

TRABAJO DE GRADO

Opción Seminario-Diplomado.

**ALGORITMO COMPUTACIONAL PARA EL ANÁLISIS Y TOMA DE
DECISIONES, PARA EL ALMACENAMIENTO DE GRANO EN LA COSECHA,
UTILIZANDO ESTRATEGIAS DE MACHINE LEARNING**

Corporación Universitaria Remington.

FACULTAD INGENIERIA
INGENIERÍA INDUSTRIAL

Estudiante:

Jaime Eduardo Torres García

Tutor: Juan Carlos Briñez de León

Opción de Trabajo de grado Seminario-Diplomado.

2024.

Dedicatoria.

A mi novia y mi familia quienes han podido ver mi proceso de estudio, y han sido mi motor, porque sin ellos no tendría sentido.

Agradecimientos.

Primero a Dios quien nos da la vida con su infinita misericordia, a todos mis allegados de la Universidad y aquella persona que comparte sus conocimientos con el fin de enseñar y causar un impacto importante en nuestras vidas “el profesor del seminario”.

Tabla de contenido

Marco conceptual y contextual.	6
Pregunta problematizadora:	8
Acercamiento de datos.	8
Objetivos:	14
Procesamiento de datos.	17
Modelo de toma de decisiones.	20
Conclusiones.	22
Referencias.	23

Palabras clave.

- Producción.
- Cosecha.
- Logística.
- Análisis de datos.
- Procesamiento.
- Machine Learning.

Marco conceptual y contextual.

La transformación digital ha emergido como un tópico de interés en la actualidad, dentro de las comunidades científicas, y las organizaciones que la establecen como una práctica recurrente, que se encuentran sometidas a los nuevos modelos de negocio, así como las demandas personalizadas de los usuarios en un entorno de hiperconectividad.

Llegar a entender la transformación digital, sus componentes y los posibles caminos para adoptarlas, además de sus impactos. Es una necesidad a la cual en la actualidad estamos sujetos. (Fernández, T. D. (2020). *Taxonomía de transformación digital. Revista Cubana de transformación digital, 1(1), 4-23.*).

En los últimos veinte años, el país ha venido tomando una postura más seria, con respecto a querer incursionar en el mercado internacional, mediante la firma de tratados de libre comercio. Aunque este sector se ha visto muy afectado, su bienestar es muy importante, a lo cual se ha vinculado activamente, puesto que el arroz es el tercer producto más importante del país.

En un estudio realizado, se pudo evidenciar los diferentes factores económicos y no económicos, teniendo como objetivo principal estudiar el impacto del TLC en la producción del país además de su sostenimiento y competitividad. Este estudio compara la industria colombiana con la estadounidense, llegando a evidenciar el avance tecnológico, mercantil y de manejo político más eficiente por parte del país americano.

Esto quiere decir que el impacto del TLC con Estados Unidos en el mercado del arroz, es negativo para la producción nacional por que así el agricultor tendrá una desventaja competitiva, ya que los precios son demasiado altos con respecto a las importaciones de dicho producto desde Estados Unidos (Chica, J., Tirado, Y. C., & Barreto, J. M. (2016). *Indicadores de competitividad del cultivo del arroz en Colombia y Estados Unidos. Revista de Ciencias Agrícolas*, 33(2), 16-31.).

En cuanto a la logística de almacenamiento, podemos ver que conforme pasan los años y evoluciona el fenómeno de la logística, se ha ido cambiando el concepto ampliando el ámbito de la responsabilidad que repercute. El almacenamiento se caracteriza por ser una unidad infraestructural del proceso de producción de cualquier industria, además de un soporte muy importante para la eficiencia de este. (APAZA PACO, J. P., CHÁVEZ LIZÁRRAGA, G. A., & HERRERA CHOQUE, V. A. N. I. A. (2015). *Logística de almacenamiento de materia prima en la industria farmacéutica. Revista Con-ciencia*, 3(1), 101-113.).

La automatización en los procesos industriales, se plantea desde la necesidad de las industrias en el incesante avance de la competitividad y eficiencia del mercado. Este básicamente nos habla de la incorporación de elementos tecnológicos eficientes que logren un impacto notable, además de asegurar el control y calidad de los mismos procesos, para ayudar a prever posibles situaciones de las cuales pueda disponer de todos los recursos a su alcance. (Moreno, E. G. (2001). *Automatización de procesos industriales. Valencia: Alfaomega*.).

Pregunta problematizadora:

¿cómo desarrollar una estrategia computacional que ayude a la predicción del comportamiento de la cosecha de arroz, y poder tener un estimado, para así tomar los ajustes necesarios en cuanto a la capacidad de almacenamiento de la planta, teniendo en cuenta los datos del comportamiento de los últimos cinco días, y predecir los cinco días futuros, sabiendo que este periodo de cosecha es de tres meses aproximadamente en el año, planteando algoritmos de Machine Learning?

Acercamiento de datos.

Esta base de datos, se basó en un consolidado de datos tomado de las planillas de entrada de carga o grano el cual recibe diariamente el molino (**Granos y cereales de Colombia**) el cual se dedica a la compra, procesamiento y comercialización del arroz en la región de los llanos orientales, explícitamente en la ciudad de Yopal, departamento del Casanare.

Esta empresa pertenece a un grupo empresarial aliado a los almacenes de cadena Olímpica, donde se vende este producto, además de sus diferentes marcas y presentaciones distribuidas en la costa del Caribe Colombiano.

Estos datos se toman diariamente dependiendo de la cantidad de toneladas de arroz que entran a diario, para tener un control de los granos y poder así tener un aproximado de grano que debe ser procesado y comercializado. Además de poder tomar

decisiones de los espacios donde se deben almacenar también dependiendo de la capacidad de almacenamiento con la que se cuenta en la planta.

Se debe tener en cuenta que este proceso de recibo de arroz, lo que se llama coloquialmente la cosecha es de solo tres meses aproximadamente.

Los datos fueron obtenidos de un coordinador del área de recibo, quien nos pudo suministrar la información correspondiente de los días en los cuales se recibió grano.

Descripción de las variables.

Se tomo una tabla, la cual relaciona la carga del grano día a día durante los tres meses de cosecha, además del porcentaje de humedad con el que llega el grano, teniendo en cuenta que este es de entre 14 a 15, para así poder saber dónde es más optimo secarlo y poder almacenarlo, llegando a obtener las siguientes variables.

- Periodo, lo que se relaciona con el día de llegada de cargue.
- Cantidad, lo que representa el peso de la carga.
- Humedad, nos representa el porcentaje de humedad con la que llega el grano para poder ser procesado.

Posibles aplicaciones.

- Poder predecir la carga que llegara en los días posteriores.
- Llegar a obtener una tendencia de los posibles porcentajes de humedad que puedan llegar.
- Llegar a tomar decisiones de los tiempos en los cuales se tomará el secado de arroz dependiendo del porcentaje de humedad con el que llega la carga.
- Llegar a tomar la decisión de donde poder almacenar, y cuánto tiempo tener el grano en reposo dentro de los silos metálicos.

Aproximación de los gráficos.

	PERIODO	CANTIDAD(TONELADAS)	PORCENTAJE DE HUMEDAD(%)
0	1/05/2024	350	14
1	2/05/2024	350	14
2	3/05/2024	350	14
3	4/05/2024	350	14
4	5/05/2024	450	14
5	6/05/2024	450	14
6	7/05/2024	450	14
7	8/05/2024	450	14
8	9/05/2024	450	15
9	10/05/2024	450	15
10	11/05/2024	450	15
11	12/05/2024	450	15

Ilustración 1 Análisis del consolidado.

En esta ilustración 1 podemos apreciar las fechas, las cantidades de grano que llegan en toneladas cada día y el porcentaje de humedad promedio que presentaba cada carga del día.

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 74 entries, 0 to 73
Data columns (total 3 columns):
#   Column                                Non-Null Count  Dtype
---  -
0   PERIODO                                74 non-null     object
1   CANTIDAD(TONELADAS)                   74 non-null     int64
2   PORCENTAJE DE HUMEDAD(%)              74 non-null     int64
dtypes: int64(2), object(1)
memory usage: 1.9+ KB
```

Ilustración 2 Análisis de las variables.

Acá en esta ilustración 2 vemos las características de cada variable, donde vemos que la variable periodo tiene caracteres diferentes a los numéricos. Además, que el consolidado tiene 74 datos y tres columnas.

	CANTIDAD(TONELADAS)	PORCENTAJE DE HUMEDAD(%)
count	74.000000	74.000000
mean	387.972973	15.216216
std	115.176811	1.252617
min	220.000000	14.000000
25%	322.500000	14.000000
50%	380.000000	15.000000
75%	450.000000	16.750000
max	670.000000	17.000000

Ilustración 3 Resumen del consolidado.

Podemos ver en esta ilustración 3, el resumen del consolidado, en donde apreciamos que el porcentaje máximo con el que llegaron las cargas fue de 17%, además que las toneladas que ha recibido máximo en un día la planta es de 670.

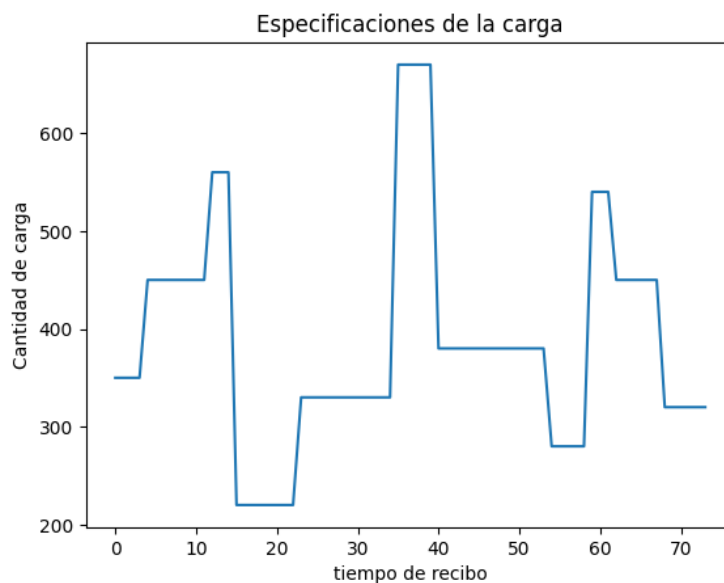


Ilustración 4 Variación de llegada del grano en los días.

En esta ilustración 4 podemos ver que hay un periodo en el que se presentó baja llegada del grano, el cual fue entre el día 15 y 23 aproximadamente dentro del primer mes de la cosecha, lo cual pudo haber sido debido a las lluvias presentadas, así como podemos apreciar que en un periodo entre el día 35 y 40 aproximadamente se presentó el mayor recibo del grano en lo que va de tiempo de cosecha.

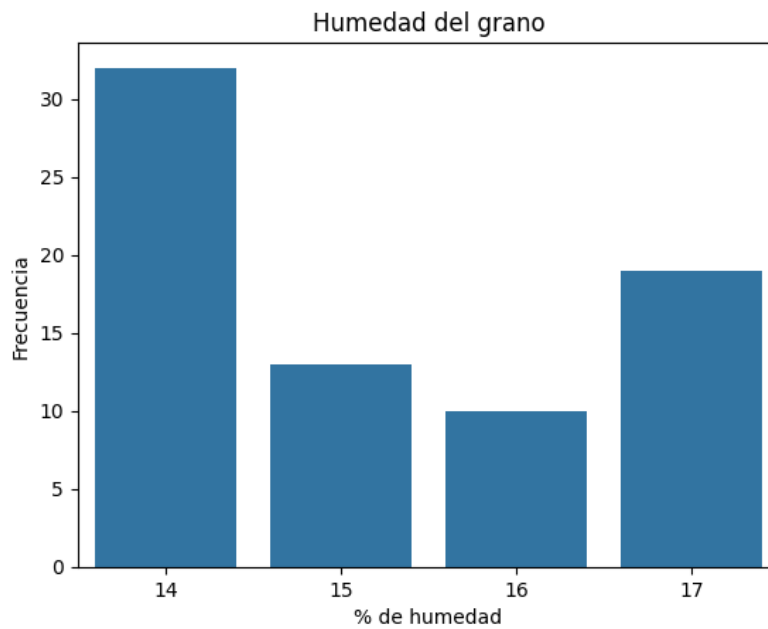


Ilustración 5 Frecuencia de la humedad en la carga.

Acá en esta ilustración 5 podemos ver la frecuencia con la que se presenta cada porcentaje de humedad en cada cargue de grano, pudiendo evidenciar que 14% es el más común. Lo que nos indica que el arroz que se está recibiendo viene con una humedad optima, y es beneficioso para el agricultor en cuanto al precio.

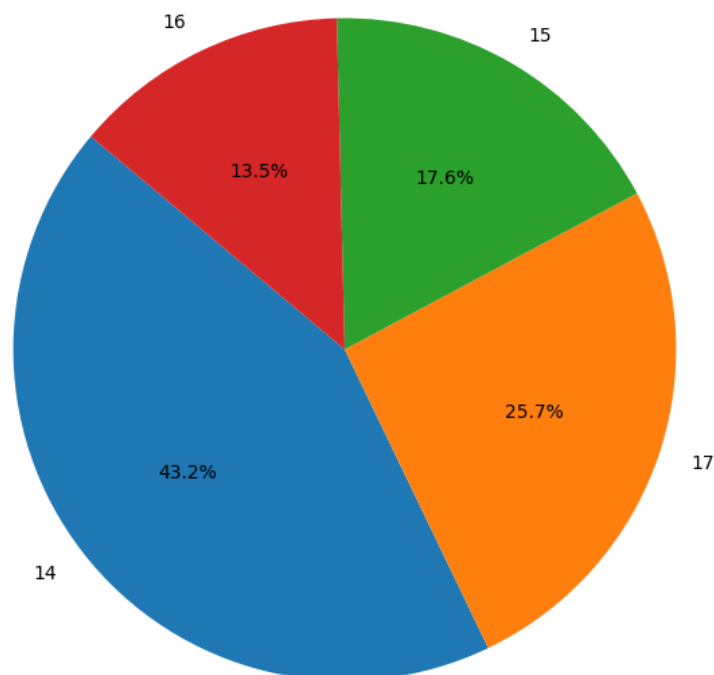


Ilustración 6 Diagrama de análisis de la humedad y el porcentaje de carga.

En esta ilustración 6 apreciamos que el arroz con mayor porcentaje dentro del total recibido es de 43,2% con una humedad óptima de 14 %, y el de menor porcentaje es de 13,5% con una humedad del 16%.

Objetivos:

Objetivo general.

Implementar un algoritmo computacional que analice y ayude a tomar decisiones, a partir de datos de recibo de arroz día a día en una planta comercializadora de este

grano, que nos ayude a poder tomar decisiones de cómo, cuándo y dónde almacenarlo, utilizando estrategias de machine learning.

Objetivos específicos.

- Caracterizar y procesar datos de interés, con miras a la toma de decisiones informadas.
- Implementar un algoritmo de Machine Learning para la toma de decisiones a partir de los datos de interés.
- Evaluar y analizar el desempeño de los algoritmos utilizados para la toma de decisiones.
- Validar el funcionamiento de toma de decisiones a partir de nuevos datos.

Desarrollo e implementación del aprendizaje.

Partiendo de la tabla o planilla suministrada por el operario o encargado del área de recibo en la planta de la empresa **GRANOS Y CEREALES DE COLOMBIA**. Donde podemos apreciar los resultados diarios de la carga de grano de arroz que se descarga a diario, además del porcentaje de humedad con el que llega este grano, el cual es muy importante ya que de este depende un buen proceso de secado, trillado y comercialización. Se requiere predecir a futuro de cinco días posteriores a la fecha. De cuanto grano podría llegar, teniendo en cuenta los datos de carga descargada en los cinco días anteriores a la fecha. Y así poder tomar la decisión de donde, cuando y como

almacenar este en la planta, optimizando los tiempos y equipos a implementar en este proceso.

Del siguiente proceso de organización de los datos para poder proyectarlos en los algoritmos de regresión, obtenemos las siguientes variables de entrada:

- Carga del último día.
- Carga del sexto día.
- Carga del séptimo día.
- Carga del octavo día.
- Carga del noveno día.

El dato de salida será el último tomado en cada día, y así obtener un consolidado, que se pueda procesar en los algoritmos de regresión.

Conocer cuál va a ser el resultado de una campaña de producción de un producto agrícola, ha sido siempre un resultado muy apreciado por el sector agroindustrial. Lo cual, con el transcurso de los años, han aparecido sistemas de estimación de la cosecha en cuanto a diversos productos.

La evolución tecnológica y la disponibilidad de datos abiertos al público, han dado como fruto la implementación o desarrollo de modelos matemáticos que pueden predecir la producción del producto con meses de anterioridad.

En diferentes países desarrollados como Estados Unidos han establecido como política nacional, la disponibilidad de datos, los cuales puedan ser aprovechados por las diferentes empresas o personas para desarrollar programas que beneficien la cadena agroalimentaria, obteniendo datos como factores climáticos, satelitales, productividades históricas, entre otros datos importantes que ayudan a desarrollar programas que puedan estimar una producción eficaz o en dado caso que no sea factible, poder tomar decisiones que no afecten.

Estos datos no serían importantes sin saber cómo procesarlos o no tener herramientas para dicho proceso. Por lo cual existen métodos matemáticos asistidos con la inteligencia artificial, que son aprovechados tecnológicamente en diversos campos como en el mencionado que sería de predicción (Marqués Gozalbo, M. Á. (2022). Modelos predictivos de producción agroindustrial con Machine Learning a partir de fuentes de información pública.).

Procesamiento de datos.

CARGA DEL ULTIMO DÍA	CARGA DEL SEXTO DÍA	CAEGA DEL SEPTIMO DÍA	CARGA DEL OCTAVO DÍA	CARGA DEL NOVENO DÍA	CARGUE DEL SIGUIENTE DÍA DÍA
450	450	450	450	450	320
450	450	450	450	450	320
540	450	450	450	450	320
540	540	450	450	450	320
540	540	540	450	450	320
280	540	540	540	450	320
280	280	540	540	540	450
540	450	450	450	450	450
280	280	280	280	540	450
280	280	280	280	280	450
380	280	280	280	280	450
380	380	280	280	280	450
380	380	380	280	280	540
380	380	380	380	280	540
380	380	380	380	380	540
380	380	380	380	380	280

Ilustración 7 Organización de tabla para el análisis de regresión.

En esta ilustración 7 podemos ver la organización del consolidado teniendo en cuenta solo la carga que llego cada día, y así reagruparla tomando cada dato y dejando cinco espacios hacia atrás, tomando los siguientes cinco datos.

	CARGA DEL ULTIMO DÍA	CARGA DEL SEXTO DÍA	CAEGA DEL SEPTIMO DÍA	CARGA DEL OCTAVO DÍA	CARGA DEL NOVENO DÍA	CARGUE DEL SIGUIENTE DÍA DÍA
0	450	450	450	450	450	320
1	450	450	450	450	450	320
2	540	450	450	450	450	320
3	540	540	450	450	450	320
4	540	540	540	450	450	320
5	280	540	540	540	450	320
6	280	280	540	540	540	450

Ilustración 8 Descarga de datos en el algoritmo de regresión.

Podemos apreciar en esta ilustración 8, como procesa los datos el algoritmo, reconociéndolos como un archivo de Excel.

```
#Ajustando los datos a una matriz numérica
import numpy as np
Dataset_Matriz = np.array(Dataset)
#Dividiendo los datos en entrada y salida
Entradas = Dataset_Matriz[:,0:4]
Salida = Dataset_Matriz[:, -1]
```

Ilustración 9 Entrenamiento para predecir

Acá podemos ver como el algoritmo entrena a partir de 4 entradas a partir de cero y una salida, reagrupa los datos en una matriz, para así poder predecir la llegada del grano en cinco días al futuro.

```

▶ #Se ingresan los datos de entrada para generar una predicción
Nueva_entrada = np.zeros((1,5))
Nueva_entrada[0,0]=float(input('Ingrese el valor de la carga del ultimo dia: '))
Nueva_entrada[0,1]=float(input('Ingrese el valor de la carga del sexto dia: '))
Nueva_entrada[0,2]=float(input('Ingrese el valor de la carga del septimo dia: '))
Nueva_entrada[0,3]=float(input('Ingrese el valor de la carga del octavo dia: '))
Nueva_entrada[0,4]=float(input('Ingrese el valor de la carga del noveno dia: '))

Proyeccion_2 = Modelo_2.predict(Nueva_entrada)

print('')
print('')
print('')
print('')
print('Según los datos ingresados, la proyección de la carga para el siguiente día usando ANN será: ',Proyeccion_2[0])

```

↳ Ingrese el valor de la carga del ultimo dia: 350
Ingrese el valor de la carga del sexto dia: 350
Ingrese el valor de la carga del septimo dia: 350
Ingrese el valor de la carga del octavo dia: 450
Ingrese el valor de la carga del noveno dia: 450

Según los datos ingresados, la proyección de la carga para el siguiente día usando ANN será: 400.7000237229171

Ilustración 10 Predicción de la carga próxima a llegar.

En este análisis predictivo con el algoritmo de regresión, que utiliza un modelo de la red neuronal, con datos a partir del último día hasta el noveno días anterior, evidenciar que en los cinco días posteriores se tendrá una carga próxima a llegar de 400 toneladas de arroz.

Modelo de toma de decisiones.

La predicción del algoritmo es acertada ya que la tendencia de la llegada de grano tiene un rango entre los 350 y 450 toneladas por día, a lo cual se puede prever que en los 5 días posteriores se tendrá que tener un espacio para poder recibir arroz sin ningún tipo de embotellamiento puesto que podemos trabajar dos turnos, uno en el día y otro en la noche ya que solo se cuenta con una máquina que pre limpia el arroz, con una capacidad de 50 toneladas hora, y si se mantiene la tendencia, que es lo que se asume, tendríamos que poder recibir y almacenar en los silos metálicos de recibo, y a la vez ir pasando a la pre limpieza y no tener represamiento, para poder tener una eficiencia en el descargue y salida del grano desde las tolvas de recibo.

Implementación en contextos reales.

En base al problema planteado, el cual se remite a poder saber cómo fluir el grano dentro de los equipos de la planta, como procesarlo sin llegar a tener un embotellamiento y no poder recibir el arroz, teniendo un proceso u operación eficiente dependiendo de los tiempos y la capacidad de cada equipo ya sea de procesamiento o almacenamiento. El proceso primordial en el tiempo de cosecha es de recibir el grano, secarlo y almacenarlo, por que dependiendo de esta cantidad de grano se podrá cumplir con la demanda a la cual se sujeta la producción.

Otros casos de implementación de estos algoritmos de regresión, son en aquellos que se necesite pronosticar las ventas, como por ejemplo en los supermercados con los

productos que tengan gran acogida puede ser en este caso el arroz que se compre, saber si los usuarios o compradores tienen preferencia por este arroz que tiene una marca establecida.

Resultados adicionales.

Pudimos evidenciar que la humedad relativa con la cual llega el grano a la planta es casi la óptima, lo cual agiliza la operación de recibo un poco más, por que solo hay que tenerlo en reposo un tiempo corto y llevarlo al almacenamiento, no sin antes ventilarlo en el estado de reposo.

Resumen.

Hablando de temas de métodos y tiempos dentro de una planta de procesamiento de arroz, es de suma importancia tener bajo control todos los aspectos que puedan generar un impacto dentro de este. Por lo cual en este caso planteado del proceso de recibo de este grano se requiere poder predecir que cantidad de arroz pueda llegar en cinco días a futuro, para así poder contar con el espacio y tiempo, sin llegar a tener ningún imprevisto, puesto que la cosecha del arroz es muy incierta conforme pasa el tiempo dentro de los tres meses que se presenta esta.

Dentro de este mismo planteamiento también podemos llegar a establecer que con el tipo de planteamientos de logística debemos poder mejorar con el paso del tiempo en

cada cosecha, con respecto a el tema de infraestructura dentro del proceso, además de poder innovar.

Conclusiones.

- Con este algoritmo de regresión podemos aproximar una predicción del comportamiento del grano con respecto a datos anteriores, lo cual nos llevara a tener una información confiable dentro de un rango aceptable.
- Este tipo de planteamiento me permite ver como se comporta la cosecha del arroz.
- Podemos evidenciar más a fondo que la cantidad de arroz que llega a una planta también puede depender de la infraestructura con la que cuente la planta.

Referencias.

- Fernández, T. D. (2020). Taxonomía de transformación digital. *Revista Cubana de transformación digital*, 1(1), 4-23.
- Chica, J., Tirado, Y. C., & Barreto, J. M. (2016). Indicadores de competitividad del cultivo del arroz en Colombia y Estados Unidos. *Revista de Ciencias Agrícolas*, 33(2), 16-31.
- APAZA PACO, J. P., CHÁVEZ LIZÁRRAGA, G. A., & HERRERA CHOQUE, V. A. N. I. A. (2015). Logística de almacenamiento de materia prima en la industria farmacéutica. *Revista Con-ciencia*, 3(1), 101-113.
- Moreno, E. G. (2001). *Automatización de procesos industriales*. Valencia: Alfaomega.
- Marqués Gozalbo, M. Á. (2022). Modelos predictivos de producción agroindustrial con Machine Learning a partir de fuentes de información pública.