

## MÁQUINA DE APRENDIZAJE PARA OPTIMIZACIÓN VÍA SIMULACIÓN DE ACIDULACIÓN DE RESIDUOS ORGÁNICOS HÚMEDOS

(*MACHINE LEARNING FOR OPTIMIZATION VIA SIMULATION OF WET ORGANIC WASTE ACIDULATION*)

*Enrique C Ávila, Marlene Brizuela.*

Universidad Nacional Experimental de los Llanos Occidentales “Ezequiel Zamora” UNELLEZ. Subprograma de Estudios Avanzados Ciencias del Agro y del Mar. San Carlos-Estado Cojedes, Venezuela.

*enriqcÁvila@gmail.com / marlenebrizuela@hotmail.com*

*Recibido: 30-09-2021/ Aceptado: 09-10-2021*

### RESUMEN

Esta investigación experimental exploratoria estocástica y determinística de machine learning en el ajuste de un modelo de red neuronal artificial con ajuste por el método de Gauss Newton para optimización predictiva multifactor-multiobjetivo vía simulación de un acidulante (mezcla de ácidos orgánicos carboxílicos) de residuos orgánicos no persistentes con alta humedad (RONoP-AH) de consumo animal, residuos estandarizado al 2 % de sal (NaCl) consumo animal estándar y 2 % de melaza de caña de azúcar; el acidulante como tecnología de control de variables indicadoras de cambios fisicoquímica para la estabilización biológica de los RONoP-AH. El diseño de muestreo estadístico fue un Gaussiano IMSE optimal (GIMSEOp). Para las restricciones de variabilidad de las respuestas de calidad en la optimización vía simulación se utilizaron funciones de reglas difusas multimétrica: minimizar la acidez iónica (pH: índice de moles H<sup>+</sup>/litro) y la acidez titulable total (ATT: ml NaOH 1N), target el potencial redox (POR: mV) entre +0 y +200 mV y maximizar la capacidad buffer (CB: ml NaOH 1N±1ΔpH); bajo la hipótesis fisiológica de que un POR negativo cercano a cero (alejado de -450 mV) inhibe el metabolismo respiratorio de microorganismos (MRM) anaerobios y un POR cercanos a 0 mV (alejados de +850 mV) inhibe el MRM aerobios. La mezcla óptima operativa de ácidos obtenida fue: 3,46 g de Citrato de Na (~35,67%), 0,62 g de Benzoato de Na (~6,39%), 3 ml de Acido Fórmico (~30,93%) y 2,62 ml de Ácido Acético (~27,01%), con una deseabilidad de calidad del acidulante de ATT: 290,1, CB: 147,6 POR: 22,9 y pH: 2,15. Validándose la mezcla óptima acidulante en RONoP-AH, con excelentes resultados, lográndose estabilizarlos biológicamente y almacenarlos al ambiente sin refrigeración ni biocidas, por periodos de más de 3 meses.

**Palabras clave:** Diseño experimental, Redox, Modelo, Optimización.

## ABSTRACT

This exploratory experimental stochastic and deterministic of machine learning research on the adjustment of an artificial neural network model with Gauss Newton method adjustment for multifactor-multiobjective predictive optimization via simulation of an acidulant (mixture of carboxylic organic acids) of non-persistent high moisture organic wastes (RONoP-AH) from animal consumption, waste standardized to 2 % salt (NaCl) standard animal consumption and 2 % sugar cane molasses; acidulant as a technology to control variables indicating physicochemical changes for the biological stabilization of ROnoP-AH. The statistical sampling design was a Gaussian IMSE optimal (GIMSEOp). For the variability constraints of the quality responses in the optimization via simulation, multimetric fuzzy rule functions were used: minimize the ionic acidity (pH: moles H<sup>+</sup>/liter ratio) and total titratable acidity (ATT: ml NaOH 1N), target the redox potential (POR: mV) between +0 and +200 mV and maximize buffer capacity (CB: ml NaOH 1N±1ΔpH); under the physiological hypothesis that a negative POR close to zero (away from -450 mV) inhibits anaerobic microorganism respiratory metabolism (MRM) and a POR close to 0 mV (away from +850 mV) inhibits aerobic MRM. The optimal operating acid mixture obtained was: 3,46 g Na Citrate (~35,67%), 0,62 g Na Benzoate (~6.39%), 3 ml Formic Acid (~30.93%) and 2,62 ml Acetic Acid (~27.01%), with an acidulant quality desirability of ATT: 290,1 CB: 147,6 POR: 22,9 and pH: 2.15. Validating the optimal acidulant mixture in ROnoP-AH, with excellent results, achieving biological stabilization and storing them in the environment without refrigeration or biocides, for periods of more than 3 months.

**Keywords:** Experimental design, Redox, Model, Optimization.

## INTRODUCCIÓN

Hay una miríada de excelentes herramientas auxiliares de machine learning, computación inteligente emergente, software, modelación, simulación de procesos y muy baja cultura de datos, que en general son ignoradas y por tanto muy poco explotadas en el mundo académico, solo aparece en algunas investigaciones, por intereses particulares; herramientas que están siendo ampliamente adoptadas y aplicadas en cada una parte de las cadena de procesos productivos y administrativos vigentes; una de las razones de la poca instrucción académica en estas herramientas es la analfabetismo tecnológico digital de los administradores docentes de aula y de los creadores de pensum académicos; factores que ha generado una analfabetización digital también de empresas, fundamentalmente PYMEs

(Grupo eDISTeL. Coaching tecnológico, 2021; AD ALFABETIZACIÓN DIGITAL, 2021; BLOGRH.Amedirth, 2021).

Así, se plantea una brecha social de alfabetización digital emergente en cuanto a instrucción en competencias digitales y cultura de DICs de los nuevos profesionales universitarios, una deficiencia fundamentalmente en los países subdesarrollados, una educación digital requerida para direccionar el desarrollo y crecimiento de la industria y sociedad global 5.0, de la Big DICs. (Datos, Información y Conocimientos masivos), una nueva división social, cultural, económica y educativa (tokioschool, 2021; BLOGRH.Amedirth, 2021).

En lo que respecta producción de residuos orgánicos no persistentes agroindustriales con alta humedad (RNoP-AH) que están impactando los bienes y servicios ambientales (BySA), el BID (2018), reporta que el 61 % de los residuos sólidos urbanos (RSU) domésticos son orgánicos (17 % papel y 44 % alimentos); sin embargo se hace énfasis que al incluir los residuos fecales, agrícolas de campo, los provenientes del matadero de animales y los residuos de campo de galpones de cría de animales; estas estadísticas se vuelven mucho más elevadas y, son un problema vigente y futuro. Y, específicamente en Venezuela (Ávila, 2019b), se ha detectado que en épocas de esplendor de disponibilidad de divisas extranjeras, para el año 2009, el consumo de carnes de solo tres (03) rubros animal, fue. Pollos broiler, 42 Kg/persona\*año; Bovinos, 24 Kg/persona\*año y Cerdo, 13 Kg/persona\*año y; según las proyecciones de INE (2019), para el año 2050, la población venezolana serán, > 40.500.761 habitantes.

En estas condiciones se producirán para el año 2050 (Ávila, 2019b), más de 2.831.753.198,7 Kg de residuos orgánicos húmedos, que representaran un problema grave, causal de impacto ambiental, si no se gestionan adecuadamente; sin embargo pueden representar una oportunidad de negocio (ONU Programa para el medio ambiente, 2020), si se implementan gestiones como las que propone esta investigación.

Basado en la problemática dual planteada, la investigación se enfocó como objetivo general en implementar máquina de aprendizaje (machine learning), de modelo de red neuronal artificial para modelar y optimizar vía simulación (usando un simulador comercial, con interfaz gráfica interactiva), donde la tecnología de acidulación es utilizada para la estabilización de ROnoP-AH, provenientes de procesos agroindustriales del beneficio y proceso de vegetales y animales.

## MATERIALES Y MÉTODOS

Para el cumplimiento del objetivo general de la investigación, se desarrollaron las siguientes fases:

1. Generación de un arreglo de muestreo de tratamientos usando el algoritmo Proceso Gaussiano IMSE Optimal del software SAS JMP 8.

2. Con los tratamientos de muestreo elaborados (mezclas de CitratoNa, BenzoatoNa, Ácido Fórmico y Ácido Acético), se ejecutó la experimentación y se determinaron las respuestas, acidez iónica (pH: índice moles H<sup>+</sup>/litro), Potencial Redox (PR: mV), Capacidad Buffer (CB: ml de NaOH 0,1N  $\pm 1 \Delta$ pH) y Acidez Titulable (ATT: ml de NaOH 0,1N), usando los métodos descritos en la A.O.A.C. (2012), adaptados y estandarizados durante las pruebas pilotos.

3. Ajuste del modelo de red neuronal artificial, concertando los parámetros con entrenamiento por el método de validación cruzada (K-Fold crossvalidation, recomendado cuando la cantidad de tratamientos de muestreo es pequeño (ejemplo, entre 5 y 20), método de Gauss Newton (para modelos de hasta 40 coeficientes regresores), esto para encontrar modelos que generalicen bien al ser introducidos nuevos datos. Para ello en cada iteración de reajuste del modelo, se reservó grupos de 05 tratamientos distintos aleatorizados, que se utiliza para verificar el modelo generado por los datos de entrenamiento no retenidos; los valores iniciales de arranque (valores cebadores) de los coeficientes regresores del modelo

no lineal (modelo de red neuronal artificial) para cada iteración, se estiman con regresión usado todos los datos.

Finalmente se crea una red de ecuaciones intrínsecas (un metamodelo de simulación), comprobando su bondad de ajuste con el coeficiente de regresión ( $R^2$ ) parciales y las gráficas de valores medidos vs. valores predichos de cada modelo para cada respuesta con el cual se emuló el proceso de elaboración del acidulante para control redox de residuos orgánicos no persistente, con alta humedad.

4. Co-optimización del proceso de creación del acidulante, vía simulación, con el simulador Prediction Profiler con interfaz gráfico interactivo del software SAS JMP 8 (SAS JMP, 2012; 2013), según las recomendaciones de Derringer (1994) y Derringer y Suich (1980), bajo las siguientes restricciones de deseabilidad de calidad (funciones de reglas difusas multimétrica) para las cuatro respuestas medidas, 1. Minimizar la respuesta acidez iónica (pH: índice moles  $H^+$ /litro), 2. Target el potencial redox entre 0 mV y 200 mV, 3. Minimizar la acidez titulable soluble total (ATT: ml NaOH 0,1 N) y 4. Maximizar la capacidad buffer (CB: ml NaOH 0,1 N  $\pm 1\Delta pH$ ), para cada una de los tratamientos diseñados, con la mezcla de ácido fórmico, ácido acético, cítratoNa y benzoatoNa.

5. Validación de la mezcla óptima del acidulante; aplicándola a residuos orgánicos no persistentes con alta humedad (RONoP-AH), residuos industriales de beneficio animal, estandarizados estos con adición de sal (NaCl) de consumo animal estándar a dosis fija de 2 % y melaza de caña de azúcar a dosis fija de 2 %.

## RESULTADOS Y DISCUSION

En lo siguiente se describen los datos, los resultados de su análisis estadístico-matemático, análisis tecnológicos y su discusión.

En este propósito, en la tabla 1 se muestra Matriz “D” de diseño de muestreo y las respuestas medidas (promedio aritmético de tres repeticiones).

**Tabla 1.** Diseño Proceso Gaussiano IMSE Optimal y respuestas

Tratamientos	Factores experimentales				Respuestas			
	Ácido Fórmico	Ácido Acético	Citrato de sodio	Benzoato de sodio	Y <sub>1</sub>	Y <sub>2</sub>	Y <sub>3</sub>	Y <sub>4</sub>
	X <sub>1</sub> [ml]	X <sub>2</sub> [ml]	X <sub>3</sub> [g]	X <sub>3</sub> [g]	pH	PR	CB	ATT
1	2,964	2,464	3,536	0,616	2,11	22,55	145,63	291,22
2	3,461	2,961	3,983	0,867	2,28	29	176,94	276,98
3	2,514	2,947	3,986	0,636	2,38	29,24	145,54	274,26
4	2,517	2,961	3,039	0,633	2,37	29,44	134,7	276,08
5	3,463	2,963	3,037	0,633	2,32	40,07	182,61	275,09
6	3,461	2,017	3,039	0,867	2,50	32,85	124,52	264,81
7	2,522	2,022	3,978	0,646	2,38	29,08	111,35	271,80
8	3,447	2,014	3,986	0,864	2,51	35,27	130,05	266,92
9	2,964	2,464	3,536	0,884	2,34	29,55	144,26	276,77
10	2,514	2,014	3,053	0,636	2,13	15,96	112,08	290,01
11	2,514	2,014	3,053	0,864	2,51	30,24	114,44	267,50
12	3,463	2,963	3,037	0,867	2,46	41,78	164,65	251,4
13	2,973	2,473	2,955	0,750	2,48	37,92	151,63	270,01
14	2,973	3,045	3,527	0,750	2,40	34,43	160,87	273,01
15	3,447	2,014	3,986	0,636	2,37	44,77	146,37	273,59
16	2,514	2,947	3,986	0,864	2,40	29,64	141,37	274,24
17	3,461	2,017	3,039	0,633	2,35	32,15	121,81	276,69
18	2,517	2,961	3,039	0,867	2,51	33,5	135,82	264,42
19	3,545	2,473	3,527	0,750	2,36	29,69	115,76	277,01
20	3,461	2,961	3,983	0,633	2,35	42,14	175,1	274,52
21	2,522	2,022	3,978	0,854	2,43	40,52	133,08	271,67

Fuente: Ávila (2021).

En la tabla 2, se muestra las condiciones de entrenamiento de la red neuronal artificial (RNA) y en la tabla 3, la historia del entrenamiento de ajuste de la RNA.

**Tabla 2.** Condiciones de entrenamiento de la RNA

Característica	Valor
Capas ocultas	1
Nodos escondidos	3 a 7
Penalización por sobreajuste	0,01 a 0,04
Numero de tours	
(Número de puntos óptimos locales recorridos en búsqueda de un óptimo máximo global)	16 a 25
Iteraciones máximas	75 a 100
Criterio de convergencia	
(Cambio en el valor de la función objetivo  change  / min(10 <sup>-8</sup> , max(10 <sup>8</sup> , obj)) )	0,00001

Fuente: Ávila (2021).

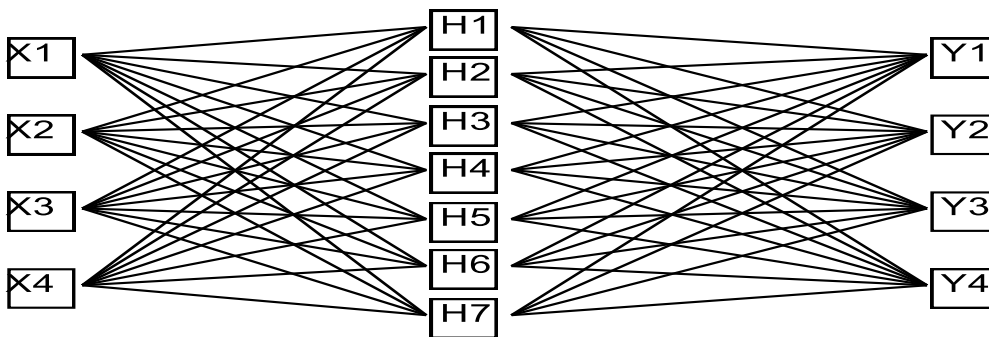
La historia del entrenamiento del metamodelo RNA, arrojó que el mejor  $R^2$  (98,767 %) se consigue con 1 capa oculta y 7 nodos ocultos y, la figura 1, muestra los  $R^2$  para c/u modelo ajustado de cada respuesta.

**Tabla 3.** Resultados de bondad de ajuste de la RNA para cada modelo de cada respuesta

$Y_i$	SSE	RMSE	SSE Scaled	RMSE Scaled	RSquare (%)
Y1	0,01266406	0,03007617	1,06544328	0,27586789	94,672784
Y2	30,9562467	1,48699714	0,6606732	0,21723476	96,696634
Y3	564,978544	6,35260657	1,21776871	0,29492962	93,911156
Y4	76,7141839	2,34085125	1,10688797	0,28118219	94,46556

**Fuente:** Ávila (2021)

La figura 1, muestra el diagrama de la RNA, un perceptrón monocapa multinodo (7 nodos).



**Figura 1.** Diagrama esquemático de la RNA

**Fuente:** Ávila (2021)

En la figura 2 y 3 siguientes, se muestra las estructuras del metamodelo de RNA, un modelo para cada respuesta y las respectivas formulas  $H_i$ , todos encriptados en la hoja de cálculo del software donde cada  $Y_i$  y cada  $H_i$  como un vector columna, los cuales están híper vinculados, ya que cada  $Y_i = f(H_i)$ .

Modelo para cada una de las respuestas objetivos $Y_i$	
$\hat{Y}_1$ [pH]	$(1.138 + 4.079*H1 - 0.384*H2 - 1.450*H3 + 2.65 *H4 - 3.303*H5 - 2.415*H6 - 1.080*H7) * 0.109 + 2.378$
$\hat{Y}_2$ [POR]	$(-0.023 + 4.332 *H1 + 2.302 *H2 - 0.920*H3 + 0.518 *H4 - 0.735 *H5 - 2.18*H6 - 2.898 *H) * 6.845 + 32.847$
$\hat{Y}_3$ [CB]	$(-2.182 + 1.549*H1 + 2.202*H2 + 2.596*H3 - 1.796* H4 + 2.042*H5 - 1.743*H6 + 1.038*H7) * 21.539 + 141.361$
$\hat{Y}_4$ [ATT]	$(-1.00 - 3.666*H1 + 0.284*H2 + 1.367*H3 - 1.96*H4 - 0.072*H5 + 2.105*H6 + 1.487*H7) * 8.325 + 273.238$

Figura 2. Metamodelo de RNA ajustado

Fuente: Ávila (2021)

Formulas $H_j$ de activación para cada uno de los 07 nodos ocultos	
H1	$\text{Squish}((-2.282) - 0.034 * ((X1 - 3.010) / 0.438) + 1.310 * ((X2 - 2.51) / 0.4385) - 2.729 * ((X3 - 3.40) / 0.438) + 1.242 * ((X4 - 0.75) / 0.111))$
H2	$\text{Squish}(2.316 + 3.026 * ((X1 - 3.010) / 0.4385) + -0.581 * ((X2 - 2.510) / 0.439) + 0.358 * ((X3 - 3.489) / 0.4385) + 1.617 * ((X4 - 0.75) / 0.111))$
H3	$\text{Squish}((-1.726) + -0.349 * ((X1 - 3.010) / 0.4385) + 0.7414 * ((X2 - 2.510) / 0.4385) + -1.697 * ((X3 - 3.490) / 0.4385) + -3.663 * ((X4 - 0.75) / 0.111))$
H4	$\text{Squish}((-2.598) + 2.412 * ((X1 - 3.010) / 0.438) + -1.543 * ((X2 - 2.510) / 0.4385) + -2.194 * ((X3 - 3.490) / 0.4385) + 1.370 * ((X4 - 0.75) / 0.111))$
H5	$\text{Squish}((-3.942) + 1.701 * ((X1 - 3.010) / 0.4385) + 0.981 * ((X2 - 2.510) / 0.4385) - 1.089 * ((X3 - 3.490) / 0.438) + 1.562 * ((X4 - 0.75) / 0.111))$
H6	$\text{Squish}(1.061 + 0.040 * ((X1 - 3.010) / 0.438) + -1.269 * ((X2 - 2.510) / 0.438) + -3.753 * ((X3 - 3.401) / 0.4385) + -0.572 * ((X4 - 0.75) / 0.111))$
H7	$\text{Squish}((-0.723) + 0.881 * ((X1 - 3.010) / 0.4385) + 1.016 * ((X2 - 2.510) / 0.438) + -1.241 * ((X3 - 3.490) / 0.4385) + 3.112 * ((X4 - 0.75) / 0.111))$

Figura 3. Metamodelo de RNA ajustado

Fuente: Ávila (2021)

Donde:

$$\hat{Y}_k = S_Y \left[ d_x + \sum_{j=1}^{N_H} (b_{jk} H_j) \right] \quad (\text{Ec-1})$$

$$\text{Squish}() = S_H(x) = \frac{1}{1 + e^{-x}} \quad (\text{Ec-2})$$

$$H_j = S_H \left[ c_j + \sum_{i=1}^{N_x} (a_{jk} X_i) \right] \quad (\text{Ec-3})$$

$\hat{Y}_k$  = Respuesta predicha

$S_Y(x)$  = La función identidad,

$N_H$  = El número de nodos escondidos

$H_j$  = Formula de la función de activación

$\text{Squish}()$  = La función logística de activación,

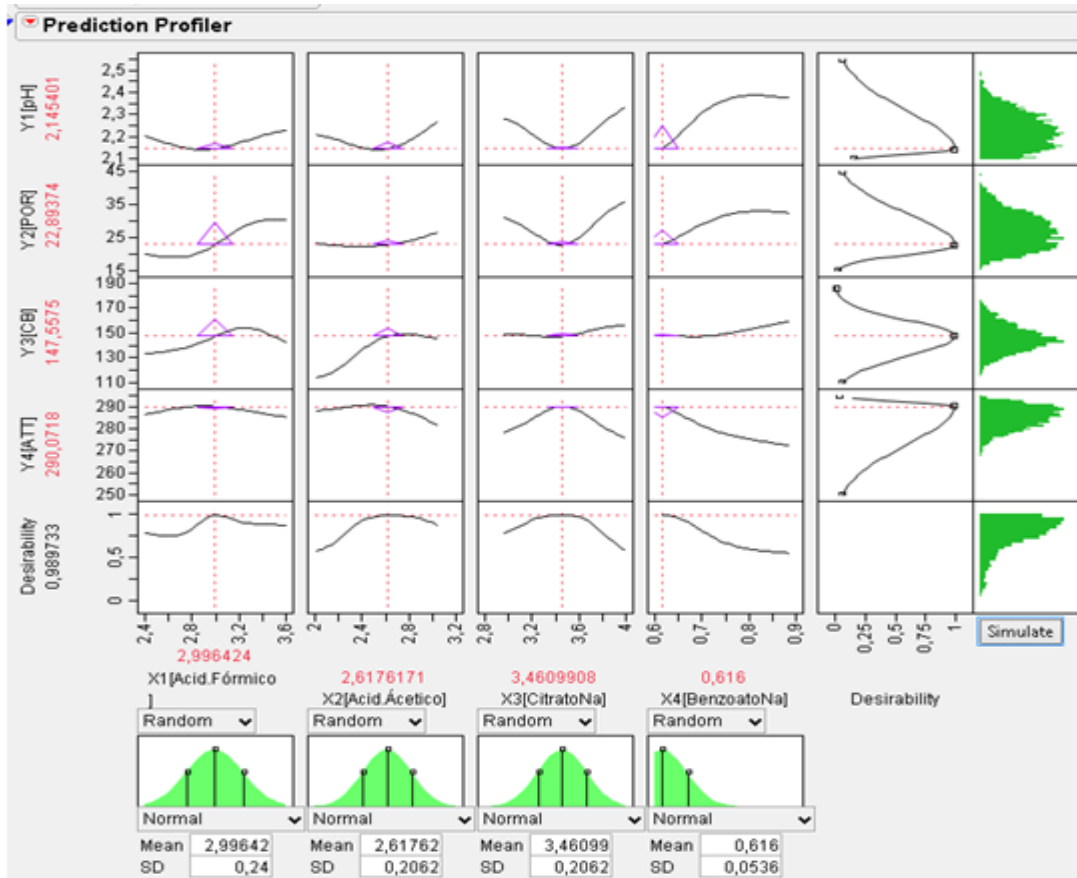
$N_x$  = el número de variables X,

$a, b, c$  y  $d$  = Coeficientes estimados



En la figura 4, se observa el prototipo de la interfaz interactiva del simulador SAS JMP 8, muestra los resultados de la optimización de la mezcla para la deseabilidad optimal del acidulante: 3,46 g de CitratoNa (~35,67%), más 0,62 g de BenzoatoNa (~6,39%), más 3 ml de Acido Fórmico (~30,93%) y 2,62 ml de Ácido Acético (~27,01%), obteniéndose un acidulante con la siguiente deseabilidad optimal (D de Derriger), ATT: 290,1, CB: 147,6 POR: 22,9 y pH: 2,15.

Por su parte, el ajuste de regresión global del metamodelo de RNA es, D = 98,97%, dando un DPMO de 337 (~3,37 no conformes por cada lote de 10.000 mezclas elaboradas).



**Figura 4.** Prototipo de la interfaz del simulador del programa SAS JMP v8

**Fuente:** Ávila (2021)

## CONCLUSIONES

Luego de los resultados obtenidos en el desarrollo de la presente investigación se han establecido las siguientes conclusiones:

1. El diseño de muestreo Proceso Gaussiano IMSE Óptimo, muestra ser económico para experimentación en laboratorio (experimentación estocástica y en experimentación determinística) y excelente en el ajuste de modelos de red neuronal con alta bondad de ajuste.

2. El proceso de machine learning, un metamodelo no lineal tipo red neuronal artificial ajustado, fue excelente en la optimización vía simulación, usando funciones de reglas difusas multimétrica, bajo el algoritmo de Derringer-Suich, implementado en el software SAS JMP 8.

3. El acidulante experimental creado, estabiliza materia orgánica altamente perecedera (rica en nutrientes y alta húmeda) casi instantáneamente, cuando se utilizan las dosis del acidulante experimental, permitiendo su preservación por más de 3 meses en condiciones ambientales sin refrigeración ni uso de biocidas.

## REFERENCIAS BIBLIOGRAFICAS

A.O.A.C. (2012). "Official methods of analysis of AOAC international. "William Horwitz y George W. Latimer. Ed.19. Washington, DC.

AD ALFABETIZACIÓN DIGITAL. (2021). Tecnología digital. El nuevo idioma obligatorio en la educación. Disponible en: <https://www.alfabetizaciondigital.redem.org/tecnologia-la-nueva-habilidad-obligatoria-en-la-educacion/>

- Ávila, E. C. (2019b). Optimizado de la revalorización con acidulado, de residuos del beneficio de pollos, con un DCC vía simulación. Guía de estudio. Laboratorio LITA y Laboratorio de Informática II. UNELLEZ-VIPI.
- BID. (2020). Investigaciones BID. Disponible en: <https://www.iadb.org/es>
- BLOGRH.Amedirth. (2021). ¿Cuál es el impacto de un analfabeta digital ante los riesgos y vulnerabilidades en el ciberespacio?. Disponible en: <https://www.amedirh.com.mx/blogrh/recursos-humanos/cual-es-el-impacto-de-un-analfabeta-digital-ante-los-riesgos-y-vulnerabilidades-en-el-ciberespacio/>
- Derringer, G. (1994). A balancing act: Optimizing a product's properties. Make intelligent trade-offs through desirability optimization methodology Quality progress. Disponible en: <http://www.statease.com/pubs/derringer.pdf>
- Derringer, G. y Suich, R. (1980). Simultaneous optimization of several response variables. Journal of quality technology. 12:214-219.
- Grupo EDISTEL. Coaching tecnológico. (2021). ¿La alfabetización digital es importante?. Disponible en: <http://www.grupoedistel.com/alfabetizacion-digital/>
- INE. (2019). Censos de población y vivienda. Disponible en: <https://www.google.com/search?q=INE+venezuela+2019&oq=INE+venezuela+2019&aqs=chrome..69i57j0i333l3.14573j0j15&sourceid=chrome&ie=UTF-8>
- ONU Programa para el medio ambiente (2020). Residuos y su valorización. Disponible en: <https://www.unep.org/es>
- SAS JMP. (2012). Software SAS Institute Inc. JMP business unit SAS.JMP: Jonhn's macintosh program.
- SAS JMP. (2013). SAS Institute Inc. Using JMP 11. Cary, NC: SAS Institute Inc. Inc. Disponible en: [http://www.jmp.com/support/downloads/pdf/jmp11/Using\\_JMP.pdf](http://www.jmp.com/support/downloads/pdf/jmp11/Using_JMP.pdf)
- Tokioschool. (2021). La inteligencia artificial y la alfabetización digital. Disponible en: <https://www.tokioschool.com/noticias/alfabetizacion-digital/>