

Redes Neuronales Artificiales: factores que determinan la cosecha de caña en la industria azucarera

Italo, Mendoza-Haro^{1*}; Hiram, Marquetti-Nodarse²

Resumen

La investigación muestra lo importante de las redes neuronales artificiales dentro de la industria azucarera, como una herramienta útil para la predicción del cultivo de la caña de azúcar, tomando como entradas la información climatológica: temperaturas máximas y mínimas, oscilación térmica, precipitaciones, heliofanía, humedad relativa, evaporación y hectáreas de los cultivos sembrados, para obtener una salida: toneladas de caña. Se desarrolló una herramienta de trabajo predictiva con resultados confiables, comparados con métodos tradicionales utilizados, como los aforos de expertos para la cosecha de la caña de azúcar. Se analizó la base de datos histórica de la organización, mediante un software MATLAB, herramienta matemática, que ofrece un entorno de desarrollo integrado (IDE) con lenguaje M de programación propio. La investigación se desarrolló en Compañía Azucarera Valdez S.A. Ubicada en la Ciudad de Milagro-Provincia del Guayas-Ecuador.

Palabras Clave: Heliofanía, IDE, industria azucarera, lenguaje M, MATLAB (MATrix LABoratory), Red neuronal artificial.

Artificial Neural Networks: factors that determine the cane harvest in a sugar industry

Abstract

The research shows the importance of artificial neural networks within the sugar industry, as a useful tool for the prediction of the cultivation of sugarcane, taking as input the climatological information: maximum and minimum temperatures, thermal oscillation, rainfall, heliophany, relative humidity, evaporation and hectares of crops planted, to obtain tons of cane as an output. A predictive work tool with reliable results was developed, compared with traditional methods used, such as expert assessment for sugarcane harvesting. The historical database of the organization was analyzed through MATLAB software, a mathematical tool which offers an integrated development environment (IDE) with its own M programming language. The research was developed at Compañía Azucarera Valdez S.A. located in the City of Milagro-Province of Guayas-Ecuador.

Keywords: Artificial neural network, heliophany, IDE, language M, MATLAB (MATrix LABoratory), sugar industry.

Recibido: 17 de noviembre de 2018

Aceptado: 07 de enero de 2019

¹ Director de mantenimiento industrial en Compañía Azucarera Valdez S.A. (CAVSA); Docente de la Facultad Ciencias de la Ingeniería Universidad Estatal de Milagro UNEMI; emendozah@unemi.edu.ec; <https://orcid.org/0000-0002-6492-6991>

² Doctor en ciencias económicas; profesor titular y consultor en el Centro de Estudios de la Administración Pública Cubana; Universidad de La Habana; email: marquetti@ceap.uh.cu

*Autor para correspondencia: emendozah@unemi.edu.ec

I. INTRODUCCIÓN

La globalización a nivel mundial se va tornando en un ambiente más competitivo y rezagando a muchas industrias que no se ajustan a su ritmo, por lo cual a diario se torna en un campo más hostil para aquellas industrias que no realizan cambios en sus estructuras tecnológicas e innovan permanentemente sus procesos, que requieren herramientas nuevas y eficientes para cumplir dichos objetivos.

La historia de la caña de azúcar en América Latina data de la mitad del siglo XVI, donde su cultivo comenzó a difundirse bajo un sistema de haciendas, con un proceso que abarcaba desde la producción agrícola hasta la elaboración de piloncillo y azúcar mascabada, además de la destilación de aguardiente. Durante los inicios del siglo XX, en el periodo Porfiriato (México) se comenzaron a establecer los sistemas de "ingenios", que consistía en una fábrica y los campos bajo el modelo de "plantación", ambos propiedad de un mismo dueño o empresa. En los años treinta se implementó una reforma agraria para dejar la producción de caña en manos de ejidatarios y pequeños propietarios. (Banko, 2005).

El desarrollo del proceso de industrialización significó un impulso adicional al desarrollo de la industria azucarera, ya que propició la ampliación de la base mecánica y tecnológica de la agroindustria, y la construcción de nuevas instalaciones para la producción de subproductos de la caña, entre otros (Marquetti, 2016).

La evolución de la tendencia a la ampliación de los diferenciales interterritoriales puede considerarse como una resultante de la minoración del carácter estratégico de la industria manufacturera y, sobre todo, del redimensionamiento del sector azucarero, un proceso que contribuyó, de forma directa, a que se produjeran importantes regresiones en los ámbitos tecnológico y productivo. (Marquetti, 2016).

La industria azucarera del Ecuador de acuerdo con la investigación predice sus resultados mediante aforos realizados por sus ingenieros y profesionales con experiencia, base para calcular los presupuestos y programas de cosecha de la caña, estos resultados

por lo general tienen variaciones importantes en la parte cuantitativa y en muchos de los casos conducen a tomar malas decisiones, generando pérdidas de productividad y económicas a las organizaciones industriales. El objetivo de la investigación será aplicar las redes neuronales artificiales utilizando Matlab, permitiendo una alternativa eficiente en la predicción para resultados de cosecha.

Predecir el rendimiento de un cultivo de caña de azúcar obedece a la necesidad de maximizar la relación inversión-ganancia y la disponibilidad de esa información, con anticipación, permite tomar decisiones sobre el manejo de una plantación, de forma específica para definir presupuestos de los productos que se van a elaborar con la materia prima . (Avila, Rodriguez, & Hernandez, 2012)

Los expertos en predicciones sobre el rendimiento mediante aforos y la experiencia del agricultor son importantes y cuando no se cuenta con métodos actualizados para la estimación, esta se convierte en el único recurso; sin embargo, estas aproximaciones pueden ser insuficientes, y lo que se necesita es información sistemáticamente almacenada que contemple, por ejemplo: registros históricos promedio de producción; y, variaciones de rendimiento por manejo agrícola o por factores climáticos entre otros, para disminuir cualquier sesgo o error. (Avila et al. 2012)

En los últimos años, la técnica de red neuronal artificial se ha desarrollado rápidamente. Se puede utilizar como una herramienta de modelado madura para tratar una gran cantidad de datos que contienen una relación matemática no lineal. (Sbarbaro, 2005).

“Los programas informáticos inteligentes capaces de estimar y predecir estados en el futuro serían útiles como "sensores de software" cuando se trata de bioprocesos caracterizados por incertidumbres y complejidad. Se ha demostrado que la lógica difusa es una herramienta valiosa para tratar con información vaga e incompleta, y para incorporar el conocimiento de expertos humanos en modelos de procesos. Los programas de redes neuronales capaces de aprender de la experiencia pasada son útiles cuando no se

dispone de información matemática exacta sobre el proceso que se está investigando". (Eerikäinen, Linko, Linko, Siimes, & Zhu, 1993)

La capacidad de almacenamiento de la información es uno de los mecanismos más importantes en el aprendizaje de redes neuronales recurrentes, recurrent neural networking (RNN). Juegan un rol crucial en aplicaciones prácticas, como el aprendizaje de secuencias. Con un buen mecanismo de memoria, la historia a largo plazo se puede fusionar con la información actual y, por lo tanto, puede mejorar el aprendizaje de RNN. (Wang, Zhang, Guo, & Zhang, 2017)

Una red backpropagation neuronal fue entrenada para predecir las cargas en los ejes inducidos del transporte cañero, utilizando vehículos industriales, con una red que consta de dos, cuatro u ocho unidades de procesamiento en las capas de entrada, ocultas y de salida, respectivamente. Las entradas a la red eran cargas útiles y cargas por eje remolque vacío. Las salidas correspondían a las cargas de eje trasero del remolque y tractor medidos. (Kanali, 1997).

El módulo de evaluación de desempeño de la cosecha de caña de azúcar, se integra con varias tecnologías, como la red neuronal artificial, la cual realiza un análisis integral, para evaluar el rendimiento de corte, el rendimiento de limpieza y otros factores para la cosecha de caña de azúcar. (Ma, 2002).

El uso de redes neuronales artificiales para ayudar a la mezcla de productos y decisiones de inversión en las centrales azucareras brasileñas, evidenciando que la inversión en proyectos para incrementar la recuperación de energía a partir de residuos puede presentar un aumento de la eficiencia en la compensación riesgo-beneficio de las bio-refinerías. (Mutran, 2018)

La red neuronal artificial se puede aplicar para predecir los rendimientos globales de glucosa en diferentes condiciones operativas tanto de pretratamiento como de hidrólisis enzimática. La confiabilidad del modelo fue evaluada a través de una sensibilidad análisis, que mostró las

condiciones operativas necesarias para mejorar el rendimiento de glucosa: concentración de biomasa inicial relativamente baja y concentración de ácido; alta concentración de enzimas; y 72 h de hidrólisis enzimática. Los resultados experimentales mostraron claramente que la celulosa menos reactiva para la hidrólisis enzimática dependió de la concentración de ácido. (Laura Plazas Tovar, 2017)

Las redes neuronales se han convertido en una herramienta eficaz en el campo azucarero, desarrollando proyectos innovadores en distintas áreas de las industrias azucareras, gestionando procesos innovadores para la eficiencia en los mismos.

El enfoque de la investigación desarrollada es cuantitativo; toma como centro de su proceso la información estadística histórica obtenida de los últimos 30 años, referidas a parámetros climatológicos como: Temperaturas máximas y mínimas, oscilación térmica, precipitaciones pluviales, heliofanía, humedad relativa de la caña de azúcar, evaporación anual, y hectáreas cosechadas por zafra de los cultivos agrícolas de Compañía Azucarera Valdez S.A.(CAVSA.) haciendo énfasis en los últimos 11 años, periodo en el cual los parámetros son homogéneos en cantidad e inversiones realizadas en el sector agrícola de la industria.

Según la homogeneidad y las inversiones realizadas, (Valdés, 2004), menciona que la caña de azúcar es un cultivo que requiere de un estudio preciso de los recursos climáticos y de las condiciones meteorológicas. Está demostrado que las limitaciones fundamentales para el crecimiento y desarrollo de esta planta se deben al componente clima, que generalmente se comporta de forma homogénea.

A nivel sudamericano existen pocos estudios que permitan estimar o diagnosticar objetivamente las toneladas de caña en un campo agrícola, mucho menos modelos probados y validados que se puedan replicar en sectores con similares características agronómicas y/o productivas. Los diferentes factores climáticos que actúan sobre un lugar determinado condicionan en gran medida las fases del ciclo

productivo de la caña y los resultados finales de esta. A cada lugar corresponde un rendimiento máximo dependiente de las condiciones climáticas del año. A la media de esas condiciones climáticas corresponde una media de rendimiento máximo, o rendimiento potencial específico. (Fauconnier, 1975).

II. DESARROLLO

1. Metodología

La investigación se realizó en la ciudad de Milagro, que se encuentra en la zona occidental de la Provincia del Guayas. Coordenadas geográficas de Milagro, Ecuador con una latitud: 2°08'02" S, longitud: 79°35'38" O y altitud sobre el nivel del mar: 14 m, como se muestra en la (figura 1).

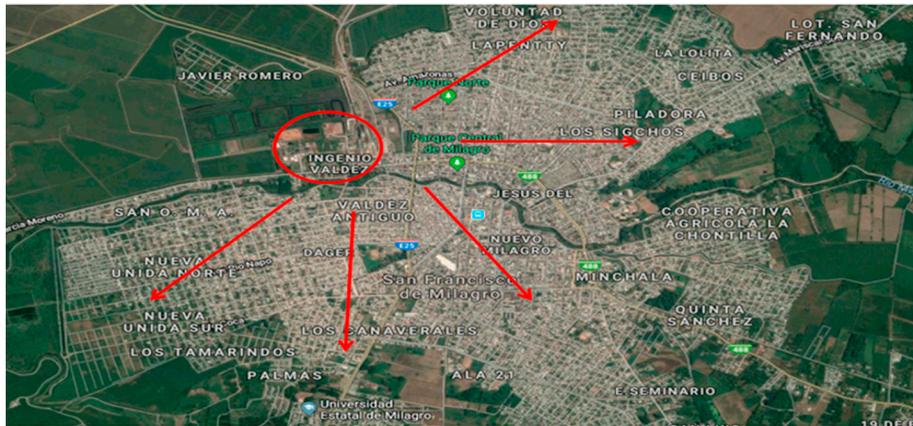


Figura 1: Localización geográfica de Milagro.
Fuente: Instituto Geográfico Militar del Ecuador.

La descripción geográfica de la ciudad está conformada por suelos fértiles, numerosos ríos y esteros, por bosques, plantíos y zonas residenciales; haciendas, fincas y otras propiedades. Al ubicarse en una zona tropical, posee gran biodiversidad y un clima cálido - húmedo todo el año. De acuerdo con los tipos de clima la variedad en los vegetales es

diferentes. Las condiciones de clima y del sustrato varían de un lugar a otro y estas variaciones tienen que reflejarse en la existencia de comunidades vegetales (NOGUEZ, 1993). De acuerdo con la zona donde se realizó el estudio, se tienen los siguientes datos climáticos como se muestra en la Tabla 1 es un clima apto para el cultivo de la caña de azúcar.

Tabla 1. Estadística y parámetros climáticos de la ciudad de Milagro.

Parámetros	Valores
Precipitación Anual	1298,3 mm
Temperatura media	25,2°C
Temperatura máxima	29,8°C
Oscilación térmica	8,9°C
Humedad relativa	80%
Heliofanía anual	1036,5 horas
Evaporación anual	1309,7 mm.
Viento predominante	SW
Velocidad del viento	3,9 km/h
Presión atmosférica	1012,4 mb

Fuente: Instituto Nacional de Meteorología e Hidrología del Ecuador.

La información estadística para el desarrollo de la instigación fue obtenida del archivo agrícola de CAVSA utilizando el utilitario Excel para clasificarla y promediar datos de 30 años de historia, haciendo énfasis en los últimos 11 de años de producción, periodo homogéneo en la cosecha de caña de azúcar tanto propia (de la organización) como la obtenida a través de cañicultores, lo que permitió realizar un análisis comparativo de los parámetros analizados en la investigación.

Durante la investigación se utilizaron como variables: la historia climatológica (datos de los 30 años de producción), fuente de ayuda bibliográfica, fuentes de consulta documental (páginas web, libros, revistas periódicos); información empírica, proveniente de técnicos vinculados con el problema, a los cuales se les aplicaron los instrumentos de investigación de campo. Todo esto permitió obtener, analizar y describir la información relacionada al proyecto de investigación. Se combina lo documental, en cuanto a citas bibliográficas, revistas y la parte de campo, que son los datos de las variables, ambas se relacionan y complementan. Así como menciona Zorrilla (1993), que la investigación mixta es aquella que participa de la naturaleza de la investigación documental y de la investigación de campo.

La investigación es de campo y exploratoria, porque desarrolla un plan piloto con datos climatológicos en la parte del cultivo y cosecha, correlacionan las variables dependientes (predicción de las toneladas de caña en zafra) con las variables independientes (Temperatura mínimas, máximas, oscilación térmica, precipitaciones, heliofanía, humedad relativa, evaporación anual y hectáreas de caña sembradas). La investigación es explicativa porque busca la razón de los hechos, estableciendo relaciones de causa- efecto. (Sampieri, 2010).

De acuerdo con las variables (independientes y dependiente), tomadas para la investigación concuerdan con lo mencionado. Subiros (1995), menciona que las variables independientes a analizar son temperatura ambiente máxima,

temperatura ambiente mínima, precipitación y radiación solar ya que estos figuran entre los principales agentes climatológicos que afectan y determinan la producción cañera.

2. Método

El procesamiento de la información se realizó mediante una red neuronal artificial. MATICH (2001) y FREEMAN (1993), mencionan que una red neuronal es un modelo computacional que pretende simular el funcionamiento del cerebro a partir del desarrollo de una arquitectura que toma rasgos del funcionamiento de este órgano sin llegar a desarrollar una réplica del mismo. El cerebro puede ser visto como un equipo integrado por aproximadamente 10 billones de elementos de procesamiento (neuronas) cuya velocidad de cálculo es lenta, pero que trabajan en paralelo y con este paralelismo logran alcanzar una alta potencia de procesamiento.

El primer método usado para la investigación es el Método Hipotético-Deductivo, donde el investigador propone una hipótesis como consecuencia de sus inferencias del conjunto de datos empíricos o de principios y leyes más generales. La hipótesis utiliza procedimientos inductivos y procedimientos deductivos. Es la vía primera de inferencias lógico deductivo para arribar a conclusiones particulares a partir de la hipótesis y que después se puedan comprobar experimentalmente. (Chagoya, 2008).

El segundo es el Método Descriptivo siendo un auxiliar del científico, es imprescindible, pues permitió describir los procesos que se aplican para desarrollar cada uno de los objetivos propuestos en la presente investigación. Según Tamayo (1998), en su libro *Proceso de Investigación Científica*, la investigación descriptiva “comprende la descripción, registro, análisis e interpretación de la naturaleza actual, y la composición o proceso de los fenómenos. El enfoque se hace sobre conclusiones dominantes o sobre grupo de personas, grupo o cosas, se conduce o funciona en presente”.

El tercero es el Método Analítico. Este método fue fundamental ya que permitió realizar una

adecuada recopilación análisis e interpretación de la información que se representa mediante cuadros estadísticos, pudiendo así, tener la información clasificada con sus respectivas variables. Según Sánchez (1990), lo define como aquel “que distingue las partes de un todo y procede a la revisión ordenada de cada uno de los elementos por separado “Este método es útil cuando se llevan a cabo trabajos de investigación documental, que consiste en revisar en forma separada todo el acopio del material necesario para la investigación.

3. Técnica

La técnica que se empleó en la investigación se divide en dos etapas: la primera en la recolección y adecuación de la información histórica obtenida durante los últimos 30 años, y procesamiento de los datos en la hoja de Excel. La segunda procesa la información obtenida de los históricos como entradas para inferir una salida a través de una red neuronal artificial utilizando un software Matlab.

En la primera etapa se validaron los resultados asegurando que la información obtenida, correspondía a los últimos 30 años de las variables independientes, apreciándose los cambios y variaciones climatológicas, durante las épocas

de cosecha de la caña, estos datos se clasificaron de acuerdo con el impacto de crecimiento y afectación del cultivo en los campos de siembra. En este sentido, la caña de azúcar posee un período vegetativo muy variable, cuya duración depende básicamente de las características del material genético utilizado y de la influencia de factores limitantes agroclimáticos que ejercen en su proceso biológico (Miceli, 2002).

Para obtener respuestas de una red neuronal que sea coherente con los valores de entrada, es necesario una adecuada selección de la arquitectura de la red a utilizar, así como del algoritmo de aprendizaje que se ejecutará.

Para la segunda etapa se construyó una tabla con sus respectivas variables dependientes e independientes, que contienen valores de temperaturas máximas y mínimas del terreno, oscilación térmica, precipitaciones, heliofanía, humedad, evaporación y hectáreas de caña sembradas, quedando la matriz conformada por doce filas (años de cosecha) y por nueve columnas que registran los parámetros que influyen en la calidad de la caña de azúcar, como se muestran en la Tabla 2 generando un archivo con formato de Excel (xlsx).

Tabla 2. Información Estadística de Cosecha CAVSA

AÑO	TEM MAX [°C]	TEM MIN [°C]	OSC [°C]	PRECIP [mm]	HELOFANIA [Horas]	HUMEDAD [%]	EVAPORA [mm]	HECTAREAS [Ha]	TONELADAS [Tn]
2018	29,80	22,20	7,90	969,10	643,90	79,00	1.115,30	24.023,00	0,00
2017	30,42	22,84	7,60	2.232,00	733,20	81,00	1.184,40	22.518,02	1.796.591,86
2016	30,63	22,93	7,70	1.318,80	861,60	81,00	1.212,00	23.434,65	2.020.471,99
2015	30,60	23,40	7,20	1.244,40	807,60	80,00	1.177,20	16.862,25	1.786.515,35
2014	30,14	22,60	7,50	1.161,60	722,40	80,00	1.171,20	19.247,16	1.845.955,33
2013	26,36	21,89	8,00	1.030,80	703,20	80,00	1.186,80	20.804,62	1.720.344,91
2012	25,75	22,08	8,20	2.078,40	876,00	79,00	1.118,40	18.692,12	1.629.832,97
2011	26,90	21,99	8,10	912,00	910,80	78,00	1.220,40	19.776,46	1.627.735,15
2010	29,82	22,10	7,70	1.348,80	681,60	82,00	1.117,20	18.198,48	1.423.312,37
2009	30,17	22,04	8,10	1.137,60	976,80	79,00	1.327,20	20.962,94	1.336.962,60
2008	29,68	21,97	7,70	2.086,80	856,80	80,00	1.244,40	20.295,62	1.433.157,11
2007	30,63	21,67	8,10	979,20	856,80	80,00	1.266,00	18.419,51	1.480.263,50
MAX	30,63	23,40	8,20	2.232,00	976,80	82,00	1.327,20	24.023,00	2.020.471,99
MIN	25,75	21,67	7,20	912,00	643,90	78,00	1.115,30	16.862,25	1.336.962,60

Fuente: Elaboración del autor.

Fue necesario normalizar la información obtenida en la Tabla 2, para trabajar con el software Matlab en el proceso de entrenamiento, para dicho procedimiento se utilizó la siguiente formula. Tabla 3:

Tabla 3. Formula de Normalización basada en la unidad.

$X' = \frac{X - X_{\min}}{X_{\max} - X_{\min}}$	X'	Valor normalizado
	X	Valor por normalizar
	X _{mín}	Valor mínimo de la muestra
	X _{máx}	Valor máximo de la muestra

Fuente: Elaboración del autor.

Donde se especifican los límites superiores e inferiores del valor de salida y luego se obtienen los valores máximos y mínimos de cada variable (tomando los datos de entradas y salidas de los 11 últimos años). Por último, se normaliza cada elemento de entrada y salida y se lo guarda para uso posterior.

La metodología trae todos los valores al rango

entre [0,1] números binarios. Denominada, normalización basada en la unidad, acorde a los datos de máximos y mínimos. Obtenidos los datos normalizados de acuerdo con la formulación aplicada, se procede a realizar una nueva tabla para la agrupación de las variables dependientes e independientes dentro del software. Como se presentan en la Tabla 4.

Tabla 4. Datos Normalizados de cosecha CAVSA

AÑO	2017	2016	2015	2014	2013	2012	2011	2010	2009	2008	2007
TEM MAX [°C]	0,9570	1,0000	0,9939	0,8996	0,1250	0,0000	0,2357	0,8340	0,9057	0,8053	0,8156
TEM MIN [°C]	0,6763	0,7283	0,0000	0,5376	0,1272	0,2370	0,1850	0,2486	0,2139	0,1734	0,0000
OSC [°C]	0,4000	0,5000	0,0000	0,3000	0,8000	1,0000	0,9000	0,5000	0,9000	0,5000	0,9000
PRECIP [mm]	0,9542	0,3388	0,2887	0,2329	0,1447	0,8507	0,0647	0,3590	0,2167	0,8563	0,1100
HELOFANIA [Horas]	0,1621	0,5653	0,3957	0,1281	0,0678	0,6106	0,7198	0,0000	0,9271	0,5503	0,5503
HUMEDAD [%]	0,7500	0,7500	0,5000	0,5000	0,5000	0,2500	0,0000	1,0000	0,2500	0,5000	0,5000
EVAPORA [mm]	0,2376	0,3352	0,2122	0,1909	0,2461	0,0042	0,3649	0,0000	0,7426	0,4498	0,5262
HECTAREAS [Ha]	0,6189	0,7193	0,0000	0,2610	0,4314	0,2003	0,3178	0,1462	0,4488	0,3757	0,1704
TONELADAS [Tn]	0,3952	0,5877	0,3865	0,4376	0,3296	0,2518	0,2500	0,0742	0,0000	0,0827	0,1232

Fuente: Elaboración del autor.

Los pasos siguientes son fundamentales para el tratamiento de la información en el software

1. Se ingresa en el módulo de importación del archivo (Excel) en Current Folder, indicando el tipo de archivo (xlsx), en el cual se van a generar los valores de entrada y salida de la neurona que se va a interactuar.
2. Una vez ingresada la información se pasa al módulo de construcción de la RNA, se construyen las matrices en el Workspace,

en la cual se define la entrada y salida a la red, indicando que las entradas fueron las variables independientes (Temperatura mínimas, máximas, oscilación térmica, precipitaciones, heliofania, humedad, evaporación y hectáreas de caña sembradas) y la variable dependiente (toneladas de caña).

Aceptados los datos, se seccionaron en tres matrices que fueron: IN (archivo de entrenamiento),

OUT (archivo de salida) y S2018 (archivo de prueba o simulación), para ser procesados según las fases siguientes.

En la primera fase se crean las matrices de entrada y salida, y generara la neurona con sus parámetros, los cuales incluyen el modo de entrenamiento y el número de capas que se decidió tomar para minimizar los errores. De acuerdo con la tipología que describe SARLE (1994), quien menciona que las RNA son una amplia gama de regresiones no lineales, modelos discriminantes, modelos de reducción de datos y sistemas dinámicos no lineales.

La segunda fase así mismo se compone de la topología o arquitectura de las redes neuronales la cual se basa en la administración y colocación de las neuronas en la red para formar grupos de

neuronas o capas que no se encuentren cercanas de la entrada ni a la salida de la red. En base a esta dirección, los parámetros fundamentales de la red son: el número de capas, el número de neuronas por capa, el grado de conectividad y el tipo de conexiones entre neuronas. (SALAS, 2000). De acuerdo con la (figura 2), pueden apreciarse los parámetros de la neurona.

En la tercera fase se selecciona el algoritmo de aprendizaje que actualmente está más extendido, el algoritmo o regla BackPropagation, es una generalización de la regla LMS (Least Mean Square); por tanto, también se basa en la corrección del error. Básicamente el proceso BackPropagation consiste en dos pasadas a través de las diferentes capas de la red, una pasada hacia adelante y una pasada hacia atrás. (Coello, 2015).

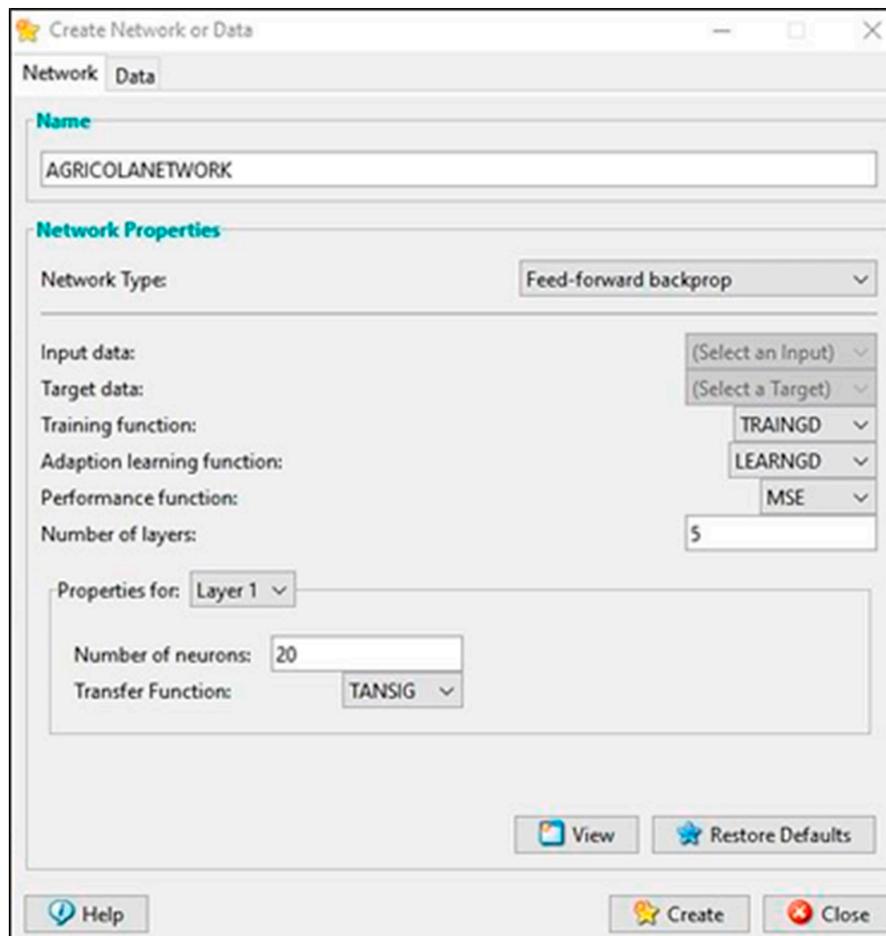


Figura 2: Parámetros de la Red Neuronal.

Fuente: Elaboración del autor.

En la ventana de CREATE NETWORK OF DATA, se adjuntan y caracteriza la red que se desea establecer, para dicha investigación se seleccionó en NETWORK TYPE la opción de FREED-FORWARD BACKPROP que permite, que el margen de error sea propagado hacia atrás desde la capa de salida. Haciendo que el margen sea mínimo y así la red pueda ser más eficaz al momento de la predicción.

En la investigación establecida, la red es de multicapas. Basualdo (2001), menciona que las multicapas son las que colocan una serie de neuronas ordenadas en distintos niveles. La forma en que se puede distinguir la capa en la que se encuentra alguna neurona consiste en fijar punto de partida de las señales que recibe a la entrada y el punto de llegada de la señal de salida. Por lo general, todas las neuronas de una capa reciben señales de entrada por parte de la capa más cercana a la entrada de la red, es decir, de una capa anterior, y envían señales a la capa que se encuentra más cercana a la salida de la red, debido a esto, una capa posterior, a estas conexiones son denominadas conexiones feedforward.

La función de entrenamiento será Tangencial-Sigmoidea, cuyos valores de salida que proporciona esta función están comprendidos dentro de un rango que va de 0 a 1. Al modificar el valor de g se ve afectada la pendiente de la función de activación.

La neurona creada contiene cuatro capas ocultas, de doce neuronas cada una. Así mismo su entrenamiento como se mencionó en el párrafo anterior es Tangente Sigmoidea, como describe Tsoukalas L.H. y Uhrig R.E (1997) quienes acreditan que la potencia de las redes neuronales reside en que los nodos individuales se encuentran en diferentes capas, formando redes altamente interconectadas, con una arquitectura inspirada en la corteza cerebral, que permiten el aprendizaje de patrones no lineales de comportamiento

El software nos permite tener una imagen detallada de las descripciones numéricas y la apreciación grafica de la red como se muestra en la (figura 3). La cual puede reconocer su funcionamiento y la transición de los datos por cada capa oculta que la contiene, hasta un resultado analítico final.

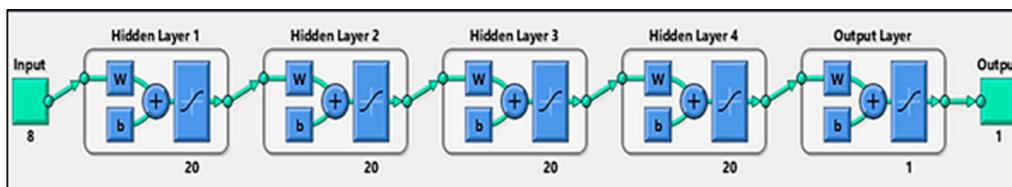


Figura 3: Red Neuronal "RED AGRICOLA."
Fuente: Elaboración del autor.

Para entrenar la red hay que deslizarse en la parte de TRAIN, que contiene parámetros analíticos los cuales se pueden modificar y subscribir el número de interacciones que se desea establecer para entrenar la Red. Existen siete parámetros asociados con este tipo de entrenamiento, son:

- Ephocs: Define el máximo número de épocas de entrenamiento que puede tener nuestro proceso de aprendizaje
- Show: Indica a Matlab la forma de visualización que deseamos tener durante el entrenamiento de la red. Si su valor es Nan quiere decir que no se quiere ningún tipo de visualización.
- Goal: Este variable indica un valor mínimo límite que puede alcanzar la función error de la red. Si esta alcanza dicho valor el entrenamiento se parará automáticamente.
- Time: Este parámetro indica el tiempo máximo en segundos que puede durar el entrenamiento de la red. Una vez que el tiempo del proceso alcance dicho valor el entrenamiento se detendrá.
- Min_grad: Determina el valor mínimo necesario que debe tener el gradiente para detener el algoritmo.
- Max_fail: Es el máximo número de iteraciones que puede incrementarse el error de validación antes de detenerse el

entrenamiento.

- Lr: Es la ratio de aprendizaje α .

Parámetros en que se entrenó la red, en una cantidad de 15,000 Epochs (Define el máximo

número de periodos de entrenamiento que puede tener nuestro proceso de aprendizaje) y con la misma cantidad en `max_fail` (Es el máximo número de iteraciones que puede incrementarse el error de validación antes de detenerse el entrenamiento), como se muestra en la (figura 4).

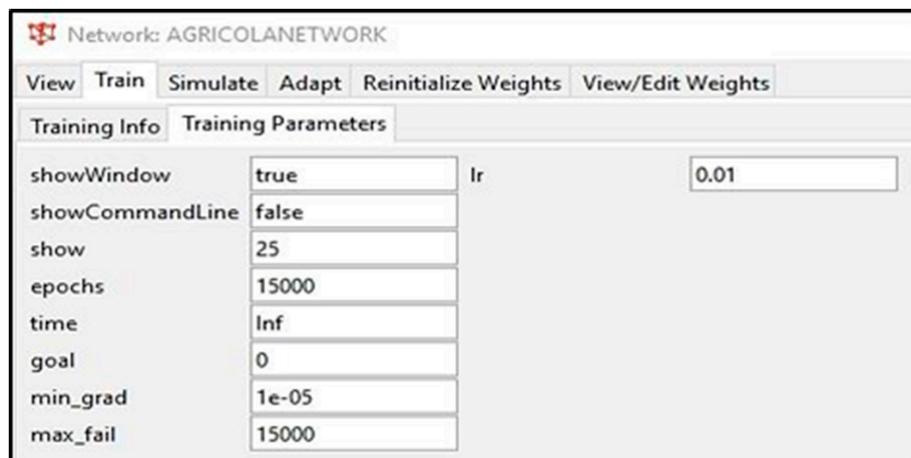


Figura 4: Parámetros de entrenamiento RED AGRICOLA.

Fuente: Elaboración del autor.

De acuerdo con los parámetros establecidos, se entrena la red de manera supervisada, para obtener un patrón de aprendizaje con los datos establecidos. De tal forma que al realizar las interacciones se puedan obtener tendencias en cuanto a los resultados obtenidos. Esta forma de entrenar se la relaciona con la ramificación de las neuronas dentro del cerebro humano, como mencionan Y.Shachmurove (2002) y Tkacz (1999), que las redes neuronales artificiales son sistemas de procesamiento de información, desarrolladas por científicos cognitivos con el propósito de entender el sistema nervioso biológico e imitar los métodos computacionales del cerebro y su impresionante habilidad para reconocer patrones.

III. RESULTADOS

Con base en los objetivos planteados y a la metodología utilizada se obtuvieron los siguientes resultados.

Fase de entrenamiento

La arquitectura final de la red neuronal, que dio mejor resultado, después de probar varios algoritmos de entrenamiento y diferentes números

de capas, quedó conformada por cuatro capas; doce neuronas en la capa de entrada, doce neuronas en la primera capa oculta, doce neuronas en la segunda capa oculta y una neurona en la capa de salida. Cada una de estas capas contiene una función de activación, (Tansig) para las capas ocultas.

Para esta etapa de la red se utilizó el algoritmo de retro propagación (Feed- forward backprop), dicho proceso se corrió sobre los datos de entrenamiento que constan de 99 datos obtenidos de los registros climáticos y de cosecha de la empresa durante treinta años, de los cuales los últimos once años han sido homogéneos en cuanto a las inversiones y reformas tecnológicas desarrolladas dentro de la empresa durante el periodo analizado, al final se realiza la verificación y predicción con un vector de entrada de ocho variables (unidades o neuronas).

La red neuronal encontró los pesos adecuados para la generalización de patrones hacia toda la población, observándose que las variables con mayor relevancia fueron en primer lugar: las variaciones de temperaturas, hectáreas de caña sembradas, seguido de la humedad y

consecuentemente las precipitaciones.

Se desarrolló como presentación de la red una interfaz, utilizando las herramientas de Matlab, con un lenguaje de programación propio del software.

Dicha ayuda permitió interactuar con la red de manera amigable sin necesidad de abrir el programa, para ingresar los datos en las distintivas variables dependientes, como se muestra en la (figura 5).



Figura 5: Interfaz Red Neuronal “AGRICOLANETWORK”.
Fuente: Elaboración del autor.

Se procedió de manera sistemática a subir la información de las variables, a través del sistema de entrenamiento propio de Matlab, se obtuvo el aprendizaje de la red, obteniendo como resultado las toneladas de caña de los últimos once años de forma predictiva, comparándolo con los datos reales de las toneladas de caña del periodo

investigado.

Se observa que los datos reales y de la red neuronal, son similares y su margen de error está dentro de lo permisible. Corroborando de manera analítica que las Redes Neuronales Artificiales, son herramientas eficaces para la industria azucarera. Ver (figura 6).

Mendoza y Marquetti. Redes Neuronales Artificiales: factores que determinan la cosecha de caña en la industria.

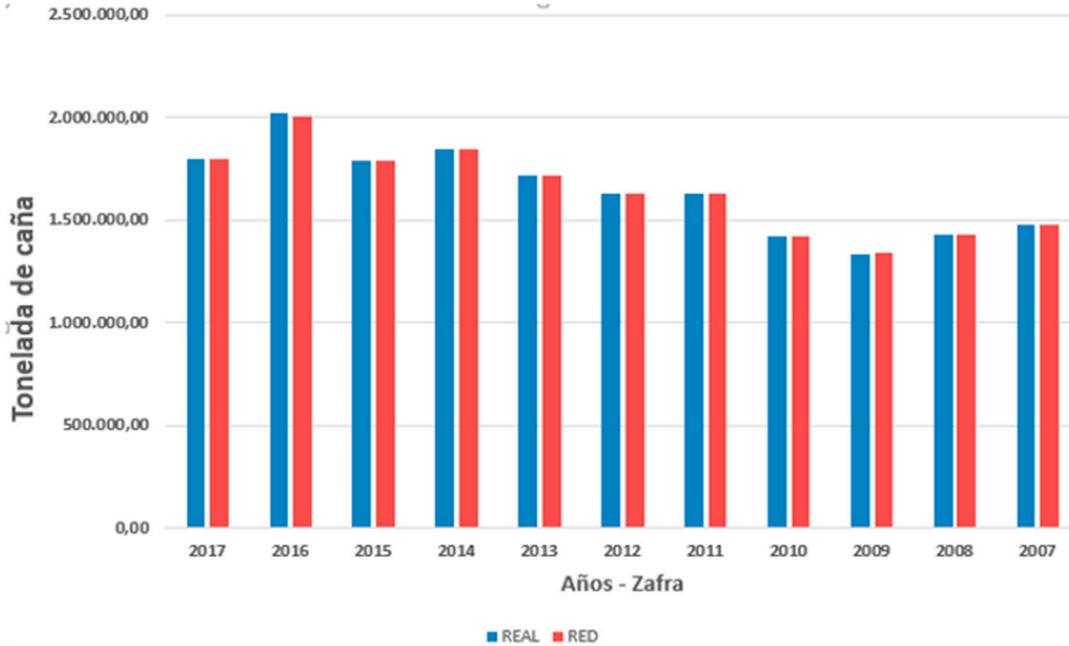


Figura 6: Gráfico comparativo Real Vs Red de la producción de caña en CAVSA, últimos 11 años.
Fuente: Elaboración del autor.



Figura 7: Gráfico Error comparativo Real Vs Red de la producción de caña en CAVSA, últimos 11 años.
Fuente: Elaboración del autor.

De acuerdo con la efectividad de la red neuronal, se lograron predecir para el año 2018 las toneladas de caña que se van a obtener, comparando con

la predicción-Aforo, lo cual se comprueba la magnitud de predicción. Ver la (figura6-7).

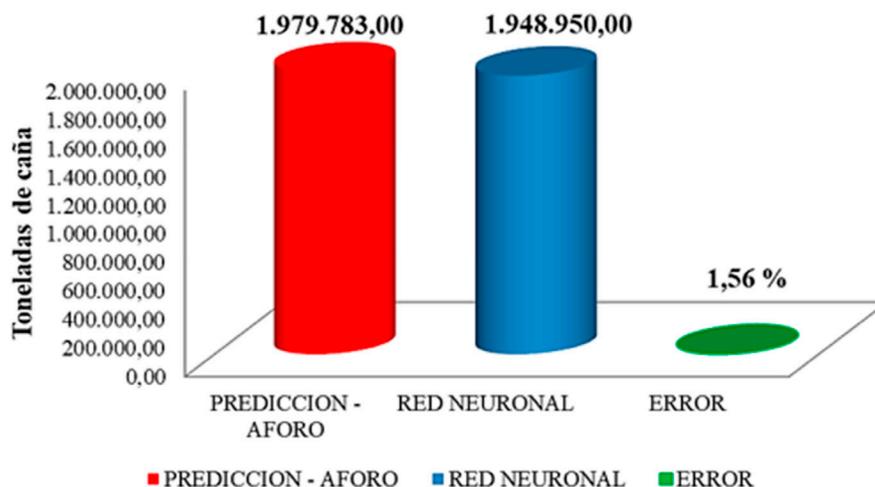


Figura 8: Predicción año 2018.
Fuente: Elaboración del autor.

IV. CONCLUSIONES

La creación de la red neuronal artificial y su entrenamiento para predecir la cosecha de la caña de azúcar en CAVSA, a través de datos climatológicos históricos de los cultivos agrícolas de caña de azúcar, se constató que el valor final predictivo arrojado por la red, es bastante aproximado al resultado obtenido por el método tradicional de aforo, realizados por los técnicos del departamento agrícola de cultivo y cosecha para el año 2018, obteniendo un resultado satisfactorio con un margen de error de 1,56%.

Comprobando la hipótesis con dicho resultado, las redes neuronales son herramientas de predicción eficientes en cuanto a su funcionabilidad ya que garantizan resultados para tomarlos como referentes en temas de planificación de presupuestos y productividad en una organización agroindustrial como CAVSA.

Matlab es un software matemático diseñado como “traje a la medida” para aplicaciones como la investigada y trae en su configuración un Interfaz GUIDE que es amigable y permite interactuar con las variables de entrada de la Red Neuronal

entrenada, obteniéndose resultados eficaces para analizarlos y cuestionarlos.

Durante el trabajo de investigación a través de la interfaz se llegó a determinar que las variables climatológicas como: temperatura máxima, temperatura mínima, y Precipitaciones afectan positivamente al rendimiento de la caña de azúcar aumentando las toneladas de producción, mientras que la Heliofanía por su parte tiene un efecto negativo en el rendimiento cañero, análisis realizado individualmente los parámetros mencionados. Si se analizan las variables de entrada en forma de conjunto, se puede concluir que ellas están relacionadas entre sí y sería motivo de otra investigación profundizar en su estudio.

Alcanzan niveles de eficacia por encima del 90% evidenciándose que se mantienen los resultados del modelo acordes a los de expertos humanos. (López, 2016).

V. RECOMENDACIONES

La investigación puede ser tomada como base para experimentar e involucrar los procesos fabriles, para predecir la producción de azúcar

y sus derivados, misma que servirá como guía para presupuestar la cantidad de caña cosechada y producción final de azúcar por zafras, que se destinarían como materia prima para producir azúcar, sus derivados y alcohol, producto de actualidad con el tema de los biocombustibles.

VI. REFERENCIAS

- ARIAS. (2004). El proyecto de Investigación, Guía para su Elaboración. CARACAS,. VENEZUELA .
- Avila, R., Rodriguez, V., & Hernandez, E. (2012). Predicción de rendimiento de un cultivo de platano mediante redes neuronales de regresión generalizada. Publicaciones en Ciencias y Tecnología, VI(1), 31-40.
- Banko, C. (2005). La industria azucarera en México y Venezuela. Un estudio comparativo. Carta Económica Regional, 41-54.
- Chagoya, E. R. (2008). Métodos y técnicas de investigación. Gestipolis, 16.
- Coello, L. (4 de FEBRERO de 2015). Redes neuronales artificiales en la producción de tecnología. Discover scientific knowledge. Obtenido de https://www.researchgate.net/publication/283847824_Rednes_neuronales_artificiales_en_la_produccion_de_tecnologia
- Eerikäinen, T., Linko, P., Linko, S., Siimes, T., & Zhu, Y.-H. (1993). Lógica difusa y aplicaciones de redes neuronales en la ciencia y tecnología de los alimentos. Elsevier Ltd, IV, 237-242.
- Fauconnier, R. (1975). La caña de azúcar. Técnicas agrícolas y producciones tropicales. Editorial Blume., 405.
- FREEMAN, J. A. (1993). Redes neuronales: algoritmos, aplicaciones y técnicas de programación, . New York, USA,: Addison-Wesley Iberoamerican.
- Kanali. (30 de julio de 1997). Predicción de cargas por eje inducidas por la caña de azúcar utilizando vehículos industriales. 2-7.
- Laura Plazas Tovar. (2017). Prediction of overall glucose yield in hydrolysis of pretreated sugarcane bagasse using a single artificial neural network: good insight for process development.
- López, S. T. (2016). Red neuronal multicapa para la evaluación de competencias laborales. Revista Cubana de Ciencias Informáticas. Obtenido de http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S2227-18992016000500016
- Ma, F. L. (2002). Knowledge Acquisition Based on Neural Networks for Performance Evaluation of Sugarcane Harvester. 4-7.
- Marquetti, H. (13 de DICIEMBRE de 2016). El desarrollo económico local en Cuba. La Habana, Cuba. Obtenido de <file:///C:/Users/Usuario/Downloads/Desarrollo-EconomicoLocalCuba.pdf>
- MATICH, D. J. (2001). Redes neuronales, Conceptos básicos y aplicaciones. Mexico: Ed. Universidad Tecnológica Nacional, .
- Miceli, G. (2002). Regulación enzimática de la acumulación de sacarosa en cañas de azúcar (Saccharum spp.). Agrobiencia, 411-419.
- Mutran, V. M. (2018). Bioenergy investments in sugarcane mills: an approach combining portfolio theory with neural networks. 3-7.
- NOGUEZ, H. R. (1993). Comparación entre imágenes Landsat Thematic. Tesis de licenciatura. División de Ciencias Forestales, Universidad Autónoma Chapingo, 115.
- SALAS, M. I. (2000). Redes neuronales artificiales en la medición de temperatura y humedad relativa. Mexico: Universidad Autónoma Chapingo.
- Sampieri, D. R. (2010). Metodología de la investigación. México D.F: McGRAW-HILL / INTERAMERICANA EDITORES, S.A. DE C.V.

- Sbarbaro, D. (2005). Supervisión de válvulas de control en circuitos de flotación basados en redes neuronales artificiales. ICANN'05 Proceedings of the 15th international conference on Artificial neural networks: formal models and their applications, Volume Part II, págs. 451- 456. Springer , Berlín, Heidelberg, alemán.
- Subiros, F. (1995). El cultivo de Caña de Azúcar. Editorial Universidad Estatal a Distancia., 441.
- Valdés, M. (2004). Determinación del periodo de crecimiento en el cultivo de la caña de azúcar. CAI Jose Martí, 1562-3297.
- Wang, J., Zhang, L., Guo, Q., & Zhang, Y. (21 de Marzo de 2017). Redes neuronales recurrentes con unidades de memoria auxiliar. IEEE Transactions on Neural Networks and Learning Systems, 1652 - 1661.
- Zhang, Y. (2018). Predicción del grado de concentrado de fosfato basado en el modelado de redes neuronales artificiales., (págs. 625-628).
- Zorrilla, A. (1993). Introducción a la metodología de la investigación. Aguilar Leon y Cal, Editores, 11ª Edición., 43.