

Ensayo de investigación

Plataforma de industria alimentaria 4.0 usando aplicación web y modelos de ML/DL para clasificación de calidad de dulces Food industry platform 4.0 using Web application and ML/DL models for candy quality classification

Ignacio Alredo Badillo¹, Daniel Pacheco Bautista^{2*}, Francisco López Huerta³
Juan Crescenciano Cruz Victoria⁴, Ángel Vergara Betancourt⁵

¹CONAHCyT-INAOE

²Universidad del Istmo, Campus Tehuantepec

³Universidad Veracruzana

⁴Universidad Politécnica de Tlaxcala

⁵TecNM / ITS Zacapoaxtla

*Autor de correspondencia:

dpachecob@hotmail.com

Recibido:03-03-2024 Aceptado: 18-12-2025 (Artículo Arbitrado)

Resumen

En la actualidad, hay un importante incremento de plataformas que proporcionan funciones de Industria 4.0, como consecuencia de la digitalización, conectividad, automatización avanzada, inteligencia artificial, fabricación aditiva, ciberseguridad industrial y búsqueda de sostenibilidad, impulsando una transformación integral en la fabricación y la cadena de suministro. Uno de los grandes problemas que se presentan es al momento del diseño e integración de estas funciones, especialmente cuando se agregan operaciones de modelos generados por el aprendizaje máquina (ML, Machine Learning) y el aprendizaje profundo (DL, Deep Learning). En este trabajo se presenta una plataforma de Industria 4.0 usando una aplicación web y modelos de ML/DL para la evaluación y clasificación de la calidad de dulces, basado en el análisis del vector de características. Esta plataforma tiene el conocimiento de las recetas de los dulces mediante el modelado, lo que le permite predecir/clasificar la calidad de éstos a través del ingreso de las variables actuales del proceso de producción.

Palabras clave: Aprendizaje profundo, Aprendizaje máquina, Industria 4.0, Calidad de dulces.

Abstract

Currently, there is a significant increase in platforms that provide Industry 4.0 functions, as a result of digitalization, connectivity, advanced automation, artificial intelligence, additive manufacturing, industrial cybersecurity and the pursuit of sustainability, driving a comprehensive transformation in manufacturing and the supply chain. One of the major challenges arises during the design and integration of these functions, especially when incorporating operations from machine learning (ML) and deep learning (DL) models. This paper presents an Industry 4.0 platform using a web application and ML/DL models for the evaluation and classification of candy quality, based on feature vector analysis. This platform has knowledge of candy recipes through modeling, which allows it to predict/classify the quality of these by inputting the current variables of the production process.

Keywords: Deep learning, Machine learning, Industry 4.0, Candy quality.

Introducción

El diseño e implementación de plataformas utilizadas en la Industria 4.0, son un reto en la actualidad por la gran cantidad de conceptos y tecnologías que pueden integrarse, entre los más importantes se encuentran: a) sofisticados modelos de Inteligencia Artificial (IA) usando principalmente Machine Learning (ML) y Deep Learning (DL), b) Internet de las cosas

(IoT), una tecnología que permite la recopilación y análisis de datos masivos en forma muy eficiente, c) Tecnología 5G, para el despliegue de conexiones con velocidades más rápidas y mayor capacidad para respaldar aplicaciones como el IoT y la realidad aumentada, d) *Blockchain* y Criptomonedas, que ofrecen servicios de contratos inteligentes y seguimiento de

la cadena de suministro, e) Computación Cuántica, una tecnología que promete proporcionar una capacidad potencial para resolver problemas complejos de manera más eficiente que las computadoras clásicas y f) Ciberseguridad, cuya importancia se debe al aumento constante de las amenazas ciberneticas, lo que exige actualizar u ofrecer nuevos servicios para proteger sistemas y datos.

En particular, la IA ha evolucionado en forma apresurada en los últimos años, teniendo una serie de ramificaciones como la Visión Artificial, Sistemas Expertos, Procesamiento del Lenguaje Natural, Aprendizaje Máquina, Robótica, entre otras. Los modelos de ML son parte de la IA, mientras que los modelos de DL son un subconjunto de ML donde las redes neuronales profundas se utilizan para tareas como el reconocimiento de imágenes, procesamiento de lenguaje natural y juegos. Recientemente, éstas han ganado amplia popularidad, impulsando el desarrollo de distintos modelos con arquitecturas y configuraciones representativas. El ML tiene una clasificación de técnicas en función de la forma en que sus modelos son entrenados, como puede ser: el Aprendizaje Supervisado, Aprendizaje no Supervisado, Aprendizaje por Reforzamiento, Aprendizaje Semi-Supervisado y Aprendizaje Profundo. Debido a la naturaleza de este trabajo, enfocado a la clasificación de la calidad de un producto, el aprendizaje supervisado es de gran interés ya que permite dos estrategias fundamentales: la Regresión, en la que el modelo predice un valor continuo como la temperatura o el precio de una acción, y la Clasificación, donde el modelo responde a una etiqueta, clase, grupo o categoría, como puede ser “alta calidad”, “mediana calidad” o “baja calidad”.

Hay una gran cantidad de técnicas para desarrollar modelos de ML/DL, entre las más representativas se pueden listar las siguientes: Regresión Lineal, Regresión Logística (RL), Máquinas de Soporte Vectorial (SVM), Árboles de Decisión, Random Forest, K Vecinos Más Cercanos (K-NN), Redes Neuronales Artificiales y Naive Bayes. También existen diversas tecnologías web que pueden utilizarse en el desarrollo de plataformas para la Industria 4.0, cuyo auge ha hecho que constantemente haya nuevos desarrollos y otros sean obsoletos. Las principales tecnologías web abarcan un conjunto diverso de herramientas, entre las más importantes se encuentran: HTML (Hypertext

Markup Language), CSS (Cascading Style Sheets), JavaScript, React, Angular, Vue.js, Node.js, Express.js, Django, RESTful APIGraphQ, MySQL y MongoDB.

Las plataformas de Industria 4.0 se han implementado en diversas áreas de trabajo, creando una interacción entre dispositivos conectados al internet e incluyendo diversos modelos de IA con ML y DL (ver la Figura 1). Por ejemplo, Sánchez-Torres et al. (2022), utilizaron ML para la toma de decisiones, generando grupos y sugiriendo productos candidatos que pueden ser sustituidos por otros con mejor costo-beneficio. Candanedo et al. (2018), usaron los sistemas de acondicionamiento para controlar el clima interior de un proceso: temperatura del aire, humedad y presión, etc. Por su parte, Balfaqih (2023) analizó las necesidades del mercado para la implementación de sistemas de gran escala para logística inteligente y su impacto en empresas, industrias y compañías. Mientras que Soumyashree et al. (2023) emplearon modelos de árboles de decisión para monitorear el uso de energía de máquinas y electrodomésticos, prediciendo de este modo su eficiencia.

En particular, en el control de la calidad de productos del sector alimentario también pueden encontrarse una diversidad de aportaciones: Verma et al. (2021) presentaron una plataforma que utiliza diferentes tecnologías web como *React Js*, el framework *Django*, *Scikit-learn* y bases de datos, los autores utilizaron mo-



Figura 1. Interacción figurada en la Industria 4.0, considerando IA, robótica, IoT, aplicaciones, entre otros componentes.

Fuente: Generada en www.fotor.com.

delos de PLN (Procesamiento de Lenguaje Natural) para hacer órdenes de comida de manera automática sin contactar con otro ser humano. Otra plataforma de la industria alimentaria muestra a una aplicación web enfocada a la nutrición y asistencia para dietas usando DL (modelos VGG y ResNet) para identificar y recomendar la dieta y valorizar su valor nutricional (Buvaneswari et al., 2023). Los autores trabajaron con herramientas como Javascript, Tensorflow, Keras y MySQL. Sahoo et al. (2019), presentaron una plataforma para el reconocimiento de imágenes de alimentos mediante DL (redes convolucionales) para el registro inteligente de alimentos e identificaron 752 tipos de éstos, usando tecnologías como Caffe, Flask, MongoDB, Javascript, entre otras. Ferrini et al. (2023), describieron una plataforma para predecir la probabilidad de un trastorno del SNC (Sistema Nervioso Central) de origen infeccioso o inflamatorio en ganado con discapacidad neurológica, utilizando varias técnicas de ML como LR, SVM, KNN, MLP (*Multi-layer Perceptron*), GB (*Gradient Boosting*) y RF, junto con tecnologías como Streamlit y Python para la creación de la plataforma web. Kruthika et al. (2023), reportaron una plataforma web para optimizar inventarios, proporcionar predicción de ventas y permitir la donación de alimentos por parte de restaurantes, realizando una predicción de las exigencias de alimentos utilizando el ML (*XGBoost*, *Gradient Boosting Regressor* y RF) y análisis mundial del desecho de alimentos, operando con tecnologías web como Javascript, PHP y MySQL. Fracarolli et al. (2020) presentan una recopilación de trabajos donde se combina el DL con visión artificial para la industria alimentaria y la calidad de sus productos, en ésta se describen aplicaciones como la inspección, monitoreo de procesamiento, detección de objetos extraños, calificación de la calidad del alimento, detección de ubicaciones de puntos defectuosos, etc. Hassoun et al. (2023) informaron de la creciente digitalización y automatización del análisis de alimentos utilizando tecnologías avanzadas en la industria alimentaria, tales como técnicas no destructivas de dactiloscopya, herramientas bioinformáticas, IA y Big Data, para revolucionar la calidad de los alimentos. El trabajo de Cubillas et al. (2022), describió un enfoque de calidad para la predicción temprana del rendimiento de los cultivos utilizando ML (Modelos Generalizados y SVM con Kernel lineal y Gaussiano), y una plataforma

web en la nube aplicado a un caso de estudio de un olivar y un conjunto de datos recopilados durante 8 años. En Nordin et al. (2019), los autores presentaron el reconocimiento de imágenes de alimentos para el cálculo de precios justos mediante una Red Neuronal Convolutacional (CNN). Por su parte, Calle et al. (2023) generaron un modelo para detectar y cuantificar adulteraciones en miel de alta calidad utilizando Vis-NIR (espectroscopia visible del infrarrojo cercano) en combinación con aprendizaje automático (SVM y RF). Finalmente, entre muchos otros trabajos, Jagadeesh et al. (2022) definieron una clasificación de frutas para identificar adulteración usando forma y complejión, empleando enfoques de DL como CNN.

En este mismo sentido, el presente trabajo se enfoca en la resolución de un problema real dentro de una empresa reconocida en México dedicada a la producción de dulces, donde actualmente se realizan pruebas para el desarrollo de un nuevo producto. Durante este proceso, se han registrado las variables de entrada (vector de características) junto con la calificación correspondiente de la salida. A partir de esta información, la plataforma propuesta se orienta al almacenamiento y gestión de los datos de entrada, así como al desarrollo de un modelo de clasificación basado en técnicas de ML y DL, con el objetivo de determinar la calidad del dulce producido. Dicho modelo permite la clasificación del producto y el ajuste de parámetros en función de sus salidas, abstrayendo el conocimiento obtenido y facilitando una toma de decisiones adecuada.

Materiales y método

La plataforma propuesta de Industria 4.0 usando una aplicación web y modelos de ML y DL para la evaluación de la calidad de los dulces, combina las tecnologías HTML, CSS, Javascript, PHP, MySQL y bibliotecas de Python para crear, almacenar y utilizar los modelos de ML y DL.

El diseño del sistema se fundamenta en dos servicios en la nube: para el almacenamiento de la base de datos y para el hospedaje de la aplicación web que da soporte a la plataforma de Industria 4.0. Estos elementos permiten subir y almacenar los datos de las diferentes fuentes en la base de datos. Adicionalmente, la aplicación web en el servidor permite funciones corporativas para actividades específicas para

cada tipo de usuario. Los datos almacenados son filtrados y procesados por este grupo específico interno del corporativo, además, se pueden visualizar tablas, monitorear el uso de la aplicación y generar reportes con diversos filtros. El dashboard y el sistema de reportes incluidos permiten una presentación visual y analítica de los datos almacenados.

Asimismo, se implementa un esquema de seguridad para proporcionar: i) confidencialidad, para que los datos almacenados en la base de datos no sean guardados de manera plana, sino cifrados, ii) integridad, para asegurar que los datos en la base de datos sean exactamente los que se almacenaron, iii) control de accesos, para permitir acceder solo a quienes deben ingresar al sistema, iv) perfiles, para proporcionar diferentes alcances de los recursos en el sistema, v) sesiones, para conocer los ingresos y la navegación de los usuarios dentro del sistema, vi) trazabilidad, para analizar cómo han sido capturados los datos y dar seguimiento a los reportes, y vii) inmutabilidad, para que a través de hash lineal se pueda conocer el almacenamiento de los datos con respecto a estampas de tiempo y encadenamiento de bloques de los diferentes grupos de registro.

De manera general, la metodología utilizada considera los siguientes pasos:

- Análisis de requerimientos específicos y adquisición de los datos.
- Diseño e implementación de los diferentes contextos de pruebas, tanto unitarias como integrales.

c. Evaluación, investigación, implementación y validación de los modelos de ciencia de datos usando ML/DL.

d. Diseño, desarrollo, implementación y pruebas de la base de datos del sistema.

e. Diseño y validación de las interfaces de usuario y pruebas y análisis conforme a la experiencia de usuario.

f. Diseño, desarrollo, implementación, pruebas unitarias y pruebas integrales del sistema integral de calidad mediante dashboard para el mapeo y visualización.

f. Documentación de manuales, pruebas y reporte del sistema, así como la capacitación en el uso de éste.

De manera específica, el inciso c requiere de una serie de pasos: i) obtención de los datos, ii) selección de características, iii) pre-procesamiento de los datos, iv) selección de un representante de cada una de las familias de ML/DL, v) ejecución de los modelos en un proceso de validación cruzada 10 por 1, vi) selección del mejor de acuerdo con un conjunto de métricas: precisión y exactitud, vii) entrenamiento del modelo seleccionado, variando hiper-parámetros y viii) validación del modelo final.

La plataforma está diseñada para cuatro tipos de usuarios: administradores, cocineros, ingenieros de calidad y supervisores (ver Figura 2); todos ellos consumen los recursos computacionales de los servicios en la nube para la aplicación web y la base de datos. Los administradores tienen el control de todos

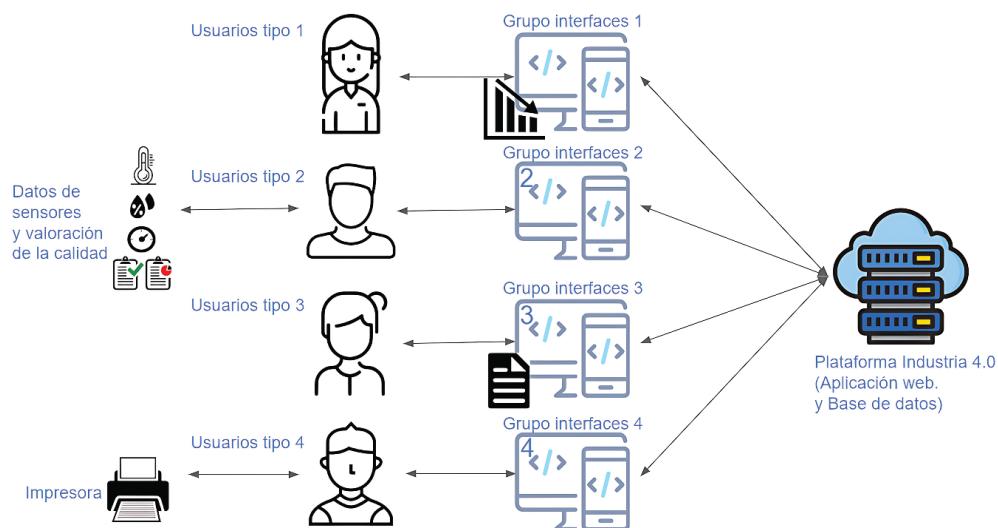


Figura 2. Actores y actividades con interfaces y la plataforma de Industria 4.0.
Fuente: Elaboración propia.

los procesos, aunque su actividad está limitada a conocer los estados de los procesos de registros, dashboard, gráficas y reportes. Los cocineros serán los encargados de ingresar los registros con los datos de entrada (vectores de características) y dar la calificación del dulce producido, así como también ellos podrán utilizar el modelo de ciencia de datos para hacer pronósticos para la clasificación del producto. Los ingenieros de calidad serán los encargados de revisar los reportes y rendimientos de lo producido, mientras que los supervisores dan el visto bueno y confirman los procesos capturados.

Aunque se tienen una gran cantidad de interfaces que incluyen varios catálogos (usuarios, personal, clientes, máquinas, productos, etc.), destacan siete: a) dashboard, b) registro de control procesos, c) control diario de temperatura y humedad, d) eventos, e) pronóstico de calidad, f) central de control procesos y g) central de control diario. En la Figura 3 se muestran dos de estas interfaces: (a) el dashboard que reporta la producción registrada por día y por turno, así como la temperatura y humedad en un lapso de tiempo y (b) una interfaz para el registro de los eventos, donde se presenta el registro de Control Procesos indicando un conjunto de métricas para almacenarlos en la

base de datos. El registro de Control Diario de temperatura y humedad es muy similar a los mostrados en la Figura 3.

Cada registro se almacena y puede ser consultado para conocer su estatus dentro de una interfaz que concentra y agrupa esos registros. En la Figura 4(a) se muestra la central de reportes para el Control Procesos, y en la 4(b) se muestra el registro de eventos del sistema, donde se conoce qué procesos, interacciones o registros se han hecho con estampa de tiempo y usuario.

Por otro lado, cada uno de los registros debe ser manejado y presentado de manera visual y digital mediante archivos PDF. En la Figura 5 se presentan las vistas de estos documentos para (a) el reporte de Control Procesos y (b) el reporte de Control Diario.

Finalmente, se tiene el modelo de IA, el cual se ha generado usando bibliotecas en Python para encapsularlo y pueda ser ejecutado en la plataforma web, de tal manera que se solicitan datos en una interfaz y basado en el vector de características que se forma a partir de 24 variables de entrada, se produzca la clasificación, además de un código de colores y una imagen representativa.

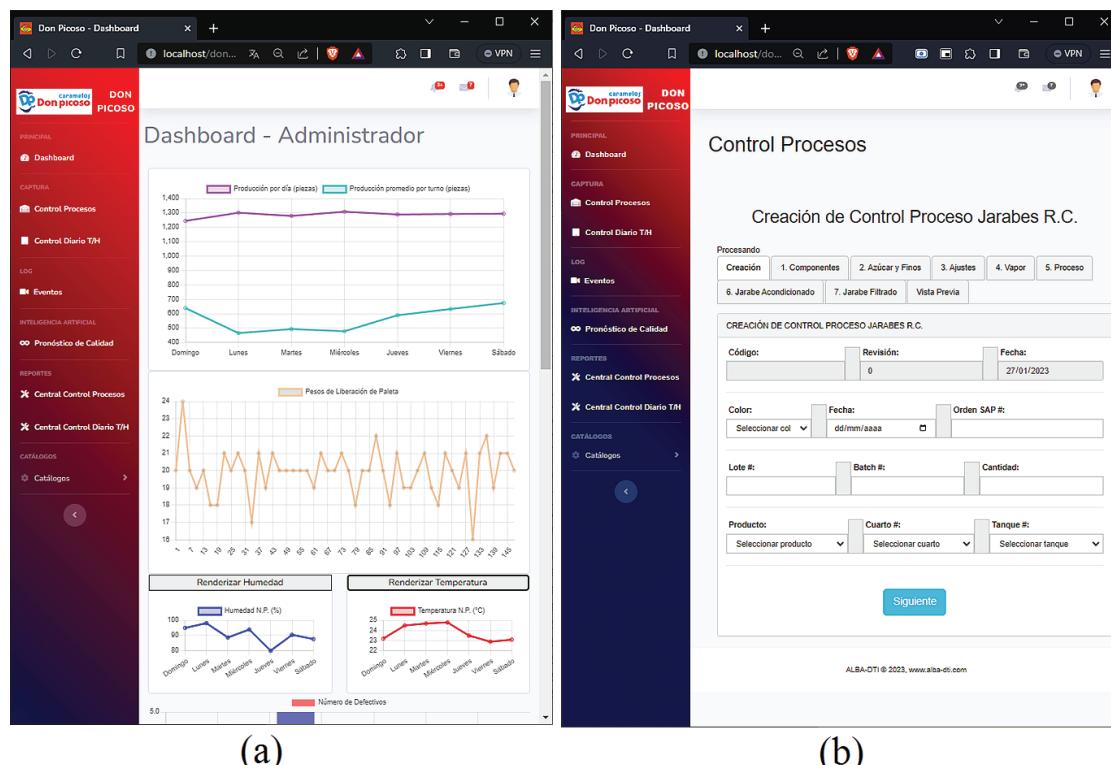


Figura 3. Interfaces: (a) dashboard y (b) registro de Control Procesos.
Fuente: Elaboración propia.

(a) Central de Reportes de Control Procesos

ID	Lote	Estampa	Piezas a fabricar	Turno	Planeado	Fecha Planeación	Documentos Planeación
1	DO-13823-01	2023-09-18 15:00:00	9,000	alto	✓	2023-09-18 20:03:16	1
2	DP-13823-11:24:00	2023-09-19	6,000	alto	✓		
3	DP-053823-01	2023-09-22 09:13:00	4,000	alto	✓	2023-09-22 09:11:27	1
4	DP-033823-01	2023-09-22 09:11:00	5,000	baj	✓	2023-09-22 09:12:46	1
5	DP-023823-01	2023-09-22 09:13:00	5,000	alto	✓	2023-09-22 09:14:36	1
6	DP-033823-01	2023-09-22 10:15:00	6,000	alto	✓	2023-09-22 09:17:10	1
7	DP-033823-01	2023-09-22 09:20:00	3,000	alto	✓	2023-09-22 09:18:34	1
8	DP-013823-01	2023-09-22 16:23:00	5,000	alto	✓	2023-09-22 10:22:42	1
9	DP-013823-01	2023-09-22 14:49:00	10,000	alto	✓	2023-09-22 10:38:01	1

Mostrando registros del 1 al 9 de un total de 9 registros

(b) REGISTRO DE EVENTOS DEL SISTEMA

ID	Evento	Usuario	Timestamp
64	Liberación 1 de Actividades CP: id_cp=5 con 3 registros	admin@donpicoso.com	2023-09-22 09:29:02
63	El usuario gestorcalidad7@dp.com ha completado la actividad 25 de la OP: 4	gestorcalidad7@donpicoso.com	2023-09-22 09:25:49
62	El usuario gestorcalidad7@dp.com ha completado la actividad 26 de la OP: 4	gestorcalidad7@donpicoso.com	2023-09-22 09:25:49
61	Liberación 1 de Actividades CP: id_cp=6 con 6 registros	admin@donpicoso.com	2023-09-22 09:21:37
60	Liberación 1 de Actividades CP: id_cp=4 con 2 registros	admin@donpicoso.com	2023-09-22 09:21:10
59	Creación de CP (no liberado): id=7; estampa=2023-09-22 09:18:34; número de ac	supervisor1@donpicoso.com	2023-09-22 09:18:34
58	Creación de CD: id=6; estampa=2023-09-22 09:17:10; número de ac	cocinero3@donpicoso.com	2023-09-22 09:17:10
57	Creación de CD: id=5; estampa=2023-09-22 09:14:36; número de ac	cocinero3@donpicoso.com	2023-09-22 09:14:36
56	Recálculo de CP (librada en producción): id=1; estampa=2023-09-19 11:11:59;	admin@donpicoso.com	2023-09-22 09:12:40
55	El usuario gestorcalidad7@dp.com ha supervisado la actividad 16 de la	gestorcalidad7@donpicoso.com	2023-09-22 09:11:27

Mostrando registros del 11 al 20 de un total de 74 re Anterior 1 2 3 4 5 ... 8 Siguiente

(a)

(b)

Figura 4. Interfaces: (a) central de registro de Control Procesos y (b) registro de eventos del sistema.
Fuente: Elaboración propia.

(a) Control Procesos Report

REPORTE DE CONTROL DE PROCESOS JARABE R. C.									
Fecha:		Orden SAP #:		Lote #:		Código		Num. de Rev.	
2023-12-05	Batch #:	133	Cantioso:	25000	Cuarto #:	134	Tanque #:	5	Ficha:
1.- COMPONENTES DE INICIO									
BLANCO 64*48° BRIX									
Características del jarabe a usar:									
PH: 23 Brix: 24 Capacidad: 25 Turbidez: 26									
2.- AZUCAR Y FINOS AGREGADOS									
Descripción: # de Costales: Cantidad Agregada Lote #									
Azúcar Jarabe: 30 Kg 31 Kg 32 Color Caramelo 33 Kg 34 Kg 35 Aromatizante 35 Kg 36 Fines 37 Kg									
3.- AJUSTE DE pH HIDROXIDO DE SODIO (NaOH) o ACIDULANTE									
Alcalino Lote Cantidad pH Acidulante Lote Cantidad pH Temp. del Jarabe									
NaOH 60 61 62 63 64 65 66 Temp. 67 60 71 72 73 74 75 76 77 60 82 83 84 85 86 87 88 90 91 92 93 94 95 96 97									
Total NaOH Agregado: 110 Total Acidulante Agregado: 120 pH Final: 130									
4.- INFORMACIÓN DE VAPOR									
Hora de Inicio de Vapor: 00:23 Hora Final de Vapor: 03:23									
5.- DATOS DE PROCESO (Lectura Final para Siembra)									
Hora inicio Lanzado a Charlors: 04:23 Hora Final a Charlors: 05:23 HORA Jarabe: 05:23 Hora Final: 05:24 Hora de Inicio: 07:23 Hora Final: 09:24 Temp. del Cuarto: 350 Temp. del Cuarto: 400									
6.- JARABE ACONDICIONADO									
Jarabe Bombeado 500 Agua Blanca 510 Fines Blancos 520 Fines Amarillo 530 *BRIX 540 pH 550									
7.- JARABE FILTRADO									
Cantidad: 600									

(b) Control Diario de Humedad y Temperatura

CONTROL DIARIO DE HUMEDAD Y TEMPERATURA											
Humedad Máxima 22.5					Temperatura Máxima 35.6						
Área: Cuarto #											
2023-09-14		2023-09-15		2023-09-16		2023-09-17		2023-09-18			
Hora*	H (%)	T (°C)	Hora*	H (%)	T (°C)	Hora*	H (%)	T (°C)	Hora*	H (%)	T (°C)
19:00	16	56.4	07:00	19	43.7	07:15	18	38.4	07:35	30	34.8
21:00	15	49.8	09:00	19	42.0	09:35	23	35.2	09:45	21	34.5
11:00	15	49.7	11:00	20	41.0	12:35	31	33.9	11:55	30	31.9
01:00	16	47.5	13:00	20	40.6	14:20	26	33.6	13:25	37	31.8
03:00	17	45.8	15:00	20	40.1	15:10	24	33.3	15:30	38	31.1
05:00	18	44.8				16:00	38	31.7			
Productos											
3 tinas con Crystals											
3 tinas de Crystals											
3 tinas p/Vapor											
Cuarto Vacío											
Tinas Limpias											
N.A.											
Observaciones											
Se abre puerta p/hemp N.A.											
Se bombeó Jarabe											
Se rompen Crystals											
Limpieza de tinas											
Se tejen tinas											
Rociando hilos											
Gestor de Calidad: _____ Supervisor: _____											
*El control de humedad y temperatura se lleva a cabo cada dos horas											

(a)

(b)

Figura 5. Documentos PDFs generados por los registros de los reportes: (a) Control Procesos y (b) Control Diario de Temperatura/Humedad.
Fuente: Elaboración propia.

Las 24 variables son: *fecha*, *hora*, *temperatura_tiempo_real*, *humedad_tiempo_real*, *temperatura_promedio*, *humedad_promedio*, *número_tanque*, *producto*, *brix_jarabe_acondicionado*, *brix_final*, *humedad_jarabe_acondicionado*, *finos*, *pH_inicial*, *pH_final*, *naoh_para_ajustar_pH*, *hora_inicio_vapor*, *hora_final_vapor*, *tiempo_cocción_vapor*, *número_cuarto*, *temperatura_cuarto*, *tiempo_inicial_chr*, *tiempo_final_chr*, *tiempo_sembrado*, y *resultado_paleta*. Para el modelo, se utilizan sólo 19 de éstas, eliminando *fecha*, *hora*, *hora_inicio_vapor*, *hora_final_vapor* y *producto*, por considerarse irrelevantes.

El conjunto de datos está conformado por 1000 registros de 24 variables, clasificado en cuatro grupos: Paleta OK, No crece paleta, Paleta se engrana y Paleta se pega a moldes.

Resultados

Como evidencia de los resultados arrojados por el sistema, en la Figura 6(a) se muestra una clasificación de un producto de máxima calidad, representado por el texto “La Paleta está OK”, la imagen de una paleta de alta calidad y el color verde como parte del semáforo visual. mientras que en la Figura 6(b)

se muestra una clasificación de un producto donde se tiene que la paleta no creció, representado por el texto “La Paleta No Crece”, la imagen de una paleta que no se formó correctamente y el color rojo como parte del semáforo visual. Es importante notar que la Figura 6 no indica el nombre de todas las métricas o parámetros, porque la empresa solicitó no presentarlas. No obstante, en la plataforma configurada para la empresa dentro del servidor, se muestra cómo es el proceso de pronóstico de calidad, el nombre de cada variable (Temperatura, Humedad, Tiempo, etc.) y el placeholder de cada campo de entrada indicando las unidades como deben ingresarse (kg, %, °C, etc.).

Se evaluaron dos grupos de modelos: (a) ML y (b) DL, que están basados en redes neuronales artificiales. En el primer grupo, se evaluaron cuatro modelos de ML para clasificar las 3 diferentes salidas de la paleta; en la Figura 7 se muestran los resultados con el conjunto de prueba.

La Tabla 1 enumera numéricamente la exactitud promedio y la desviación estándar de la exactitud para cada modelo de ML utilizado.

El conjunto de datos ha sido pre-procesado para eliminar valores nulos o características que no apor-

Ingresar variables o parámetros a simular

Procesando
Clasificar

CONTROL DIARIO DE TEMPERATURA Y HUMEDAD

Variable 1: En unidades 1	Variable 2: En unidades 2	Variable 3: En unidades 3
...		
Variable 25: En unidades 25	Variable 26: En unidades 26	Variable 27: En unidades 27

Pronóstico de Calidad

Respuesta: La paleta está OK

(a)

Ingresar variables o parámetros a simular

Procesando
Clasificar

CONTROL DIARIO DE TEMPERATURA Y HUMEDAD

Variable 1: En unidades 1	Variable 2: En unidades 2	Variable 3: En unidades 3
...		
Variable 25: En unidades 25	Variable 26: En unidades 26	Variable 27: En unidades 27

Pronóstico de Calidad

Respuesta: La paleta no crece

(b)

Figura 6. Salidas visuales en la plataforma de Industria 4.0 para la clasificación del modelo.
Fuente: Elaboración propia.

tan al conocimiento, además, se han evaluado varios modelos y se presentan los cuatro más interesantes, donde la metodología de selección se fundamentó en: a) dividir en conjuntos de entrenamiento y de pruebas, b) con el conjunto de entrenamiento se explora un espacio de análisis con siete modelos a través de una validación cruzada con diez divisiones, c) análisis estadístico para la selección de cuatro modelos, d) ajuste de hiper-parámetros para una mejor sintonización de estos modelos, y e) selección del modelo final mediante mejor exactitud y menor desviación estándar de la exactitud con el conjunto de pruebas. La métrica de exactitud fue elegida, porque se deseaba un modelo que clasificara de mejor manera a cada una de las salidas esperadas enfocadas en la calidad del producto, por lo que el mejor modelo fue utilizando Regresión Logística que, a través de distintas corridas de la metodología, fue el que mejor resultado presentó con una exactitud del 81.05 % ante vectores no vistos integrados en el conjunto de pruebas.

Siguiendo una metodología similar de selección, se evaluó el segundo grupo de modelos basados en

redes neuronales artificiales. Se tuvo una gran cantidad de configuraciones, y para la elección de los mejores modelos se utilizó el comportamiento a través de la exactitud promedio y su desviación estándar a nivel de exploración del tamaño del batch, así como, a través de la exactitud promedio y su desviación estándar a nivel de configuración; las más representativas se presentan en Tabla 2.

Para cada una de las seis configuraciones anteriores se exploró el espacio de diseño con los siguientes hiper-parámetros:

- Tamaño del batch = [5, 10, 25, 50]
- Épocas = [50, 100, 250, 500, 1000]

Las Tablas 1 y 2 permitieron hacer la selección final del mejor modelo entre los grupos ML y DL, siendo los últimos los que muestran mejor desempeño; resaltando que, para algunos autores, los modelos de DL son aquellos que además de las capas de entrada y salida, presentan más de una capa oculta, aunque en este trabajo se considera que los modelos DL son los basados en redes neuronales artificiales. Los resultados muestran que las redes neuronales artificiales

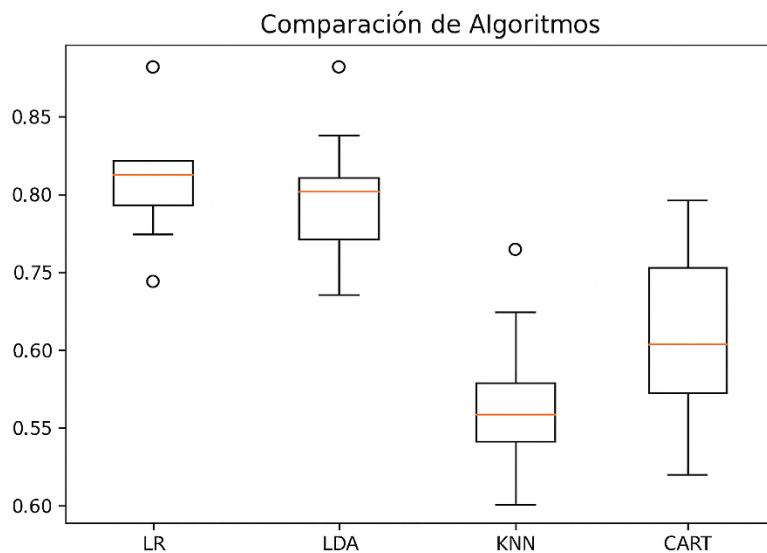


Figura 7. Modelos evaluados con el conjunto de pruebas.
Fuente: Elaboración propia.

Tabla 1. Resultados de la exactitud de los diferentes modelos analizados.

Modelo	Exactitud (Acc) (%)	Desviación Estándar (Acc) (%)
LR (Regresión logística)	81.05	3.38
LDA (Análisis de Discriminantes Lineales)	79.42	4.36
KNN (K vecinos más cercanos)	66.12	4.72
CART (Árboles de clasificación y regresión)	70.40	5.65

Fuente: Elaboración propia.

les se comportan de mejor manera cuando se tienen configuraciones de 1 capa de entrada, 1 capa oculta y 1 capa de salida, a diferencia de cuando se configuran más capas ocultas. Adicionalmente, los modelos ML1, ML2 y ML3 reportan una mayor exactitud promedio con una menor desviación estándar, siendo entre éstas la configuración ML3 la que reporta los mejores resultados, es decir la configuración de 1 capa de entrada con 14 neuronas, 1 capa de salida con 3 neuronas (función de activación *sigmoid*) y 1 capa oculta con 15 neuronas (función de activación *relu*), abstrae el conocimiento y lo modela de mejor manera con el conjunto de datos proporcionado.

Conclusiones

Las plataformas de Industria 4.0 son herramientas importantes en el manejo de procesos actuales. Como tal, la plataforma descrita en este documento refleja el conocimiento de los cocineros, supervisores y auditores para obtener un modelo que permite la clasificación de la calidad de un producto (paletas con innovación en su estructura). Este modelo aporta en dos sentidos, por un lado, permite conocer la salida con las variables o características de entrada, por otro lado, ofrece ideas de nuevas variables para tener un mayor y mejor control de la calidad del producto.

La plataforma da ventajas en cómo se registran los reportes de los distintos usuarios de ésta, reduciendo el tiempo de registro, almacenamiento, procesamiento y análisis de los distintos datos capturados. Ade-

más, permite mantener seguro el almacenamiento de estos mediante contraseñas, sesiones, perfiles, privilegios y registro de eventos del uso de la plataforma. Un requerimiento que puede cambiar en el futuro es que el proceso se mantendrá de manera local, ya que se está modelando un producto de innovación a nivel mundial. Los directivos de la empresa consideraron que es adecuado mantener el proceso dentro de servidores locales, para evitar la fuga de información y lograr tener un mejor control de lo que se puede ver hacia dentro de la empresa.

Agradecimientos

Los autores agradecen al CONAHCYT (ahora SE-CIHTI), por el apoyo al proyecto de Estímulo Fiscal a la Investigación y Desarrollo de Tecnología de la Convocatoria EFIDT 2023 con número 1900477, así como el estímulo para llevar a cabo las investigaciones a través del programa SNII y al Proyecto Cátedras CONAHCYT número 882.

Referencias

- Balfaqih, H. (2023). Artificial Intelligence and Smart Logistics Systems in Industry 4.0, *In Proceedings of the International Conference on Industrial Engineering and Operations Management*, Manila, Philippines, March 7-9, 2023.
- Buvaneswari, M., Aswath, S., Karthik, P y Mohammed M. R. (2023). Web based nutrition and diet assistance using machine learning, *International Research Journal of Engineering and Technology*, 10(5), 983-987.
- Calle, J. L. P., Punta-Sánchez, I., González-de-Peredo, A. V., Ruiz-Rodríguez, A., Ferreiro-González, M. y Palma, M. (2023). Rapid and

Tabla 2. Exactitud promedio y la desviación estándar de los modelos DL ante el conjunto de pruebas.

Modelo	Capa de entrada	Capa de salida	Capa(s) ocultas	Exactitud promedio	Desviación estándar
ML1	14 neuronas	3 neuronas FA Sigmoid	1) 10 neuronas FA Relu	94.2%	9.2%
ML2	14 neuronas	3 neuronas FA Sigmoid	1) 20 neuronas FA Relu	97.2%	5.7%
ML3	14 neuronas	3 neuronas FA Sigmoid	1) 15 neuronas FA Relu	97.2%	4.4%
DL1	14 neuronas	3 neuronas FA Sigmoid	1) 10 neuronas FA Relu 2) 6 neuronas FA Relu	88.9%	18.6%
DL2	14 neuronas	3 neuronas FA Sigmoid	1) 20 neuronas FA Relu 2) 16 neuronas FA Relu	96.3%	8.3%
DL3	14 neuronas	3 neuronas FA Sigmoid	1) 10 neuronas FA Relu 2) 10 neuronas FA Relu 3) 10 neuronas FA Relu	85.9%	22.7%

Fuente: Elaboración propia.

- automated method for detecting and quantifying adulterations in high-quality honey using vis-NIRs in combination with machine learning. *Foods*, 12(13), 2491.
- Candanedo, I. S., Nieves, E. H., González, S. R., Martín, M. T. S. y Briones, A. G. (2018). Machine learning predictive model for industry 4.0. In *Knowledge Management in Organizations: 13th International Conference, KMO 2018*, Žilina, Slovakia, August 6–10, 2018.
- Cubillas, J. J., Ramos, M. I., Jurado, J. M. y Feito, F. R. (2022). A machine learning model for early prediction of crop yield, nested in a web application in the cloud: a case study in an olive grove in southern Spain. *Agriculture*, 12(9), 1345.
- Ferrini, S., Rollo, C., Bellino, C., Borriello, G., Cagnotti, G., Corona, C., Di Muro, G., Giacobini, M., Iulini, B. y D'Angelo, A. (2023). A novel machine learning-based web application for field identification of infectious and inflammatory disorders of the central nervous system in cattle. *Journal of Veterinary Internal Medicine*, 37(2), 766-773.
- Fracarolli, J. A., Pavarin, F. F. A., Castro, W. y Blasco, J. (2020). Computer vision applied to food and agricultural products. *Revista Ciéncia Agronómica*, 51(spe), e20207749.
- Hassoun, A., Jagtap, S., Garcia-Garcia, G., Trollman, H., Pateiro, M., Lorenzo, J. M., y Câmara, J. S. (2023). Food quality 4.0: From traditional approaches to digitalized automated analysis. *Journal of Food Engineering*, 337, 111216.
- Jagadeesh B N., Apoorva J. P., Chandana S., Kumar, K. y Girish, J. (2022). Web Based Application for Recognition of Deterioration in Fruits by Machine Learning. *International Journal of Creative Research Thoughts*, 10(7), 349-353.
- Kruthika, V., Lavanya H. R., Mahalakshmi E. H., Ranju P. S. R., Priyanka H.L y Sindhu, K. S. (2023). Integrated approach for food donation system, restaurant food demanding forecasting using machine learning, and global food waste analysis. *International Research Journal of Modernization in Engineering Technology and Science*, 5(7), 11-18.
- Nordin, M. J., Xin, O. W. y Aziz, N. (2019). Food image recognition for price calculation using convolutional neural network. In *Proceedings of the 2019 3rd International Conference on Digital Signal Processing*. Jeju Island Republic of Korea, February 24-26, 2019.
- Sahoo, D., Hao, W., Ke, S., Xiongwei, W., Le, H., Achananuparp, P., Lim, E. y Hoi, S. C. (2019). FoodAI: Food image recognition via deep learning for smart food logging. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. Anchorage AK USA, August 4-8, 2019.
- Sánchez-Torres, F., González, I. y Dobrescu, C. C. (2022). Machine Learning in Business Intelligence 4.0: Cost Control in a Destination Hotel. *International Journal of Interactive Multimedia and Artificial Intelligence*, 7(3), 86-95.
- Soumyashree, S., Gupta, A. y Biswas, B. (2023). An Automation Designed for Industry 4.0 Using Robotics and Sensors that Based on IoT & Machine Learning. *E3S Web of Conferences*, 405, 02014.
- Verma, A., Kapoor, C., Sharma A. y Mishra B. (2021). Web Application Implementation with Machine Learning. *2nd International Conference on Intelligent Engineering and Management (ICIELM)*, London, United Kingdom, April 28-30, 2021.