

# Horizontes de innovación tecnológica en la industria de alimentos: datos sintéticos y computación cuántica para el futuro inmediato y a largo plazo

Camilo A. Castro 

Grupo de Investigación Aplicada en Lípidos, Team Foods S.A., Calle 45A Sur 56-21, Bogotá D.C., Colombia.

\*camilo.castro@alianzateam.com, castro.lopezcamilo@gmail.com

*Fecha recepción: julio 5 de 2023*

*Fecha aceptación: septiembre 30 de 2023*

## Resumen

El artículo describe a modo de opinión la transformación tecnológica en la industria alimentaria, destacando la influencia positiva de la inteligencia artificial y la generación de datos sintéticos para mejorar la calidad de los productos y optimizar procesos. Se abordan los desafíos de la implementación de la IA, junto con la prometedora aplicación de la computación cuántica en la resolución de problemas complejos. Se presentan ejemplos concretos de la generación de datos sintéticos en la detección de enfermedades en plantas y la predicción de temperatura en la cadena de frío. Se subraya la importancia de estas tecnologías para respaldar la toma de decisiones, mejorar la seguridad y calidad de los alimentos, y se reconoce que la computación cuántica, a pesar de sus desafíos actuales, promete ser una herramienta revolucionaria en la industria alimentaria.

**Palabras clave:** *Inteligencia artificial; Transformación tecnológica; tecnología de alimentos.*

# Technological innovation horizons in the food industry: synthetic data and quantum computing for the near and long-term future

## Abstract

The article provides an opinion-based overview of the technological transformation in the food industry, highlighting the positive influence of artificial intelligence and synthetic data generation in enhancing product quality and process optimization. It addresses the challenges of implementing AI, alongside the promising application of quantum computing in tackling complex problems. Concrete examples of synthetic data generation in disease detection in plants and real-time temperature prediction in the cold chain are presented. The importance of these technologies in supporting decision-making, improving food safety and quality is underscored, with recognition that despite current challenges, quantum computing holds the potential to be a revolutionary tool in the food industry.

**Keywords:** *Artificial intelligence; Technological transformation; Food technology.*

# Horizontes da inovação tecnológica na indústria alimentícia: dados sintéticos e computação quântica para um futuro imediato e a longo prazo

## Resumo

O artigo fornece uma visão baseada na opinião da transformação tecnológica na indústria de alimentos, destacando a influência positiva da inteligência artificial e da geração de dados sintéticos na melhoria da qualidade dos produtos e na otimização dos processos. Aborda os desafios da implementação da IA, juntamente com a promissora aplicação da computação quântica na abordagem de problemas complexos. São apresentados exemplos concretos de geração de dados sintéticos na detecção de doenças em plantas e na previsão em tempo real da temperatura na cadeia de frio. É enfatizada a importância dessas tecnologias no apoio à tomada de decisões, na melhoria da segurança e na qualidade dos alimentos, com o reconhecimento de que, apesar dos desafios atuais, a computação quântica tem o potencial de ser uma ferramenta revolucionária na indústria alimentícia.

**Palavras-chave:** *Inteligência artificial; Transformação tecnológica; Tecnologia de alimentos.*

La industria alimentaria está siendo testigo de una transformación tecnológica sin precedentes, reconfigurando sus operaciones y mejorando la calidad de los productos entregados a los consumidores gracias a la inteligencia artificial (IA) [1]. Desde la optimización de la cadena de suministro para reducir el desperdicio y mejorar la eficiencia de los procesos [2], hasta la detección de problemas potenciales, como la contaminación o el deterioro de los alimentos, para garantizar la salud pública [3]. Esta convergencia de la inteligencia artificial y la innovación en tecnologías de datos está impulsando una transformación profunda y continua en la industria, abriendo el camino hacia un futuro lleno de promesas [4]. No obstante, a medida que avanzamos hacia el futuro, se vislumbra la influencia prometedora de la computación cuántica en esta revolución tecnológica. La computación cuántica presenta virtualmente beneficios en el desarrollo de moléculas para la fijación de nitrógeno [5] y en la identificación de variedades de plantas resistentes al cambio climático [6], planteando un horizonte fascinante para la industria alimentaria, temas que abordaremos en el siguiente escrito.

Aun cuando la IA presenta enormes beneficios, su implementación en la industria alimentaria se enfrenta a desafíos considerables que abarcan dos aspectos críticos: el “problema de los datos” y la calidad de los datos. El “problema de los datos” se refiere a la dificultad de contar con datos suficientes para los modelos de aprendizaje automático, lo que puede surgir cuando los conjuntos de datos disponibles son pequeños y cuando el etiquetado manual se vuelve económicamente inviable [7], un problema que puede ser usual en las compañías productoras. Paralelamente, se encuentra el desafío de garantizar la calidad de los datos, un elemento esencial para la eficacia de los modelos de IA. Esto implica asegurar la precisión, actualización y la ausencia de sesgos en los datos a través de medidas de control de calidad [8]. Para abordar ambos desafíos, se consideran viables nuevas estrategias como la generación de datos sintéticos.

En el contexto de la inteligencia artificial aplicada a la industria de alimentos, la generación de datos sintéticos se presenta como una opción para superar la escasez de datos reales. Esta metodología implica la creación de conjuntos de datos que reproducen con precisión las

características de los datos reales, ampliando significativamente la disponibilidad de datos para entrenamiento y mejorando el rendimiento de los modelos de IA. Además de abordar la cantidad de datos, esta estrategia permite construir datos bajo estándares normalizados de calidad, reduciendo el sesgo y proporcionando una base de entrenamiento equitativa para los modelos. Más aún, la data sintética permite probar hipótesis y comparar métodos y algoritmos en un entorno controlado al generar datos con propiedades y distribuciones conocidas, lo que facilita evaluar el rendimiento de diferentes modelos y algoritmos en escenarios simulados [7]. Entonces, la generación de datos sintéticos es una opción para que las industrias no se queden rezagadas en la incorporación de la inteligencia artificial.

Para comprender la versatilidad y el alcance de la generación de datos sintéticos, examinaremos dos casos de la industria alimentaria. En el primero, en el ámbito agrícola, donde la detección temprana de enfermedades es esencial para proteger la economía de los agricultores, se presenta un caso de generación de datos sintéticos que busca superar las limitaciones de la disponibilidad de imágenes reales de un vegetal en la detección visual de daños. Esta técnica implica el uso de algunas imágenes reales para generar nuevas incluyendo el uso de Redes Generativas Adversarias (RGA). Por otro lado, en un segundo caso relacionado con la predicción en tiempo real de la temperatura en una cadena de frío de frutas, se ha utilizado la generación de datos sintéticos como recurso para entrenar una red neuronal del tipo Perceptrón Multicapa (PM). En este tipo de problemas, los datos sintéticos, generados a partir de modelos térmicos, son abundantes, pero tienden a generar sobreajuste y sobrevaloración del rendimiento de predicción. En este trabajo se comparan modelos entrenados con ambos tipos de datos, simulados y reales, para evaluar la eficiencia de la predicción. Los casos escogidos nos permitirán analizar tanto en la detección de enfermedades agrícolas como en la predicción de la temperatura en la cadena de frío, que la generación de datos sintéticos es una opción para respaldar la toma de decisiones.

En el primer caso, se aborda el desafío de la detección temprana de enfermedades en las plantas de tomate, un problema que enfrenta limitaciones en la disponibilidad de datos [9]. Para abordar este desafío, los autores proponen el uso

de una RGA que genera imágenes sintéticas de hojas de plantas de tomate. Estos datos sintéticos se emplean posteriormente en el entrenamiento de un modelo de aprendizaje profundo capaz de identificar enfermedades en las hojas. De manera complementaria, se aplican técnicas de aumento de datos como rotación, desplazamiento, volteo, zoom, escalado y normalización, para generar un conjunto más amplio de imágenes, lo que mejora el rendimiento del modelo en términos de precisión y generalización.

Un Modelo de Aprendizaje Profundo es una forma avanzada de inteligencia artificial que se inspira en la estructura y el funcionamiento del cerebro humano. Funciona mediante la creación de múltiples capas de procesamiento de datos, imitando la forma en que las neuronas procesan la información. Este modelo se somete a un proceso de entrenamiento intensivo, durante el cual analiza un extenso conjunto de imágenes de hojas de tomate para identificar patrones y características que son indicativas de enfermedades. Sin embargo, como se describió arriba una de las limitaciones clave que se enfrenta en este proceso es la disponibilidad limitada de datos reales. Para superar este obstáculo, se introduce la técnica llamada Redes Generativas Adversarias. Estas redes consisten en dos componentes: un generador y un discriminador. El generador crea imágenes sintéticas que imitan con precisión las características de las hojas de tomate, mientras que el discriminador evalúa cuán auténticas son estas imágenes en comparación con las reales. Este proceso iterativo entre el generador y el discriminador permite al Modelo de Aprendizaje Profundo mejorar su capacidad para generar datos sintéticos de alta calidad.

Los resultados del estudio presentan hallazgos significativos. En primer lugar, se demuestra que la combinación del enfoque de generación de datos sintéticos y el aprendizaje profundo tiene el potencial de mejorar sustancialmente la precisión en la detección de enfermedades en hojas de tomate. Específicamente, el modelo propuesto por Padmanabhuni & Gera [9] logra una destacada precisión del 98,5 % en la identificación de enfermedades en hojas de tomate, superando los resultados previamente documentados en la literatura científica. Además, se destaca que la aplicación de técnicas de aumento de datos ofrece mejoras adicionales en el rendimiento del modelo.

Por otro lado, en el segundo caso abordamos el desafío de la predicción en tiempo real de la temperatura en equipos de refrigeración a lo largo de una cadena de frío de frutas, con el objetivo de garantizar la calidad y seguridad de los alimentos y reducir el desperdicio [10]. El problema que se aborda es crucial, ya que la temperatura juega un papel fundamental en la conservación de manzanas específicamente. El estudio compara el uso de conjuntos de datos sintéticos y experimentales para entrenar modelos de *machine learning* y proporciona información valiosa sobre la estrategia óptima para mejorar el rendimiento en la predicción de la temperatura en tiempo real en la cadena de frío.

El estudio se llevó a cabo comparando dos estrategias en un arreglo de pallet de un nivel con 8 cajas organizadas en una disposición de 2x4, cada caja contiene 28 manzanas. La primera estrategia implicó la generación de datos sintéticos utilizando un modelo térmico basado en ecuaciones de balance de calor, que consideró propiedades térmicas de materiales, geometría del pallet, productos y condiciones ambientales. Estos datos se emplearon para entrenar modelos de *machine learning* y predecir la temperatura en tiempo real en distintas condiciones de transporte y almacenamiento de alimentos. La segunda estrategia involucra la generación de datos experimentales en una sala de pruebas de temperatura controlada, utilizando sensores inalámbricos para medir la temperatura en puntos específicos del pallet a intervalos regulares. Estos datos se utilizaron para entrenar un modelo de *machine learning* con el mismo propósito.

El modelo de red neuronal utilizado en el estudio de Loisel *et al.* [10] es el Perceptrón Multicapa, una estructura compuesta por múltiples capas de neuronas interconectadas. Las neuronas se activan bajo el resultado de una función de activación, lo que le permite capturar relaciones y patrones complejos en los datos. En particular, en este trabajo se ha empleado la función de activación Rectified Linear Unit (RELU) debido a su eficacia en la aproximación de funciones no lineales. Utilizando esta red neuronal se realiza el proceso de entrenamiento, se aplican el algoritmo de retro propagación y el método del descenso de gradiente estocástico para ajustar iterativamente los pesos y sesgos de la red, minimizando así el error cuadrado entre las predicciones del modelo

y los valores reales de temperatura. Además, se exploran diversas configuraciones de capas ocultas, desde una sola hasta tres capas, con diferentes números de neuronas en las capas, esto con el propósito de evaluar cómo la variación en la estructura de la red afecta la capacidad de predicción de la temperatura. Esta flexibilidad en la arquitectura de la red permite adaptarla a la complejidad de los datos y optimizar su rendimiento en la predicción de la temperatura en tiempo real a lo largo de la cadena de frío.

Aunque los autores encontraron que las redes neuronales entrenadas con datos experimentales superaron a aquellas entrenadas con datos sintéticos, con mejoras de rendimiento en el rango del 20 al 40 %, bien se puede utilizar la información derivada de los datos sintéticos para complementar el conjunto de datos experimental. También destacaron que la precisión de los modelos aumenta con el tiempo de medición, aunque con cierta incertidumbre en las predicciones durante las primeras horas. En última instancia, la implementación de esta metodología en la cadena de frío se vislumbra como una forma de preservar la calidad, al mismo tiempo que se reduce el desperdicio y se permite la predicción en tiempo real de las temperaturas de los productos mediante el uso de sensores inalámbricos en los pallets. Esto sugiere que, si bien los datos sintéticos no superaron a los datos experimentales en este estudio, podrían ser útiles en la mejora continua de modelos y algoritmos para aplicaciones futuras en la cadena de frío.

Ejemplificado a través de dos casos en la agricultura y la cadena de frío, se demuestra que esta estrategia puede significativamente mejorar la precisión y generalización de los modelos. Aunque en uno de los casos los datos experimentales superaron a los sintéticos, la combinación de ambos se revela como una valiosa herramienta para aplicaciones futuras en la industria alimentaria, respaldando la toma de decisiones, la seguridad y la calidad de los productos.

A pesar de la versatilidad que ofrece la inteligencia artificial, es importante reconocer que los procesos de entrenamiento y desarrollo de aplicaciones basadas en IA a menudo requieren un tiempo significativo de ejecución. Además, se enfrentan desafíos inherentes en la computación clásica, Wisnosky & Asaithambi

[11] explican tres retos actuales, el primero la alta complejidad computacional, que implica que a medida que el número de entradas aumenta, la cantidad de cálculos que deben realizarse crece rápidamente, y en casos donde esto ocurre de manera exponencial, las computadoras clásicas no pueden resolver eficazmente estos problemas. El segundo es la generación de números pseudoaleatorios ya que las computadoras, debido a su naturaleza determinista, generan números que son pseudoaleatorios y no verdaderamente aleatorios. Y por último, la falta de reversibilidad en las operaciones de computadoras clásicas dificulta la recuperación de las entradas originales a partir de los resultados, lo que presenta un obstáculo en la resolución de problemas específicos, como en el caso de las compuertas lógicas.

La computación cuántica, en principio, ofrece soluciones a los desafíos de la computación clásica. Una computadora cuántica se compone principalmente de tres elementos fundamentales: la memoria cuántica, la unidad de procesamiento cuántico y un módulo de entrada/salida. La memoria cuántica almacena qubits, que son análogos a los bits en la computación clásica, pero pueden existir en estados de superposición y entrelazamiento, lo que significa que pueden coexistir con diferentes probabilidades, aumentando así las posibilidades de definición de estado [12]. Gracias a la superposición cuántica, la capacidad para almacenar y procesar información crece de manera exponencial, lo que permite resolver problemas complejos de manera eficiente. Además, la naturaleza impredecible de las partículas cuánticas habilita la generación de números verdaderamente aleatorios, superando la limitación de los números pseudoaleatorios en las computadoras clásicas. Por último, la reversibilidad inherente a la computación cuántica ofrece la posibilidad de recuperar las entradas originales a partir de los resultados, una característica que falta en las operaciones clásicas [11]. Entonces, dada la capacidad para realizar cálculos complejos de manera más eficiente, la computación cuántica puede impactar positivamente reduciendo la energía requerida para resolver problemas específicos. Además, su capacidad para generar números aleatorios verdaderamente aleatorios también puede mejorar la seguridad de los sistemas de cifrado lo que podría contribuir a una mayor seguridad en la protección de data sensible. Sin embargo, a pesar de que la computación

cuántica ofrece un gran potencial, su adopción generalizada se ve restringida en la actualidad debido a su complejidad y costos elevados. No obstante, herramientas como IBM-Qiskit están contribuyendo a ampliar el acceso a esta tecnología como lo demuestran Faro & Jhonson [13], aunque su uso todavía plantea desafíos en términos de comprensión y aplicación.

A pesar de la creciente investigación en el desarrollo de algoritmos, actualmente no existen aplicaciones de computación cuántica disponibles en línea para la industria de alimentos que puedan ser estudiadas. Sin embargo, se ha explorado cómo la computación cuántica podría abordar desafíos en la producción de fertilizantes. Uno de los casos de estudio se centra en la molécula FeMoco, que desempeña un papel crucial en la conversión del nitrógeno del aire en amoníaco, esencial para la producción de fertilizantes. Esta investigación podría tener un impacto significativo en la seguridad alimentaria y la crisis climática al mejorar la eficiencia en este proceso que actualmente consume el 2 % de la energía mundial [5]. En otros ámbitos de la agricultura persisten desafíos especialmente en aquellas regiones afectadas por condiciones climáticas extremas y desastres naturales. Sin embargo, la computación cuántica tiene el potencial de desencadenar una auténtica revolución en la genética de las plantas al permitir la identificación directa de los genes responsables de rasgos altamente deseables, como la resistencia a la sequía, prescindiendo de la necesidad de cruzar diferentes variedades y llevar a cabo largos ensayos [6]. Aunque el alcance completo de la computación cuántica en la agricultura aún es incierto, se anticipa que los avances en esta tecnología aportarán beneficios sustanciales a la industria y contribuirán de manera crucial a abordar los desafíos continuos en la seguridad alimentaria.

En conclusión, la generación de datos sintéticos se considera una estrategia en evaluación para mitigar las limitaciones de datos reales en la industria alimentaria, mejorando la eficacia de los modelos de inteligencia artificial y respaldando la toma de decisiones fundamentales. No obstante, a medida que avanzamos hacia el futuro, la computación cuántica se perfila como una herramienta disruptiva que podría resolver de manera eficiente problemas complejos relacionados con la optimización de procesos y la resolución de

desafíos críticos en la producción de alimentos, gracias a su capacidad para procesar información de manera exponencialmente más rápida, aunque su adopción sigue siendo un reto pendiente en términos de costos y complejidad tecnológica. Estas dos vertientes tecnológicas prometen un futuro fascinante y lleno de oportunidades para la industria alimentaria.

## Referencias

- [1] Monteiro J, Barata J. Artificial Intelligence in Extended Agri-Food Supply Chain: A Short Review Based on Bibliometric Analysis. *Procedia Computer Science*. 2021;192:3020-3029. doi.org/10.1016/j.procs.2021.09.074
- [2] Ramirez E, Vilchez J, Thakar C, Phasinam K, Kassanuk T, Naved M. A review on role of artificial intelligence in food processing and manufacturing industry. *Materials Today: Proceedings*. 2022;51:2462-2465. doi.org/10.1016/j.matpr.2021.11.616
- [3] Kudashkina K, Corradini MG, Thirunathan P, Yada RY, Fraser E. Artificial Intelligence technology in food safety: A behavioral approach. *Trends in Food Science & Technology*. 2022;123:376-381. doi.org/10.1016/j.tifs.2022.03.021
- [4] Khan R. Artificial Intelligence and *Machine Learning* in Food Industries: A Study. *Journal of Food Chemistry & Nanotechnology*. 2022;7:60-67.
- [5] Webber M, Elfving V, Weidt S, Hensinger WK. The Impact of Hardware Specifications on Reaching Quantum Advantage in the Fault Tolerant Regime. *AVS Quantum Computing*. 2022;4:1-22. doi.org/10.1116/5.0073075
- [6] Byrum J. Quantum computing may help to solve the global food security problem (sitio en internet). *AgFunder News*. Disponible en <https://agfundernews.com/quantum-computing-answer-to-the-global-food-security-problem>. Acceso el 30 de septiembre de 2023.
- [7] Nikolenko SI. *Synthetic Data for Deep Learning*. Springer International Publishing; 2021. doi.org/10.1007/978-3-030-75178-4
- [8] Aldoseri A, Al-Khalifa KN, Hamouda AM. (2023). Re-Thinking Data Strategy and Integration for Artificial Intelligence: Concepts, Opportunities, and Challenges. *Appl. Sci*. 2023;13:7082-7115. doi.org/10.20944/preprints202305.1565.v1
- [9] Padmanabhuni SS, Gera P. Synthetic Data Augmentation of Tomato Plant Leaf using Meta Intelligent Generative Adversarial Network:

- Milgan. International Journal of Advanced Computer Science and Applications, 2022;13:230-238. [doi.org/10.14569/IJACSA.2022.0130628](https://doi.org/10.14569/IJACSA.2022.0130628)
- [10] Loisel J, Cornuéjols A, Laguerreb O, Tardet M, Cagnon D, Duchesne O, *et al.* *Machine learning* for temperature prediction in food pallet along a cold chain: Comparison between synthetic and experimental training dataset. *Journal of Food Engineering*. 2022;335:111156-111167. [doi.org/10.1016/j.jfoodeng.2022.111156](https://doi.org/10.1016/j.jfoodeng.2022.111156)
- [11] Wisnosky D, Asaithambi A. Solving Classical Computing Problems Via Quantum Computing – SOARS 2021. SOARS Virtual Conference. Disponible en: <https://unfsoars.domains.unf.edu/2021/posters/solving-classical-computing-problems-via-quantum-computing/>. Acceso el 30 de septiembre de 2023.
- [12] Ajagekar, A. New frontiers of quantum computing in chemical engineering. *Korean Journal of Chemical Engineering*. 2022;39:811-820. [doi.org/10.1007/s11814-021-1027-6](https://doi.org/10.1007/s11814-021-1027-6)
- [13] Faro I, Johnson B. IBM Quantum delivers 120x speedup of quantum workloads with Qiskit Runtime. IBM.com. Disponible en <https://research.ibm.com/blog/120x-quantum-speedup?lnk=ushpv18re2>. Acceso el 30 de septiembre de 2023.