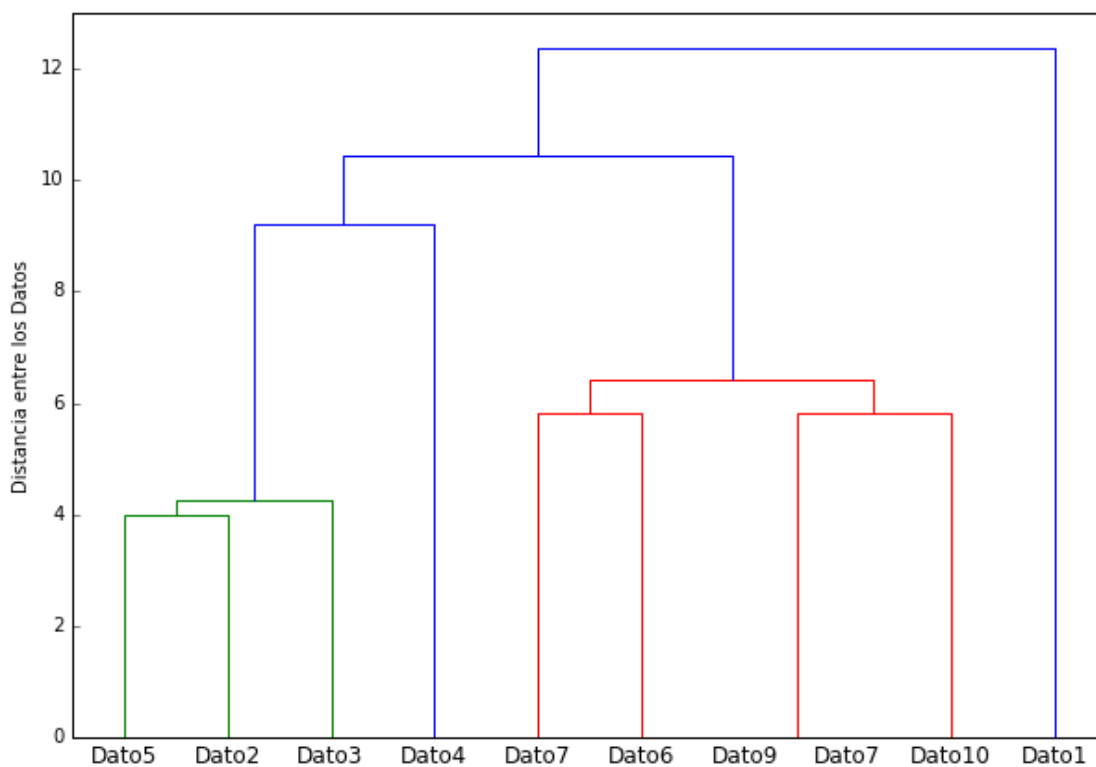
Nota. Miller (2016)

Por su parte, el *clustering* jerárquico no requiere la especificación previa del número de clústeres. En cambio, construye una jerarquía de agrupaciones que puede representarse mediante un dendrograma. Esta metodología puede ser aglomerativa o divisiva, dependiendo de si

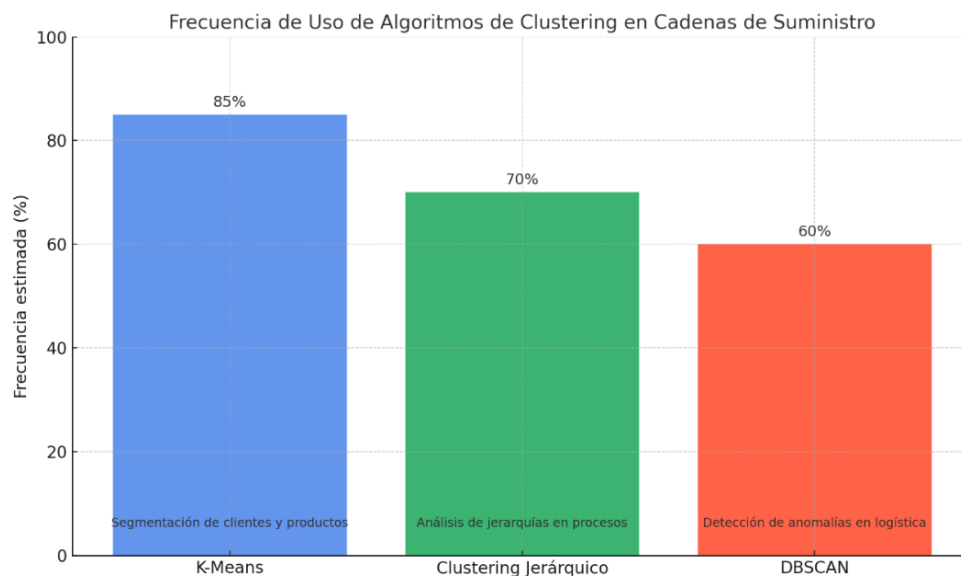
comienza uniendo elementos individuales o dividiendo el conjunto completo, respectivamente (Müller y Guido, 2016). Otro algoritmo relevante es DBSCAN, que identifica clústeres como áreas densas de puntos separados por regiones de baja densidad. Esta técnica es particularmente útil para detectar estructuras de forma arbitraria y manejar ruido en los datos (Ester et al., 1996).

Figura 5

Clustering Jerárquico



Nota. Miller (2016)

Figura 6*Frecuencia de Uso de Algoritmos de Clustering en Cadenas de Suministro****Reducción de Dimensionalidad: PCA, t-SNE***

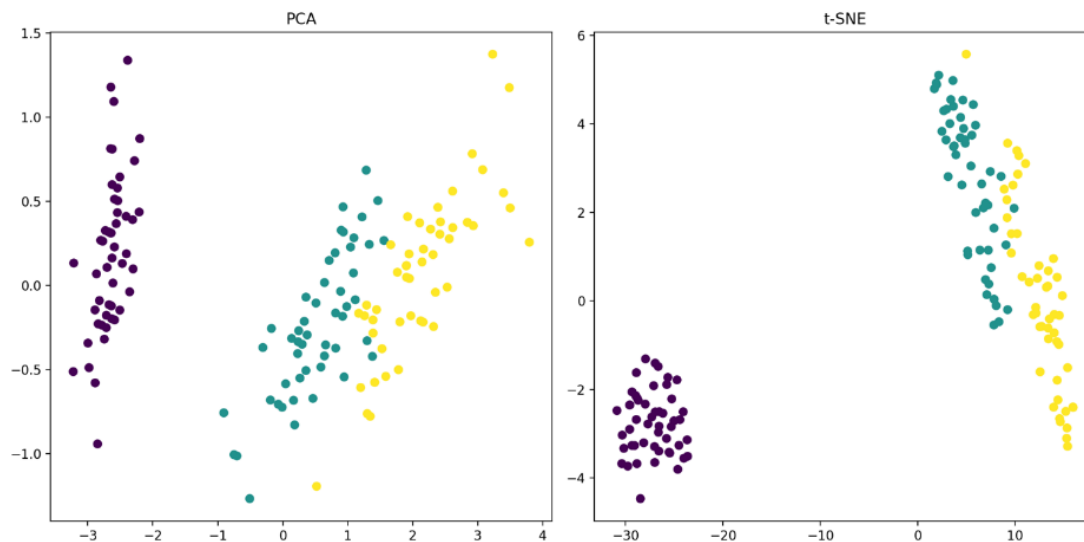
En problemas con un gran número de variables, la reducción de dimensionalidad se convierte en una necesidad para evitar el fenómeno conocido como la maldición de la dimensionalidad. La técnica más clásica en este ámbito es el Análisis de Componentes Principales (PCA), que transforma las variables originales en un nuevo conjunto de componentes lineales ortogonales que capturan la mayor varianza posible de los datos (Jolliffe y Cadima, 2016).

El PCA es ampliamente utilizado debido a su interpretabilidad y eficiencia computacional. Sin embargo, no es ideal cuando las relaciones en los datos son no lineales. Para estos casos, se ha desarrollado el algoritmo t-SNE (*t-distributed Stochastic Neighbor Embedding*), que permite representar datos de alta dimensionalidad en espacios de dos o tres dimensiones, conservando las relaciones de proximidad locales. Esta técnica ha sido aplicada

exitosamente en la visualización de datos genómicos, imágenes y texto (van der Maaten y Hinton, 2008).

Figura 7

PCA, t-SNE



Nota. Miller (2016)

Asociación: A priori

El análisis de reglas de asociación constituye otra línea de investigación clave en el aprendizaje no supervisado. El algoritmo Apriori se basa en la identificación de conjuntos frecuentes de ítems que aparecen juntos en transacciones, y a partir de estos, genera reglas de asociación. Este método ha sido utilizado principalmente en el análisis de canastas de compra, pero también se aplica en bioinformática, telecomunicaciones y diagnóstico médico (Agrawal y Srikant, 1994).

El Apriori trabaja incrementando gradualmente el tamaño de los conjuntos de ítems, verificando si su frecuencia en la base de datos supera un umbral mínimo preestablecido. Posteriormente, se generan reglas que deben cumplir con criterios de soporte, confianza y

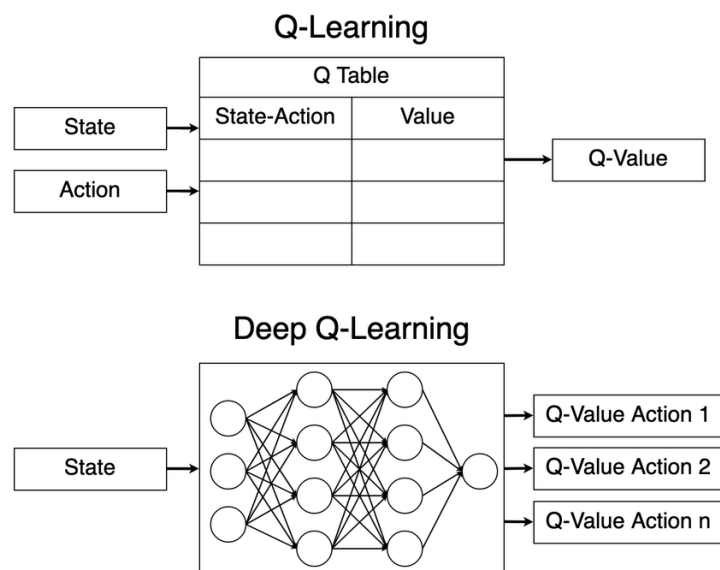
elevación, para garantizar que la relación entre los ítems no sea aleatoria. Esta técnica proporciona información valiosa para el diseño de estrategias comerciales, recomendación de productos y segmentación de mercados.

Aprendizaje por Refuerzo

El aprendizaje por refuerzo constituye una de las ramas más avanzadas y complejas del *machine learning*. A diferencia del aprendizaje supervisado y no supervisado, este enfoque se basa en la interacción continua entre un agente y su entorno, donde el agente toma decisiones con base en un sistema de recompensas y penalizaciones. El objetivo es maximizar una función de recompensa acumulada a lo largo del tiempo, mediante la selección de acciones que conduzcan a estados beneficiosos. Este paradigma resulta útil en situaciones donde las decisiones tienen consecuencias secuenciales y no están disponibles datos etiquetados de manera directa (Sutton y Barto, 2018).

Q-Learning

Entre los algoritmos fundamentales del aprendizaje por refuerzo se encuentra Q-Learning. Este método se basa en una tabla de valores Q, donde se almacenan las recompensas esperadas para cada par estado-acción. El modelo se actualiza iterativamente según una fórmula que considera la recompensa inmediata y el valor estimado del estado futuro, lo que permite al agente aprender una política óptima sin necesidad de un modelo del entorno (Watkins y Dayan, 1992).

Figura 8*Q-Learning Y Deep Q-Learning*

Nota. Silver (2017)

Q-Learning es considerado un algoritmo fuera de política, ya que aprende la acción óptima independientemente de la política seguida por el agente. Esta característica lo hace versátil en entornos donde las transiciones de estado son estocásticas. No obstante, su implementación está limitada a espacios de estado reducidos, ya que la tabla Q crece exponencialmente con el número de estados y acciones, lo que afecta la escalabilidad del modelo (Kaelbling, Littman y Moore, 1996).

Deep Q-Networks (DQN)

Para superar las limitaciones del Q-Learning clásico en entornos con espacios de estados continuos o de alta dimensionalidad, se han desarrollado los *Deep Q-Networks* (DQN). Estos modelos utilizan redes neuronales profundas para aproximar la función Q, permitiendo generalizar mejor en espacios grandes. El uso de memoria de experiencia y de objetivos diferidos

son mecanismos que estabilizan el aprendizaje, evitando oscilaciones durante el entrenamiento (Mnih et al., 2015).

Los DQN han demostrado un rendimiento sobresaliente en tareas complejas, como videojuegos del sistema Atari, donde logran superar el desempeño humano en varios títulos. Esta capacidad de aprendizaje endógeno, sin supervisión explícita, convierte al DQN en una de las metodologías más relevantes dentro del aprendizaje por refuerzo profundo. No obstante, el entrenamiento de estos modelos requiere una cantidad significativa de recursos computacionales, además de una gran cantidad de episodios para lograr una convergencia adecuada.

Policy Gradients

Otro enfoque importante en el aprendizaje por refuerzo es el uso de gradientes de política, conocidos como *Policy Gradients*. En lugar de aprender una función de valor, estos algoritmos aprenden directamente una política estocástica que asigna probabilidades a cada acción dada una observación del entorno. La actualización de la política se realiza a través del cálculo del gradiente de la función de recompensa esperada, permitiendo una mayor flexibilidad en entornos continuos o no deterministas (Williams, 1992).

Entre los algoritmos basados en esta técnica se encuentran REINFORCE, Proximal Policy Optimization (PPO) y Trust Region Policy Optimization (TRPO). Estos métodos son particularmente útiles cuando se requiere una exploración eficiente del espacio de acciones, como ocurre en tareas de control robótico. A diferencia de Q-Learning, los Policy Gradients permiten manejar entornos donde las decisiones no pueden ser representadas en forma tabular, lo que amplía su aplicabilidad en escenarios reales.

Aplicaciones de Aprendizaje Supervisado: Predicción de Precios, Detección de Spam

El aprendizaje supervisado se ha consolidado como una herramienta central en la transformación digital de múltiples sectores económicos, siendo particularmente relevante en el diseño de soluciones para la industria alimentaria y la gestión de cadenas de suministro sostenibles. Dentro de esta categoría de aprendizaje automático, los casos de uso como la predicción de precios, la detección de correos electrónicos no deseados y el diagnóstico médico han sido extensamente documentados. Sin embargo, su aplicabilidad va más allá de estos ejemplos clásicos, extendiéndose a procesos críticos relacionados con la eficiencia logística, la reducción del desperdicio y el cumplimiento normativo en sistemas complejos como las cadenas de valor agroalimentarias.

En el ámbito de la predicción de precios, los modelos de regresión han demostrado utilidad al anticipar fluctuaciones en el costo de productos agroindustriales, lo cual es esencial para la planificación de compras, la asignación de recursos y la toma de decisiones estratégicas en la cadena de suministro. Por ejemplo, modelos de regresión lineal múltiple han sido implementados para estimar el precio de insumos agrícolas con base en variables como la estacionalidad, el clima, la localización geográfica y el historial de oferta y demanda. En el contexto de la industria alimentaria, esta capacidad predictiva permite a los productores ajustar volúmenes de producción y almacenamiento de manera más eficiente, lo que a su vez contribuye a minimizar pérdidas económicas y reducir la generación de desperdicios por sobreproducción o desabastecimiento (Sarker, 2021).

La integración de modelos supervisados de predicción en sistemas de información logística permite mejorar la sincronización entre actores de la cadena de suministro. Cuando los agentes pueden anticipar los precios de productos perecederos, se habilita una programación de

entregas más eficiente y alineada con las ventanas de comercialización. Esto se traduce en una reducción de costos operativos y emisiones asociadas al transporte, alineándose con objetivos de sostenibilidad y con regulaciones internacionales como las establecidas en el Pacto Verde Europeo y los Objetivos de Desarrollo Sostenible (ODS) (FAO, 2019).

En relación con la detección de spam, si bien este ejemplo es clásico en el ámbito de la clasificación binaria, su lógica es trasladable a la identificación automática de registros o eventos anómalos en cadenas de suministro. Por ejemplo, los algoritmos utilizados para discriminar correos no deseados también se aplican para detectar desviaciones en patrones de consumo o distribución. Esta funcionalidad resulta particularmente relevante en escenarios donde el monitoreo en tiempo real permite evitar decisiones ineficientes o riesgos para la inocuidad de los productos.

A través del uso de clasificadores como árboles de decisión, bosques aleatorios o máquinas de soporte vectorial, es posible automatizar el control de calidad en las distintas fases de producción y distribución. Por ejemplo, los sistemas pueden aprender a identificar transacciones logísticas que difieren del comportamiento esperado, tales como envíos repetidos a destinos inusuales o retrasos recurrentes en tramos específicos de la cadena. Esto permite detectar fallos estructurales, corregirlos de manera oportuna y garantizar la trazabilidad de los productos. De esta forma, los modelos supervisados contribuyen directamente a la mejora continua de los procesos críticos de la cadena, alineándose con estrategias de sostenibilidad empresarial y minimizando el impacto ambiental derivado de errores operativos (Rodríguez, 2022).

En el campo del diagnóstico médico, los avances en reconocimiento de patrones han permitido la detección automática de enfermedades a partir de imágenes y datos clínicos. Este

tipo de aplicación ha sido replicado en la industria alimentaria mediante sistemas de visión computacional que identifican defectos en frutas, verduras o productos empacados. Por medio de redes neuronales entrenadas con imágenes etiquetadas, se ha conseguido mejorar la clasificación de productos, eliminando aquellos que no cumplen con estándares de calidad y evitando que lleguen al consumidor. Esta estrategia no solo mejora la percepción del cliente final, sino que también reduce desperdicios y mejora la eficiencia en la distribución, al garantizar que únicamente se transporten productos aptos para su comercialización (Meléndez Mogollón et al., 2022).

En las plantas de procesamiento, los modelos supervisados también se utilizan para predecir fallos en maquinaria mediante el análisis de variables como la vibración, la temperatura o la presión. Estos modelos permiten anticipar averías y planificar acciones de mantenimiento preventivo, lo que disminuye los tiempos de inactividad y optimiza el uso de energía. Esta reducción en el consumo energético se traduce en una menor huella de carbono, cumpliendo con criterios de eficiencia energética establecidos en marcos normativos como la Agenda 2030 y el Protocolo de Kioto.

En cuanto al análisis predictivo de la demanda, los modelos de regresión multivariada y redes neuronales han sido entrenados con datos históricos de consumo, comportamiento del cliente y tendencias del mercado. Esto permite prever volúmenes de venta y ajustar en tiempo real la producción, evitando tanto el desabastecimiento como el exceso de stock. El impacto de esta aplicación es doble: por un lado, permite reducir los costos de almacenamiento y, por otro, mejora la rotación de inventarios, reduciendo así el desperdicio de productos perecederos.

La alineación de estas herramientas con los marcos de sostenibilidad se evidencia en su contribución a la trazabilidad, eficiencia energética y uso racional de recursos. La ciencia de

datos aplicada, mediante el uso de modelos supervisados, genera un sistema de control adaptativo que permite a las empresas responder de manera ágil a las exigencias del mercado y a los desafíos ambientales. Además, permite cumplir con requisitos normativos como la Directiva de Diligencia Debida en Sostenibilidad Corporativa de la Unión Europea, que exige a las empresas evaluar y mitigar sus impactos ambientales y sociales en toda la cadena de suministro.

Asimismo, los modelos de clasificación supervisada se han empleado para predecir la aceptación de nuevos productos alimentarios. A partir de datos recopilados en encuestas y redes sociales, los algoritmos permiten segmentar a los consumidores según sus preferencias y anticipar la reacción del mercado ante un nuevo lanzamiento. Esta capacidad predictiva permite reducir riesgos de inversión, mejorar la estrategia de marketing y diseñar productos alineados con las expectativas de los consumidores, lo que a su vez favorece el posicionamiento competitivo de la empresa y su sostenibilidad financiera.

Tabla 1*Casos de Uso en Aprendizaje Supervisado*

Caso de uso	Modelo de ML aplicado	Impacto en cadena de suministro
Predicción de precios	Regresión lineal, múltiple	Optimización de compras y producción, reducción de desperdicios
Detección de spam	Clasificación (SVM, Random Forest)	Identificación de anomalías logística, mejora en trazabilidad
Diagnóstico médico	Redes neuronales, SVM	Clasificación de productos no conformes, mejora en seguridad alimentaria
Control de calidad en alimentos	Clasificadores supervisados (árboles, Random forest)	Mejora de eficiencia en distribución y trazabilidad
Mantenimiento predictivo	Regresión logística, redes neuronales	Reducción de fallos, ahorro energético, cumplimiento normativo
Análisis de demanda de productos	Regresión multivariada, redes neuronales	Optimización de inventarios, mejora en logística y reducción de residuos

Por tanto, los casos de uso del aprendizaje supervisado, cuando se contextualizan en la cadena de suministro alimentaria, superan los ejemplos tradicionales para convertirse en herramientas clave de transformación. Su implementación ha contribuido a la mejora de procesos logísticos, al fortalecimiento de sistemas de trazabilidad, a la minimización del impacto ambiental y al cumplimiento de normas internacionales (Amado, 2024). La integración de estos

modelos en las estrategias empresariales no solo optimiza el desempeño operativo, sino que también posiciona a las organizaciones frente a los desafíos de la sostenibilidad y la transformación digital.

Limitaciones

Pese a sus beneficios, el aprendizaje supervisado presenta limitaciones que deben ser consideradas. En primer lugar, requiere grandes cantidades de datos etiquetados, cuya obtención puede ser costosa, lenta y sujeta a errores humanos. Además, si el conjunto de entrenamiento no representa adecuadamente la variabilidad del entorno real, el modelo puede generalizar mal, afectando su aplicabilidad (Mitchell, 2021).

Asimismo, existe el riesgo de sobreajuste, especialmente en modelos complejos como las redes neuronales profundas, que pueden memorizar los datos de entrenamiento sin capturar las relaciones subyacentes. Este problema se agrava cuando los datos contienen ruido o patrones espurios. Finalmente, en algunos algoritmos, como XGBoost o ANN, la interpretabilidad del modelo es limitada, lo cual dificulta su aplicación en contextos que requieren transparencia y justificación explícita de las decisiones (Molnar, 2020).

Aplicaciones de Aprendizaje no Supervisado: Segmentación, Anomalías, Recomendación, Reducción

El aprendizaje no supervisado tiene aplicaciones transversales en diversas industrias. En marketing, el *clustering* permite segmentar consumidores en grupos con comportamientos similares, facilitando campañas personalizadas y estrategias de fidelización. En ciberseguridad, estos métodos se utilizan para detectar anomalías en patrones de acceso o transacciones inusuales que puedan indicar actividades fraudulentas (Chandola, Banerjee y Kumar, 2009).

En plataformas de contenido, como servicios de streaming o comercio electrónico, las reglas de asociación permiten generar sistemas de recomendación que ofrecen productos o servicios relacionados con base en interacciones pasadas. La reducción de dimensionalidad, por su parte, facilita la visualización de datos complejos, permitiendo a los analistas detectar.

Ventajas

Entre las principales ventajas del aprendizaje no supervisado se encuentra su independencia de datos etiquetados, lo que lo hace especialmente valioso en contextos donde el etiquetado manual resulta inviable por costos o volumen. Esta flexibilidad permite explorar estructuras y patrones en los datos que podrían pasar desapercibidos con otros enfoques. Además, facilita el descubrimiento de conocimientos ocultos, lo que puede conducir a nuevas hipótesis o líneas de investigación (Hastie, Tibshirani y Friedman, 2017).

Otra ventaja radica en su capacidad para detectar relaciones complejas en datos multidimensionales, gracias al uso de técnicas como la reducción de dimensionalidad o los modelos de mezcla. Esta cualidad es fundamental en áreas como la bioinformática o el análisis financiero, donde los datos suelen tener alta complejidad y correlaciones no evidentes. También permite el preprocesamiento de datos, facilitando la limpieza, compresión o preparación para modelos supervisados.

Limitaciones

A pesar de sus ventajas, el aprendizaje no supervisado presenta limitaciones importantes. Una de las más relevantes es la dificultad para evaluar la calidad de los resultados, ya que no existen etiquetas de referencia que permitan validar la agrupación o la representación obtenida. Esto hace necesario el uso de métricas indirectas como la cohesión intra-clúster, la separación inter-clúster o la interpretabilidad de las dimensiones reducidas (Tan, Steinbach y Kumar, 2018).

Además, muchos algoritmos son sensibles a los parámetros iniciales, como el número de clústeres en K-Means o los umbrales de densidad en DBSCAN. Una mala elección de estos parámetros puede conducir a resultados subóptimos o incluso engañosos. También puede ser complejo interpretar los resultados, especialmente en técnicas como t-SNE, donde las distancias en la visualización no reflejan necesariamente relaciones cuantitativas entre los datos originales. Es notable que, el aprendizaje no supervisado ofrece oportunidades poderosas para el descubrimiento de patrones y la compresión de estructuras complejas en datos no etiquetados. No obstante, su aplicación requiere consideraciones metodológicas rigurosas y una interpretación cuidadosa de los resultados para evitar errores o sobreinterpretaciones.

Aplicaciones Aprendizaje por Refuerzo: Robots, Rutas, Juegos

El aprendizaje por refuerzo ha encontrado aplicaciones significativas en diversos dominios. En robótica, se utiliza para enseñar a agentes a caminar, manipular objetos o interactuar con entornos complejos sin instrucciones explícitas. Estos modelos permiten optimizar movimientos a través de la retroalimentación continua, lo que reduce la necesidad de programación manual detallada (Kober, Bagnell y Peters, 2013).

En logística y transporte, se emplea para la optimización de rutas en tiempo real, considerando factores como tráfico, costos de combustible o disponibilidad de recursos. Esta aplicación es particularmente valiosa en redes de distribución urbana o en cadenas de suministro donde las condiciones cambian dinámicamente. En el campo de los videojuegos y la inteligencia artificial general, el aprendizaje por refuerzo ha sido clave en el desarrollo de agentes que han vencido a campeones humanos en juegos como Go, ajedrez y StarCraft, demostrando su capacidad para aprender estrategias complejas a partir de la experiencia (Silver et al., 2017).

Ventajas

Una de las principales ventajas del aprendizaje por refuerzo es su capacidad para manejar problemas de toma de decisiones secuenciales en entornos inciertos. Esta característica lo diferencia de otros enfoques de *machine learning*, ya que considera la historia de interacciones del agente y los efectos a largo plazo de sus decisiones. Además, permite la adaptación continua del comportamiento del agente, lo que resulta útil en sistemas que deben operar en condiciones cambiantes o impredecibles (Arulkumaran et al., 2017).

Otra ventaja relevante es su aplicabilidad a tareas donde no existen datos etiquetados, sino que la información proviene de la interacción directa con el entorno. Esto amplía su uso a contextos donde el aprendizaje supervisado sería inviable o insuficiente. Finalmente, su capacidad para aprender políticas óptimas a partir de recompensas escasas permite resolver problemas complejos que requieren exploración activa del espacio de decisiones.

Limitaciones

Pese a sus beneficios, el aprendizaje por refuerzo enfrenta limitaciones importantes. Uno de los principales desafíos es la eficiencia *sample*, es decir, la cantidad de interacciones necesarias para lograr una política efectiva. En muchos casos, los modelos requieren millones de episodios para aprender comportamientos adecuados, lo que representa un obstáculo en entornos del mundo real donde las simulaciones son costosas o lentas (Dulac-Arnold et al., 2021).

Además, la estabilidad del entrenamiento es un problema frecuente, especialmente en algoritmos que utilizan redes neuronales. La sensibilidad a los hiperparámetros, la exploración insuficiente y el desbalance en las recompensas pueden generar políticas subóptimas o conductas no deseadas. Por tanto, la interpretabilidad de los modelos sigue siendo un desafío, ya que las

decisiones aprendidas pueden ser difíciles de justificar ante usuarios o reguladores, especialmente en ámbitos sensibles como la salud o la seguridad.

Tabla 2

Comparativa de Tipos de Aprendizaje en Machine Learning

Tipo de Aprendizaje	Definición	Algoritmos Representativos	Aplicaciones Clave	Ventajas	Desventajas
Aprendizaje Supervisado	El modelo aprende a partir de datos etiquetados con pares entrada – salida conocidos	Regresión Lineal, Regresión Logística, Árbol de decisión, Random Forest, SVM, ANN, XGBoost	Predicción de precios, diagnóstico médico, detección de spam, control de calidad, predicción de demanda	Alta precisión con datos etiquetados, facilidad de validación y evaluación y buena interpretabilidad en algunos modelos	Requiere grandes cantidades de datos etiquetados, riesgo de sobreajuste, interpretabilidad limitada en modelos complejos
Aprendizaje No Supervisado	El modelo identifica patrones o estructura en datos no etiquetados	K-Means, Clustering jerárquico, DBSCAN, SCAN, t-SNE, Apriori	Segmentación de clientes, detección de anomalías, análisis exploratorio, reducción de dimensionalidad	No requiere datos etiquetados, útiles para explorar grandes volúmenes de datos, facilita descubrimiento de patrones ocultos	Dificultad para validar resultados, interpretación compleja, sensibilidad a parámetros de configuración
Aprendizaje por Refuerzo	El modelo aprende tomando decisiones en un entorno a través de recompensas y penalizaciones	Q-Learning, Deep Q-Networks (DQN), Policy Gradients, PPO, TRPO	Control de robots, optimización de rutas, toma de decisiones en video juegos, sistemas dinámicos	Capacidad de aprendizaje en entornos dinámicos, maneja decisiones secuenciales, útil en ausencia de datos etiquetados	Alta demanda computacional, requiere muchas iteraciones para aprender, poca interpretabilidad en redes profundas

Metodología

Esta monografía se desarrolla bajo un enfoque metodológico de revisión sistemática de nivel descriptivo-analítico, orientado a identificar, organizar e interpretar los principales aportes científicos relacionados con la aplicación del machine learning en la optimización de cadenas de suministro sostenibles dentro de la industria alimentaria. El estudio tiene un alcance descriptivo, al sistematizar casos prácticos y experiencias documentadas en artículos académicos, y un enfoque analítico, al interpretar críticamente los impactos económicos, sociales y ambientales derivados de dichas aplicaciones tecnológicas. El objetivo central es proponer un marco teórico fundamentado en la literatura que permita comprender cómo el machine learning contribuye a la sostenibilidad operativa y normativa del sector alimentario, particularmente en contextos latinoamericanos. No obstante, el estudio presenta limitaciones inherentes a su carácter documental, como la dependencia de la disponibilidad de estudios publicados en bases de datos científicas, la posible variabilidad en los criterios metodológicos de los trabajos incluidos, y la falta de medición empírica directa, lo cual restringe la generalización de los hallazgos a contextos específicos.

Criterios de Selección y Alcance de la Revisión Bibliográfica

Se realiza una revisión exhaustiva de evidencia científica actualizada que apunten a identificar los enfoques y modelos de machine learning aplicados en la optimización de las cadenas de suministro alimentarias sostenibles.

Para garantizar la rigurosidad de la revisión, se definieron criterios de inclusión como publicaciones entre 2015 y 2024, utilizando como herramienta bases de datos de Google Scholar, complementada con exploraciones académicas abiertas como DOAJ, Scielo, y Semantic Scholar. Se utilizaron combinaciones de palabras clave como:

"machine learning" + "food supply chain" + "sustainability" y "optimización logística" + "inteligencia artificial" + "industria alimentaria".

Además, se obtuvieron estudios revisados por pares entre artículos, libros académicos y tesis, de los cuales se documentaron en español e inglés que abordan el uso de *machine learning* en cadenas de suministro, sostenibilidad o industria alimentaria, con una evidencia empírica o conceptual que relacione el impacto económico, social y/o ambiental de estas tecnologías.

Por otro lado, se definieron también criterios de exclusión como: fuentes no académicas sin respaldo metodológico, estudios que no contemplen ningún eje de sostenibilidad y aplicaciones de *machine learning* en industrias no alimentarias, salvo que aporten modelos transferibles.

La mayoría de los estudios coinciden en la utilidad del *machine learning* para mejorar la eficiencia logística, reducir desperdicios y predecir la demanda. Sin embargo, divergen algunos autores donde destacan la efectividad económica, pero otros advierten sobre barreras de adopción tecnológica, falta de talento humano o viabilidad en contextos con infraestructuras limitadas.

Se identifica una limitada producción académica aplicada directamente al contexto latinoamericano, especialmente en cadenas agroalimentarias medianas o rurales. Aunque muchos modelos muestran resultados exitosos en países desarrollados, su aplicabilidad en Colombia y América Latina depende de condiciones como la infraestructura digital, el acceso a datos confiables y la formación técnica de los equipos de trabajo. En este sentido, la transferencia tecnológica debe ir acompañada de adaptaciones estructurales y culturales.

Impacto Económico, Social y Ambiental del *Machine Learning* en la Industria Alimentaria

La implementación del *machine learning* en las cadenas de suministro alimenticias ha generado transformaciones estructurales en la forma como se diseñan, gestionan y optimizan los procesos de producción, distribución, almacenamiento y comercialización de alimentos. Esta tecnología permite analizar grandes volúmenes de datos generados en distintos puntos de la red logística, identificando patrones que optimizan la eficiencia operativa y reducen los impactos ambientales. En particular, los modelos de regresión multivariada y redes neuronales han demostrado eficacia al anticipar comportamientos de consumo con base en variables como estacionalidad, clima, eventos regionales y tendencias del mercado, contribuyendo a una planificación más precisa (Meléndez Mogollón et al., 2022).

En la etapa de predicción de la demanda, el uso de datos históricos y actuales permite a los algoritmos identificar ciclos de consumo y ajustar la oferta con base en escenarios reales. Esta capacidad mejora la planeación de la producción, reduce los desperdicios derivados de sobreoferta y evita fallas de abastecimiento, elementos clave para la sostenibilidad en cadenas alimentarias (Chen, Chu y Zhao, 2024). La anticipación de la demanda, además de optimizar la operación logística, contribuye a la alineación con objetivos regulatorios como la reducción del desperdicio alimentario en la Agenda 2030 de la ONU (FAO, 2019).

En la gestión de inventarios, los algoritmos de clasificación y agrupamiento permiten categorizar productos según variables como la rotación, condiciones de conservación o susceptibilidad al deterioro. A través de modelos como K-Means y árboles de decisión, es posible estructurar almacenes de forma dinámica, asignando ubicaciones que reduzcan el tiempo de tránsito interno y la exposición a condiciones adversas. Esta estrategia conocida como “stock

inteligente” maximiza la eficiencia en el uso del espacio físico y disminuye la pérdida por caducidad, aspecto fundamental en productos perecederos (Badghish y Soomro, 2024).

Por otro lado, en la distribución y transporte de alimentos, los modelos de aprendizaje por refuerzo han permitido optimizar rutas con base en múltiples variables operativas, incluyendo distancias, horarios, consumo de combustible y condiciones de tráfico. Estos modelos aprenden a través de simulaciones y retroalimentación, proponiendo decisiones óptimas que reducen los costos logísticos y las emisiones de gases contaminantes (Ochoa-Barragán et al., 2024). En entornos urbanos, esta tecnología permite hacer ajustes dinámicos en tiempo real, mejorando la eficiencia energética y minimizando el impacto ambiental de la logística alimentaria (Ding et al., 2023).

El control de calidad también ha sido transformado por la implementación de técnicas de visión por computador. Las redes neuronales convolucionales (CNN) permiten analizar imágenes de productos durante el proceso de clasificación, detectando defectos, contaminaciones o características anómalas. Estas soluciones, entrenadas con datos visuales etiquetados, garantizan que solo los productos que cumplen con los estándares establecidos sean distribuidos, fortaleciendo la inocuidad alimentaria y reduciendo reprocesos (Hong y Xiao, 2024). El uso de CNN ha sido particularmente útil en productos frescos, como frutas y verduras, donde la inspección manual es lenta y menos precisa.

La trazabilidad automatizada es otro componente central del *machine learning* en cadenas de suministro. Algoritmos supervisados y sistemas basados en sensores permiten rastrear el recorrido del producto desde su origen hasta el consumidor final. Esta capacidad, complementada con tecnologías como blockchain, mejora la transparencia de la cadena y permite responder ante alertas sanitarias o devoluciones de producto. La trazabilidad

automatizada también fortalece la relación con el consumidor, quien exige cada vez más información sobre el origen y condiciones de producción de los alimentos (Rodríguez, 2022).

En cuanto al monitoreo ambiental, el *machine learning* facilita la evaluación de indicadores de sostenibilidad como consumo energético, generación de residuos o uso del agua. Modelos de series temporales y redes neuronales recurrentes permiten anticipar picos de consumo y diseñar estrategias correctivas. Esta información, aplicada al control de plantas de producción, ayuda a cumplir con regulaciones como el Pacto Verde Europeo y a mitigar el impacto ambiental de las operaciones logísticas (Gaur et al., 2023). El análisis predictivo también permite ajustar prácticas de producción en función de la disponibilidad de recursos naturales, como energía o materias primas.

La integración con tecnologías como el Internet de las Cosas (IoT) ha permitido recolectar datos en tiempo real desde sensores ubicados en centros de acopio, vehículos refrigerados y almacenes. Estos datos, analizados con algoritmos de *machine learning*, generan alertas automáticas en caso de desviaciones que puedan comprometer la calidad del producto. El procesamiento simultáneo de datos sobre temperatura, humedad, ubicación y tiempos de tránsito optimiza la toma de decisiones en tiempo real, mejorando la eficiencia de la cadena y garantizando la estabilidad del producto (Amado Mateus, 2024).

Desde una perspectiva económica, la automatización inteligente de las cadenas de suministro alimenticias ha permitido reducir costos operativos, redistribuir recursos humanos y mejorar la toma de decisiones estratégicas. El uso de modelos de regresión y clasificación en la planeación logística mejora la asignación de tareas, minimiza los tiempos de espera y reduce el error humano. A su vez, las predicciones de rotación y demanda permiten realizar compras más

ajustadas, lo que disminuye los costos por almacenamiento innecesario (Coronado-Hernández y García-Sabater, 2019).

En el plano normativo, los modelos de machine learning apoyan el cumplimiento de marcos regulatorios que exigen prácticas sostenibles. La recopilación y análisis de datos permite documentar la reducción de emisiones, el uso eficiente de recursos y la disminución del desperdicio alimentario. Esta trazabilidad normativa es fundamental para exportadores y empresas que operan en múltiples jurisdicciones, donde los estándares ambientales son parte del criterio de acceso a mercados (Al Zaabi et al., 2020). Además, el cumplimiento con normas de sostenibilidad fortalece la reputación de las organizaciones y mejora su posicionamiento competitivo.

En contextos de crisis, como interrupciones en la cadena por conflictos, pandemias o eventos climáticos extremos, los sistemas basados en machine learning permiten generar escenarios alternativos y planes de contingencia. A través del modelado predictivo y el análisis de riesgos, se pueden anticipar disrupciones, identificar puntos vulnerables y tomar decisiones informadas para asegurar el abastecimiento. Esta capacidad de resiliencia es clave para garantizar la seguridad alimentaria y la continuidad operativa en un entorno global cada vez más volátil (Dhiman et al., 2024).

Actualmente, los proyectos aplicados del machine learning en las cadenas de suministro alimenticias representa una solución estratégica para mejorar la eficiencia operativa, fortalecer la trazabilidad, reducir el desperdicio, garantizar la calidad y cumplir con las normativas ambientales. Este enfoque, se sustenta en el uso de datos, la transformación de la lógica tradicional de la cadena de suministro y la evidencia a la adaptación de los desafíos globales de la sostenibilidad, eficiencia y seguridad (Chen et al., 2024).

El trabajo de investigación publicado por Ochoa-Barragán et al. (2024) – “Sustainable Lime Production in Michoacán Mexico: An Optimal and Equitable Approach with Machine Learning.”, describe un proyecto aplicado que integró machine learning para optimizar la cadena de suministro de lima (cítrico) de manera sostenible. Mediante algoritmos de ML combinados con modelos de optimización, se logró reducir el impacto ambiental en la producción (uso más eficiente de energía y materias primas) al mismo tiempo que se mejoró la equidad en la distribución de beneficios a lo largo de la cadena. El caso demuestra cómo la IA puede apoyar una producción alimentaria más verde (menores emisiones y residuos), socialmente responsable (trato equitativo a productores locales) y rentable (manteniendo la viabilidad económica del negocio). El modelo de aprendizaje automático fue programado en Python, utilizando conjuntos de datos recopilados desde el año 2000, los cuales comprendían variables socioeconómicas significativas. La herramienta de predicción empleada fue el Random Forest Regressor de la biblioteca Scikit-learn, con 8 estimadores y parámetros de validación como el valor P y el R^2 ajustado. En general, el estudio propone una estrategia que combina herramientas de aprendizaje automático y optimización determinista para mejorar la cadena de suministro de limón en Michoacán, México. Se optimizó la producción y distribución de limón mexicano y persa, considerando beneficios económicos, uso de agua para riego y reducción de emisiones de CO_2eq . Además, se identificaron impactos negativos en la producción agrícola derivados de la violencia y criminalidad en la región.

Para contrarrestar estas afectaciones, se sugiere implementar un esquema de justicia de Bienestar Social, que demostró ofrecer mayores beneficios económicos a futuro. Las proyecciones indican que, con esta estrategia, se podrían exportar más de 50 mil toneladas de limón al año hacia Estados Unidos. En conjunto, el trabajo ofrece una solución innovadora para

enfrentar los desafíos económicos, sociales y ambientales de la cadena de producción de limón, contribuyendo al uso de la inteligencia artificial en la gestión eficiente de cadenas de suministro agrícolas.

Otro ejemplo significativo de aplicación de modelos de machine learning con impacto económico y social en la cadena de suministro alimentaria es el desarrollado por Niño-Chaparro et al. (2023), en el contexto del mercado agrícola colombiano. En su estudio, se emplearon algoritmos de Random Forest y árboles de decisión para anticipar municipios con alta probabilidad de experimentar choques negativos de precios en productos como la papa y la cebolla. El modelo alcanzó una precisión del 79% en la predicción de zonas en riesgo, permitiendo al gobierno nacional implementar estrategias de subsidios focalizados y preventivos.

Desde una perspectiva económica, este tipo de herramienta permitió optimizar el uso de recursos públicos, evitar pérdidas financieras y mejorar la planificación de distribución, reduciendo el impacto del sobreabastecimiento o de la caída de precios. Socialmente, el sistema ayudó a proteger el sustento de pequeños productores rurales, promoviendo la estabilidad del ingreso y la sostenibilidad de la producción. En términos de la cadena de suministro, este modelo puede integrarse como un componente de inteligencia predictiva, que guía decisiones logísticas estratégicas y reduce vulnerabilidades estructurales ante las dinámicas del mercado.

Por otro lado, en el estudio de Sánchez Jiménez y García Hernández (2019) se ofrece una visión amplia de las aplicaciones de inteligencia artificial y machine learning en la industria alimentaria colombiana, con énfasis en su contribución a la sostenibilidad operativa. El artículo documenta casos reales donde estas tecnologías han sido utilizadas en tareas como el control de calidad, clasificación automatizada de alimentos, predicción de la demanda y optimización de la logística interna. Estas implementaciones han resultado en una reducción comprobada del

desperdicio alimentario, una mayor eficiencia en las operaciones y un incremento en la calidad e inocuidad de los productos destinados al consumo.

Desde la perspectiva de la cadena de suministro, estas aplicaciones tienen un efecto directo en la mejora del desempeño logístico, la toma de decisiones predictiva y la trazabilidad de los productos. Además, al optimizar recursos y reducir desperdicios, las soluciones documentadas no solo generan beneficios económicos para las empresas, sino que también fortalecen la sostenibilidad ambiental y social del sistema agroalimentario colombiano.

También existe un aporte en Colombia del estudio de Gutiérrez Duque y Polo Martínez con una perspectiva integral sobre el papel de la inteligencia artificial en la transformación de las cadenas de suministro. Aunque no se limita exclusivamente al sector alimentario, sus conclusiones resultan altamente aplicables al análisis de esta industria. Las autoras destacan que la adopción de tecnologías basadas en machine learning permite automatizar tareas complejas, optimizar rutas logísticas, reducir errores operativos y mejorar la eficiencia en la toma de decisiones, elementos fundamentales en la gestión moderna de cadenas alimentarias.

En términos económicos, el estudio evidencia cómo estas tecnologías pueden generar una reducción significativa de costos operativos mediante la automatización de procesos y la optimización en el uso de recursos. Asimismo, se observa un impacto ambiental positivo al minimizar desperdicios y mejorar la planificación de distribución, lo que contribuye a una reducción de emisiones y del uso innecesario de materiales. Desde el punto de vista social, las tecnologías inteligentes aportan al fortalecimiento de la trazabilidad y la seguridad de los productos, beneficiando directamente al consumidor.

No obstante, el estudio también alerta sobre desafíos importantes que deben considerarse en procesos de adopción tecnológica en países como Colombia: los altos costos iniciales de

implementación, la necesidad de integración con sistemas existentes y las implicaciones éticas relacionadas con el empleo y la privacidad de los datos. Estos factores son especialmente relevantes en el caso de pequeñas y medianas empresas del sector alimentario, que pueden enfrentar mayores barreras estructurales.

En conclusión, el análisis de Gutiérrez Duque y Polo Martínez permite comprender que la incorporación de machine learning en las cadenas de suministro no debe abordarse únicamente como una cuestión técnica, sino como un proceso estratégico que involucra variables económicas, sociales y ambientales. Su aplicación en la industria alimentaria colombiana requiere no solo inversión tecnológica, sino también planificación adaptativa, regulación adecuada y formación del talento humano que permita sostener estos cambios en el largo plazo.

Finalmente, otro aporte relevante para comprender la relación entre sostenibilidad y tecnología en las cadenas de suministro alimentarias es el trabajo de Sánchez Salgado (2023), titulado “La pérdida y el desperdicio de alimentos en la cadena de suministro y la importancia de la tecnología para contrarrestarla”. Este estudio profundiza en las causas estructurales del desperdicio alimentario en las distintas etapas de la cadena (producción, almacenamiento, distribución y consumo) y argumenta que la tecnología, particularmente la analítica de datos y el machine learning, constituye una herramienta clave para reducir dichas pérdidas.

Desde una perspectiva económica, el autor señala que los desperdicios representan una pérdida significativa de recursos e ingresos para productores, distribuidores y comerciantes, afectando la competitividad del sector. La implementación de tecnologías predictivas permite anticipar fluctuaciones en la oferta y demanda, optimizar inventarios y evitar sobreproducción o vencimiento de productos perecederos. En cuanto al impacto ambiental, la reducción del

desperdicio alimentario contribuye a disminuir el uso innecesario de agua, energía y transporte, lo que se traduce en una menor huella de carbono.

A nivel social, el estudio destaca la experiencia de la startup argentino Nilus, que aplica modelos de machine learning para conectar empresas con excedentes de alimentos con comedores comunitarios, promoviendo una redistribución eficiente y ética de productos. Esta práctica no solo reduce el desperdicio, sino que mejora la seguridad alimentaria de poblaciones vulnerables, alineándose con objetivos de desarrollo sostenible.

En el contexto de la industria alimentaria colombiana, el análisis de Sánchez Salgado subraya que la adopción de tecnologías inteligentes no debe verse como un lujo, sino como una estrategia necesaria para garantizar eficiencia operativa, sostenibilidad ambiental y justicia social. Su enfoque aporta una mirada crítica sobre la necesidad de políticas públicas, inversión en infraestructura tecnológica y cambios culturales en los actores de la cadena de suministro para que estas soluciones sean realmente efectivas y escalables.

Marco Teórico Aplicado: Machine Learning y Sostenibilidad en la Gestión de la Cadena de Suministro Alimentaria

La creciente exigencia por parte de los mercados internacionales, consumidores y normativas ambientales ha impulsado a la industria alimentaria a adoptar estrategias orientadas hacia la sostenibilidad. Estas estrategias deben equilibrar los impactos económicos, sociales y ambientales en toda la cadena de suministro. En este contexto, las tecnologías emergentes, y en particular el machine learning (ML), se presentan como herramientas clave para responder a estos desafíos.

El machine learning es una subdisciplina de la inteligencia artificial que permite a los sistemas aprender automáticamente a partir de datos y mejorar su rendimiento sin intervención explícita. Su aplicabilidad en las cadenas de suministro incluye la predicción de la demanda, la optimización logística, el control de calidad automatizado, la gestión inteligente de inventarios y la reducción de desperdicio de alimentos (Sánchez Jiménez y García Hernández, 2019; Sánchez Salgado, 2023).

Uno de los marcos teóricos más relevantes para articular la relación entre ML y sostenibilidad es el enfoque de la Triple Cuenta de Resultados (Triple Bottom Line), que considera de manera simultánea los impactos económicos (eficiencia operativa y reducción de costos), ambientales (uso responsable de recursos y reducción de emisiones) y sociales (seguridad alimentaria, condiciones laborales y equidad en el acceso a alimentos) (Gaur et al., 2023; Dhiman et al., 2024).

Estudios recientes han demostrado que modelos como Random Forest, redes neuronales y árboles de decisión han alcanzado precisiones superiores al 90 % en la predicción de la demanda, lo que permite ajustar la producción a las necesidades reales, evitar sobrestock y reducir el

desperdicio de productos perecederos (Niño-Chaparro et al., 2023). Estas técnicas no solo permiten optimizar recursos logísticos, sino que también contribuyen a reducir las emisiones asociadas al transporte y almacenamiento innecesario, alineándose con los principios de la economía circular (Bains et al., 2024).

Asimismo, la trazabilidad mejorada mediante ML y tecnologías afines como blockchain refuerza la transparencia en los procesos productivos y logísticos, lo que permite a las empresas alimentarias cumplir con normativas internacionales como el Pacto Verde Europeo, la Agenda 2030, y la Ley de Responsabilidad Extendida del Productor (EPR), promoviendo así una producción más ética, segura y rastreable (Rodríguez, 2022; Gow, 2020).

Adicionalmente, estudios como el de Sánchez Salgado (2023) muestran cómo la tecnología puede favorecer la redistribución de excedentes alimentarios hacia poblaciones vulnerables, lo que representa una integración efectiva de la sostenibilidad social en la cadena de suministro. Iniciativas como la de la startup Nilus, basada en machine learning, representan modelos replicables en el contexto latinoamericano para reducir la inseguridad alimentaria y la desigualdad de acceso.

Tecnología: Aplicaciones de Machine Learning Como Herramienta Analítica y Predictiva

El machine learning (ML), como rama de la inteligencia artificial, se ha consolidado como una herramienta central para el análisis de grandes volúmenes de datos y la predicción de patrones complejos. En el contexto alimentario, su implementación permite anticipar tendencias de consumo, detectar anomalías en procesos de producción y distribución, y optimizar operaciones críticas mediante el aprendizaje continuo a partir de la experiencia histórica (Sánchez Jiménez y García Hernández, 2019).

Entre las aplicaciones más destacadas de ML en la industria alimentaria se encuentran:

- Predicción de demanda mediante modelos como Random Forest, redes neuronales y árboles de decisión, lo que facilita la planificación de producción y reduce pérdidas por sobreoferta (Niño-Chaparro et al., 2023).
- Clasificación automatizada de alimentos con visión por computador, mejorando el control de calidad y disminuyendo el desperdicio por errores humanos.
- Análisis de riesgo logístico, incluyendo predicción de interrupciones en el transporte o escasez de insumos.
- Redes neuronales para control de condiciones ambientales en almacenamiento y transporte de productos perecederos.
- Estas tecnologías no solo mejoran la eficiencia, sino que se convierten en herramientas clave para la sostenibilidad operativa, al reducir costos, consumo energético y errores en la toma de decisiones.

Gestión de la Cadena de Suministro: Procesos Logísticos, Trazabilidad, Inventario, Transporte y Calidad

La cadena de suministro alimentaria es una red compleja que abarca desde la producción primaria hasta el consumidor final. Esta cadena incluye múltiples actores, procesos y puntos de control, donde la eficiencia logística y la trazabilidad juegan un papel crucial. Según Coronado-Hernández y García-Sabater (2019), los retos principales en la gestión de esta cadena están relacionados con la coordinación, visibilidad y capacidad de reacción ante fluctuaciones de mercado o interrupciones inesperadas.

En este contexto, el ML potencia significativamente la gestión de la cadena en aspectos como:

- Optimización de inventarios: evitando excesos o escasez mediante predicciones más precisas.

- Trazabilidad avanzada: combinada con blockchain, permite monitorear el recorrido del producto desde el origen hasta el destino, garantizando calidad, inocuidad y cumplimiento normativo (Rodríguez, 2022).
- Ruteo inteligente de transporte: que minimiza tiempos, costos y emisiones mediante algoritmos adaptativos.
- Gestión de calidad y caducidad: mediante análisis en tiempo real de condiciones físicas, químicas o ambientales de los productos.

Esta visión integral, apoyada por tecnologías inteligentes, contribuye a construir cadenas de suministro más resilientes, adaptables y alineadas con los objetivos de sostenibilidad y eficiencia.

Sostenibilidad: Enfoque Integral Desde lo Económico, Social y Ambiental

La sostenibilidad en la cadena de suministro alimentaria no puede ser concebida únicamente desde la eficiencia operativa. Requiere de un enfoque triple que considere simultáneamente el impacto económico, ambiental y social, en línea con la Triple Cuenta de Resultados (Elkington, 1997) y los Objetivos de Desarrollo Sostenible (ODS) propuestos por la ONU.

- Impacto económico: El uso de ML permite reducir desperdicios, optimizar la planificación logística y aumentar la rentabilidad de las operaciones. Se traduce en ahorro de costos para productores y distribuidores, así como en mayor competitividad en mercados sostenibles (Gutiérrez Duque y Polo Martínez, 2023).
- Impacto ambiental: La mejora en la eficiencia energética, la planificación de rutas logísticas, la reducción de emisiones por transporte innecesario, y la disminución de desperdicios

alimentarios (Sánchez Salgado, 2023), contribuyen a reducir la huella de carbono del sistema agroalimentario.

- Impacto social: El ML también puede ser utilizado para fortalecer la seguridad alimentaria, garantizar transparencia y ética en la distribución de productos, y conectar excedentes con poblaciones vulnerables, como lo demuestra el caso de Nilus en Argentina. Además, tecnologías predictivas permiten a los gobiernos tomar decisiones más focalizadas y justas en contextos de crisis alimentarias (Niño-Chaparro et al., 2023).

La incorporación del *machine learning* en la gestión de la cadena de suministro alimentaria no representa solo una innovación tecnológica, sino una transformación profunda en la manera en que concebimos la sostenibilidad en el sector agroalimentario. En un contexto donde la presión por reducir el impacto ambiental, mejorar la eficiencia económica y atender las necesidades sociales es cada vez más urgente, el ML emerge como una herramienta que permite tomar decisiones informadas, predictivas y éticamente orientadas. Su integración con principios como la Triple Cuenta de Resultados y marcos regulatorios internacionales convierte a esta tecnología en un puente entre el avance digital y el compromiso con los Objetivos de Desarrollo Sostenible. Sin embargo, más allá de los algoritmos y modelos predictivos, lo verdaderamente transformador radica en cómo estas tecnologías pueden redirigir los flujos de recursos, conocimientos y alimentos hacia un sistema más justo, resiliente y consciente de sus impactos a largo plazo.

Conclusiones

El análisis de la literatura evidencia que la implementación de machine learning en la industria alimentaria no solo representa una innovación técnica, sino una necesidad estratégica para responder a las exigencias de sostenibilidad económica, social y ambiental. Modelos como Random Forest y redes neuronales han demostrado alta eficacia en la predicción de demanda, optimización logística y reducción de desperdicios, lo que permite mejorar significativamente la eficiencia operativa y disminuir los costos asociados a la sobreproducción y distribución ineficiente. Sin embargo, esta transformación digital aún se encuentra limitada por barreras estructurales, como la falta de infraestructura tecnológica y la resistencia al cambio en pequeñas y medianas empresas del sector.

La revisión bibliográfica muestra que las aplicaciones de machine learning pueden tener impactos positivos más allá del plano económico. Casos como el de Nilus, en Argentina, y el modelo aplicado por Niño-Chaparro et al. (2023) en Colombia, revelan que estas tecnologías también pueden ser utilizadas para mejorar la equidad en la distribución de alimentos, reducir la inseguridad alimentaria y focalizar subsidios públicos en poblaciones vulnerables. Asimismo, la trazabilidad mejorada y la reducción de emisiones logísticas refuerzan el cumplimiento de normativas ambientales internacionales y contribuyen al logro de los Objetivos de Desarrollo Sostenible (ODS), especialmente los relacionados con el hambre cero, la producción responsable y la acción climática.

Los estudios revisados coinciden en que la adopción de tecnologías basadas en machine learning debe estar acompañada por una visión sistémica que integre factores tecnológicos, regulatorios, organizacionales y humanos. No basta con disponer de modelos predictivos; es imprescindible que las empresas desarrollen capacidades internas para interpretar los datos,

adaptar sus procesos logísticos y alinear sus objetivos con estándares internacionales de sostenibilidad. Este enfoque sistémico requiere también políticas públicas de apoyo, inversión en formación digital y el fortalecimiento de ecosistemas de innovación que promuevan la colaboración entre gobierno, industria y academia.

Recomendaciones

Se recomienda a las organizaciones del sector alimentario avanzar en la integración gradual de técnicas de machine learning como parte de sus procesos de transformación digital. Esta adopción debe realizarse a partir de un diagnóstico de madurez tecnológica y con el apoyo de equipos multidisciplinarios que articulen conocimiento en ciencia de datos, logística, producción y normativas ambientales. La alineación entre objetivos organizacionales y capacidades analíticas es un requisito fundamental para obtener beneficios sostenibles.

Es recomendable implementar programas de formación interna en analítica de datos y aprendizaje automático. La transferencia de conocimiento técnico al personal operativo facilita la apropiación de los modelos y promueve una cultura organizacional orientada a la toma de decisiones basada en evidencia. Este enfoque contribuye a reducir la dependencia de consultorías externas y a fortalecer la autonomía tecnológica de las empresas.

También se sugiere establecer protocolos para la gobernanza de datos, priorizando la calidad, la trazabilidad y la seguridad de la información recolectada. La efectividad de los modelos de machine learning depende en gran medida de la disponibilidad de datos estructurados, actualizados y representativos. La implementación de sistemas de gestión de datos sólidos facilita el entrenamiento de algoritmos y mejora su capacidad predictiva en entornos reales.

A nivel regulatorio, se recomienda que las empresas anticipen los cambios normativos internacionales relacionados con sostenibilidad y tecnología. La Directiva de Diligencia Debida en Sostenibilidad Corporativa y los Objetivos de Desarrollo Sostenible exigen sistemas de monitoreo, reporte y verificación de impactos sociales y ambientales. La aplicación de modelos

predictivos puede facilitar el cumplimiento de estos requerimientos mediante herramientas de evaluación automatizada y seguimiento continuo.

En última instancia, se sugiere fomentar alianzas con universidades, centros de investigación y proveedores tecnológicos para el desarrollo de proyectos conjuntos. La co-creación de soluciones permite adaptar los modelos a las realidades locales, reducir los costos de implementación y generar conocimiento replicable. Este tipo de colaboración fortalece el ecosistema de innovación y promueve la sostenibilidad del sector alimentario en el mediano y largo plazo.

Referencias Bibliográficas

- Acuña, A. P. (2019). *La gestión de los stakeholders: Análisis de los diferentes modelos*. Encuentro Regional Zona Sur Adenag, 12.
<http://www.fundacionseres.org/Lists/Informes/Attachments/658/La%20gestión%20de%20los%20stakeholders.%20Análisis%20de%20los%20diferentes%20modelos.pdf>
- Al Zaabi, S., Al Dhaheri, N., y Diabat, A. (2020). Analysis of interaction between the barriers for the implementation of sustainable supply chain management. *International Journal of Advanced Manufacturing Technology*, 68(1–4), 895–905. <https://doi.org/10.1007/s00170-013-4951-8>
- Amado Mateus, M. (2024). La dualidad de la inteligencia artificial en la sostenibilidad de las cadenas de suministro: una revisión narrativa. *European Public y Social Innovation Review*, 9, 1–21. <https://doi.org/10.31637/epsir-2024-552>
- Badghish, S., y Soomro, Y. A. (2024). Artificial intelligence adoption by SMEs to achieve sustainable business performance: Application of Technology–Organization–Environment Framework. *Sustainability*, 16(5), Art. 1864.
<https://doi.org/10.3390/su16051864>
- Bains, A., Sridhar, K., Dhull, S. B., Chawla, P., Sharma, M., Sarangi, P. K., y Gupta, V. K. (2024). Circular bioeconomy in carbon footprint components of nonthermal processing technologies towards sustainable food system: A review. *Trends in Food Science y Technology*, 149, 104520. <https://doi.org/10.1016/j.tifs.2024.104520>
- Chen, P., Chu, Z., y Zhao, M. (2016). The road to corporate sustainability: The importance of artificial intelligence. *Technology in Society*, 76, 102440.
<https://doi.org/10.1016/j.techsoc.2023.102440>

- Conrad, R. (2024, 1 de marzo). AI in logistics: Driving sustainability and efficiency in supply chains. *RTS Labs*. <https://rtslabs.com/ai-logistics-sustainability-efficiency/>
- Coronado-Hernández, J. R., y García-Sabater, J. P. (2019). Complejidad en la cadena de suministro: clasificación, fuentes y medición. *Pág*, 38, 25.
- Crawford, K. (2024). Generative AI's environmental costs are soaring – and mostly secret. *Nature*, 626(8000), 693. <https://doi.org/10.1038/d41586-024-00478-x>
- Dhiman, R., Miteff, S., Wang, Y., Ma, S.-C., Amirikas, R., y Fabian, B. (2024). Artificial intelligence and sustainability – A review. *Analytics*, 3(1), 140–164. <https://doi.org/10.3390/analytics3010008>
- Ding, T., Li, J., Shi, X., Li, X., y Chen, Y. (2023). Is artificial intelligence associated with carbon emissions reduction? Case of China. *Resources Policy*, 85, 103892. <https://doi.org/10.1016/j.resourpol.2023.103892>
- FAO. (2019). *El estado de la seguridad alimentaria y la nutrición en el mundo 2019*. Organización de las Naciones Unidas para la Alimentación y la Agricultura. <https://www.fao.org/publications/sofi/2019>
- Fernández, J. L., y Bajo, A. (2018). La teoría del stakeholder o de los grupos de interés, pieza clave de la RSE, del éxito empresarial y de la sostenibilidad. *Adresearch Esic International Journal of Communication Research*, 6(6), 130–143. <https://doi.org/10.7263/adr.rsc.006.07>
- Garzón Zúñiga, M. A., y León Carlosama, R. G. (2022). Gestión sostenible de la cadena de suministro en la industria alimentaria del Ecuador a través de la tecnología blockchain (2019–2021) [Trabajo de investigación, Universidad de las Fuerzas Armadas ESPE].

- Gaur, L., Afaq, A., Arora, G. K., y Khan, N. (2023). Artificial intelligence for carbon emissions using system of systems theory. *Ecological Informatics*, 76, 102165.
<https://doi.org/10.1016/j.ecoinf.2023.102165>
- Gómez, U., y Alberto, J. (2019). Una perspectiva de la innovación tecnológica en Latinoamérica. *Revista de Tecnología y Sociedad*, 11.
- González, E. (2019). La teoría de los stakeholders. Un puente para el desarrollo práctico de la ética empresarial y de la responsabilidad social corporativa.
- Gow, G. (2020, 3 de septiembre). Environmental sustainability and #AI. [Publicación en LinkedIn].
- Gutiérrez, M. Polo, E. (2024) Inteligencia artificial dentro de la cadena de suministros. Fundación Universitaria Del Área Andina. Bogotá.
- Hong, Z., y Xiao, K. (2024). Digital economy structuring for sustainable development: The role of blockchain and artificial intelligence in improving supply chain and reducing negative environmental impacts. *Scientific Reports*, 14(1), 3912. <https://doi.org/10.1038/s41598-024-53760-3>
- Juárez, C. (2024, 17 de mayo). Nueva IA de Walmart reduce desperdicio de perecederos y textiles en anaqueles. *The Logistics World*.
<https://thelogisticsworld.com/tecnologia/nueva-ia-de-walmart-reduce-desperdicio-de-perecederos-y-textiles-en-anaqueles/>
- Melamed-Varela, E., Blanco-Ariza, A. B., y Rodríguez-Calderón, G. (2018). Creación de valor compartido: estado y contribuciones a la sostenibilidad corporativa. *Revista Escuela de Administración de Negocios*, 85, 37–56.
<https://doi.org/10.21158/01208160.n85.2018.2047>

- Meléndez Mogollón, I. C., García Silvera, E. E., Pérez Arias, A., y Camero Lozano, Y. B. (2022). Tendencias del procesamiento de alimentos en el contexto de la COVID-19 y la globalización mundial. *Biotecnología en el Sector Agropecuario y Agroindustrial*, 21(2), 178–192. <https://doi.org/10.18684/rbsaa.v21.n2.2023.2182>
- Morcillo-Bellido, J. (2018). Sostenibilidad en la cadena de suministro: evidencias del sector minorista español. *360: Revista de Ciencias de la Gestión*, 3, 18–38. <https://doi.org/10.18800/360gestion.201803.001>
- Niño-Chaparro, G. E., Niño-Chaparro, A., y Chaparro-Pesca, J. A. (2023). AI in Colombian Food Markets: Using Machine Learning to Address Price Crisis. *Económicas CUC*, 45(1), e24818
- Ochoa-Barragán, R., Serrano-Arévalo, T. L., Pulido-Occegueda, J. C., Cerda-Flores, S. C., Ramírez-Márquez, C., Nápolis-Rivera, F., y Ponce-Ortega, J. M. (2024). Sustainable lime production in Michoacan Mexico: An optimal and equitable approach with machine learning. *Journal of Cleaner Production*, 442, 141017. <https://doi.org/10.1016/j.jclepro.2022.141017>
- Quintero-Campos, L. J. (2019). Aportes teóricos para el estudio de un sistema de innovación. *Innovar*, 20(38), 57–76.
- Rivera, M. Á., y Caballero, R. (2020). Los sistemas de innovación nacional y la teoría del desarrollo. *Problemas del Desarrollo. Revista Latinoamericana de Economía*, 34(134). <https://doi.org/10.22201/iiec.20078951e.2003.134.7498>
- Rodríguez, H. (2022). Trazabilidad y sostenibilidad: el futuro de la cadena de suministro. *Revista Magazine*. <https://revistamagazine.com>

- Sanchez, A. (2021). La pérdida y el desperdicio de alimentos en la cadena de suministro y la importancia de la tecnología para contrarrestarla. Universidad Militar Nueva Granada. Bogotá.
- Sánchez Jiménez, M. E., y García Hernández, D. (2019). Inteligencia artificial aplicada a la industria alimentaria: avances y desafíos. *Revista Colombiana de Tecnología Aplicada*, 12(1), 55–72.
- Singh, R. (2022). Sustainable food supply chain framework in a circular economy. En *Proceedings of the International Conference on Industrial Engineering and Operations Management* (pp. 1234–1245).
- Vorecol. (2023). ¿Cuáles son las mejores prácticas para integrar la sostenibilidad en la cadena de suministro? *Vorecol Blog*. [Cita de estudio de McKinsey y Company].
- Jhonny Amaya. (2025, 25 de mayo). *Fase 3 Socialización Jhonny Amaya* [video]. YouTube. <https://youtu.be/gX0591HOU1M>