C&T Review Journal

Riqchary

Revista de Investigación en Ciencia y Tecnología

Volumen 05 - Número 1 -Octubre ISSN: 2810-8124 (en línea)



begin ( )

DOI: https://doi.org/10.57166/riqchary/v5.n1.2023

UNIVERSIDAD NACIONAL MICAELA BASTIDAS DE APURÍMAC

VRIN

ISSN/2810-8124



Comité Editorial

- **Editor General:** Dr. Ecler Mamani Vilca
  Universidad Micaela Bastidas Nacional de Apurímac, Perú

Equipo Técnico:

- Dr. José Luis Merma Aroni
  Universidad Micaela Bastidas Nacional de Apurímac, Perú
- M.Sc. Mario Aquino Cruz
  Universidad Tecnológica de los Andes Perú
- Dr. Manuel J. Ibarra Cabrera
  Micaela Bastidas Universidad Nacional de Apurímac, Perú
- Dr. Yalmar Ponce Atencio
  Micaela Nacional de Juliaca, Perú
- Dr. Leonel Coyla Idme
  Universidad Nacional del Altiplano, Perú
- Dr. José Luis Herrera Salazar
  Universidad Norbert Wiener, Perú
- Mag. Olga Jaramillo
  Santa Fe Community College, Estados Unidos
- Dr. Edwar Ilasaca Cahuata
  Universidad Nacional del Altiplano, Perú
- Mag. Yonatan Mamani Coaquira
  Universidad Micaela Bastidas Nacional de Apurímac, Perú

Equipo Editorial:

- Mag. Marleny Peralta Ascue
  Universidad Privada de Tacna, Perú
- Mag. Rafael Quispe Merma
  Universidad Nacional San Antonio de Cusco, Perú
- Ing. Ebert Gómez Aiquipa
  Universidad Nacional San Antonio de Cusco, Perú
- MBA. Evelyn Naida Luque Ochoa
  Micaela Bastidas Universidad Nacional de Apurímac, Perú
- Mag. Kevin Arnold Arías Figueroa
  Micaela Bastidas Universidad Nacional de Apurímac, Perú
- Mag. Alejandrina Huaylla Quispe
  Universidad Tecnológica de los Andes Perú
- Mag. Betsabe M. Ccolqque Ruiz
  Micaela Bastidas Universidad Nacional de Apurímac, Perú
- Mag. Miguel Alfaro
  Micaela Bastidas Universidad Nacional de Apurímac, Perú

Periodicidad: Semestral
Vol. 5 Núm. 1 año (2023)
Áreas temáticas: Ciencias exactas, ingenierías y de materiales
Editorial: UNAMBA – Perú
Vice Rectorado de Investigación

https://doi.org/10.57166/riqchary/v5.n1.2023
ISSN: 2706-543X- **ISSN:** 2810-8124 (e)

Editorial

Con gran satisfacción presentamos la revista de investigación C&T Riqchary, un espacio dedicado a la publicación de artículos en las áreas de ciencias exactas, ingenierías y de materiales. Esta revista, de carácter académico, tiene como finalidad promover la difusión y el intercambio de conocimientos e investigaciones en dichos campos. A partir del año 2024, la revista adopta una periodicidad semestral y un enfoque de acceso abierto, fortaleciendo su compromiso con la comunidad académica y el público en general. De este modo, se consolida como un medio que impulsa el desarrollo científico y tecnológico, contribuyendo al avance del conocimiento.

Director: Dr. Ecler Mamani Vilca





C & T Riqchary

Revista de investigación - UNAMBA

Autores

- Enciso Ortiz Sergio Elías
- Mamani Vilca Ecler
- Ordoñez Ramos Erech
- Isael Gomez Villa
- Sleyther Arturo De La Cruz Vega
- Francisco Cari Incahuanaco
- Yhon Fuentes Huamán
- Maritza Condori Quispe
- Carlos Manuel Ruano González
- Dennis Uriel Añasco Chata
- Percy Huata Panca
- Adolfo Carlos Jimenez Chura
- Pablo Cesar Tapia Catacora
- David F. Catalán Mendoza
- José A. Cárdenas Catalán
- Wilson J. Mollocondo Flores
- John A. Aguirre Carrasco
- Luis M. Alfaro Chirinos
- Alejandrina Huaylla Quispe
- Edgar Eloy Carpio Vargas
- Alicia Roxana Mayda Huanca
- D. German Rafael Espinoza Rivas
- Betsabe M. Ccolqque Ruiz
- Filio Carrasco Sauñe
- Manuel J. Ibarra Cabrera
- Norma Aguirre Quispe
- Niki Franklin Flores Pacheco
- Salvador Quispe Chipana
- Maria Luisa Flores Pacheco

- © Dr. Ecler Mamani Vilca / Director del Comite Editorial

Av. Garcilaso C.U Tamburco, Abancay Telef, +51983747138
Apurímac – Perú eclervirtual@unamba.edu.pe
Hecho el Depósito Legal en la Biblioteca Nacional del Perú N° 2020-10220 ISSN: 2709-8990
ISSN: 2955-8646 (e)- ISSN: 2709-8990
Tiraje: 500 ejemplares

Volumen 5 Número 1 – 2023
Se terminó de imprimir en El Misti EIRL
RUC 20450599108 - Cel: 950-509953, imprentamisti@hotmail.com

INDICE

01 *Pág.* Determinación de la mejor Arquitectura de Redes Neuronales Convolucionales: VGG16, ResNet50 o MobileNet para detección de la Neumonía 2023.

11 *Pág.* Análisis del afirmado con la adición de polímero acrilamida y cemento, Ocobamba, Chincheros, Apurímac 2023.

19 *Pág.* Tutoría virtual en la formación académica y humana de estudiantes de la UNAMBA.

23 *Pág.* Metodología para el diagnóstico y la gestión de mantenimiento en las subestaciones eléctrica de los grupos de generación distribuida.

29 *Pág.* Método de Box-Jenkins para el pronóstico de consumo de energía eléctrica del 2021-2023.

37 *Pág.* Análisis del comportamiento geomecánica del macizo rocoso en galería nivel 4040, mina Victoria.

43 *Pág.* Análisis del rendimiento en Clústeres basados en PelicanHPC.

49 *Pág.* La mejor opción para predecir el precio máximo de las acciones de Intel Corporation: ¿Árbol de regresión o Regresión Lineal Múltiple?

57 *Pág.* Determinación del mejor algoritmo de reconocimiento de imágenes de camélidos sudamericanos mediante machine learning, Apurímac, 2023.

63 *Pág.* Caracterización de los suelos de la comunidad de Huayllati, Apurímac

C&T Riqchary
Revista de Investigación en Ciencia y Tecnología
ISSN: 2810-8124 (en línea) / ISSN: 2706-543x

Vol. 5 Núm. 1 (2023)
Publicado: 15/10/23
DOI: https://doi.org/10.57166/riqchary/v5.n1.2023.104

Determinación de la mejor Arquitectura de Redes Neuronales Convolucionales: VGG16, ResNet50 ó MobileNet para detección de la Neumonía 2023

The Determination of the Best Convolutional Neural Network Architecture: VGG16, ResNet50, or MobileNet for Pneumonia Detection in 2023

Enciso Ortiz Sergio Elías [A], Mamani Vilca Ecler [B] y Ordoñez Ramos Erech [C]

ORCID:0009-0005-0863-4889[A], ORCID:0000-0002-3058-9683[B] y ORCID: 0000-0002-5205-3660[C]

(Recepción: 8/08/2023 y aceptación 5/09/2023)

Resumen— Las Redes Neuronales Convolucionales (CNN) se utilizan para el reconocimiento de imágenes de radiografías y otras aplicaciones. Actualmente, existen trabajos que comparan la eficacia de las arquitecturas CNN, como VGG16, ResNet50 y MobileNet con parámetros de entradas diferentes en sus entrenamientos, creando incertidumbre entre los desarrolladores de aplicaciones de clasificación de imágenes, nosotros sí aplicamos entradas iguales para el entrenamiento de las CNN en estudio. Para abordar esta falta de información, se utilizó una base de datos de Kaggle que consta de 5856 imágenes. De esa base de datos, se seleccionó una muestra sistemática de 746 imágenes de radiografías de pulmones sanos y con Neumonía. Para asegurar la normalización de las imágenes, se utilizaron las herramientas iloveimg y ReNamer. Además, se utilizó Python con Google Colab y diversas librerías, como tensorflow, matplotlib, numpy, os, cv2 y random, para ejecutar las diferentes arquitecturas. El diseño metodológico se basó en un enfoque cuantitativo, empleando tablas de comparación y las imágenes adquiridas de la base de datos de Kaggle. Los resultados obtenidos indicaron que el porcentaje de exactitud (Accuracy) fue de 80.83% para VGG16, 91.82% para ResNet50 y 78.28% para MobileNet, concluyéndose que ResNet50 es la arquitectura más precisa en este contexto.

Palabras claves: Arquitecturas de Redes Neuronales Convolucionales, MobileNet, ResNet50, VGG16.

Abstract— Convolutional Neural Networks (CNN) are used for the recognition of X-ray images and other applications. Currently, there are studies comparing the effectiveness of CNN architectures like VGG16, ResNet50, and MobileNet with different input parameters during their training, causing uncertainty among image classification application developers. In our study, we used consistent input parameters to address this lack of information. We employed a Kaggle database containing 5856 images. From this database, we systematically selected 746 images of X-rays depicting healthy and pneumonia-afflicted lungs. To ensure image normalization, we utilized tools such as iloveimg and ReNamer. Furthermore, we employed Python in conjunction with Google Colab and various libraries including TensorFlow, Matplotlib, NumPy, os, cv2, and random to execute the various CNN architectures. Our methodology was based on a quantitative approach, utilizing comparison tables and the images sourced from the Kaggle database. The results we obtained indicated an accuracy rate of 80.83% for VGG16, 91.82% for ResNet50, and 78.28% for MobileNet. We concluded that in this context, ResNet50 is the most accurate architecture.

Keywords: Convolutional Neural Network Architectures, MobileNet, ResNet50, VGG16.

A. *Enciso Ortiz Sergio Elías[A] Escuela Académico Profesional de Ingeniería Informática y Sistemas de la Universidad Nacional Micaela Bastidas de Apurímac-Perú, sergioeliasencisoortiz@gmail.com*

B. *Mamani Vilca Ecler[B] Escuela Académico Profesional de Ingeniería Informática y Sistemas de la Universidad Nacional Micaela Bastidas de Apurímac- Perú, eclervirtual@unamba.edu.pe*

C. *Ramos Ordoñez Erech[C] Escuela Académico Profesional de Ingeniería Informática y Sistemas de la Universidad Nacional Micaela Bastidas de Apurímac - Perú. eordonhez@unamba.edu.pe*

1 INTRODUCCIÓN

En el aprendizaje automático de la inteligencia artificial en adelante (IA) se basa en varios modelos computacionales y una de ellas es la Redes Neuronales Convolucionales (CNN, por sus siglas en inglés en adelante) se utilizan principalmente para tareas de procesamiento de imágenes y visión computacional, sin embargo dentro de ellas también existe diferentes arquitecturas como: Visual Geometry Group 16 (en adelante VGG16), Red Residual 50 (en adelante ResNet50) y Red Móvil (en adelante MobileNet), el problema es determinar **¿cuál de ellos tiene mejor accuracy?,** el objetivo es analizar y ejecutar las CNN VGG16, ResNet50 y MobileNet para determinar el que tiene mayor porcentaje de accuracy en el reconocimiento de Neumonía con imágenes RX mediante la métrica de evaluación (Matriz de confusión) y la hipótesis, es encontrar que exista una diferencia de accuracy entre las CNN VGG16, ResNet50 ó MobileNet, en la detección de la Neumonía.

En los trabajos de Maldonado Lezama, Lisset Fernanda y Moreano Rojas, Eder Omar **[1]**, realizó el diagnóstico de enfermedades pulmonares cuyo resultado de accuracy fue la implementación propia de 96.19%, InceptionV3 de 86.46% y VGG16 de 96.32% siendo este el mejor, según Godoy Francisco, Vady Lee **[2]**, el trabajo fue determinar el efecto del algoritmo de transferencia de aprendizaje con accuracy para resnet50 96.61%, VGG16 86.26% y RNN 95.89% donde ResNet50 resultó ser el mejor, también Caya Pérez, Jhan Carlos **[3]** realizó el requerimiento para el diagnóstico del análisis de RX de tórax, para evaluar el estado de los pulmones donde utilizó tres arquitecturas de CNN como: implementación propia, ResNet50 e InceptionV3, con accuracy de 96.59%, 97.63% y 98.48% respectivamente siendo InceptionV3 el mejor, así mismo según Cabrejos Yalán, Víctor Manuel **[4]** utilizaron CNN como: AlexNet, VGG y ResNet-50 para mejorar el diagnóstico de la Neumonía utilizando imágenes de RX, a fin de mejorar el tiempo de atención de pacientes con esta enfermedad y ayudar ante la demanda, obtuvieron resultados con accuracy del 95.22% para AlexNet que obtuvo el mejor resultado, 94.53% para VGG16 y 94.62% para ResNet50. Moura y otros **[5]** presenta la clasificación de imágenes RX de tórax de niños en dispositivos móviles utilizando CNN como MobileNetV1, MobileNetV2 e InceptionV3 con accuracy de 95.4%, 94.2% y 92.8% respectivamente siendo MobileNetV1 el mejor.

Se entiende por IA, Pino, Gómez y Martínez lo define como: "un campo de la ciencia y la ingeniería que se ocupa de la comprensión, vista de un punto informático, que se denomina comportamiento inteligente" **[6]**, en la medicina la IA es una rama de la informática capaz de analizar datos médicos complejos **[7]**, también Recuenco y Reyes define como: La base sobre la que los sistemas se centran en el comportamiento para definirlos como inteligentes, así como la primera interpretación de tipo científico donde un sistema es inteligente cuando pasa con éxito el test de Turing **[8]** además Boden define como: IA no es una dimensión única, sino un espacio profusamente estructurado de capacidades diversas para procesar la información, de modo que esta utiliza muchas técnicas diferentes para resolver gran variedad de tareas **[9]**. El aprendizaje automático (Machine Learning) es una rama de IA. Gerón **[10]** lo define como: "El aprendizaje automático es la ciencia y el arte de programar máquinas para que puedan aprender de los datos", En consecuencia, el aprendizaje automático puede ayudar a resolver problemas complejos, aprender de manera autónoma y descubrir patrones ocultos en grandes conjuntos de datos, lo que proporciona una valiosa herramienta tanto para la automatización como para el análisis de información. Este consta de tres partes: el algoritmo computacional, las variables y las funciones que conforman la decisión y el conocimiento base según el cual se sabe la respuesta que permite aprender al sistema (lo entrena). Así también se clasifican en: Aprendizaje supervisado, no supervisado y por refuerzo, el primero es automático se le proporciona una serie de datos de entrada cuya variable de salida es conocida, que tiene la finalidad de construir la regla general. El aprendizaje no supervisado no se utiliza la clase de salida esperada para los datos, el algoritmo no sabe nada. Por lo tanto, se tiene que estudiar el comportamiento de un conjunto de patrones de datos de entradas para determinar comportamientos parecidos que se agrupan y por último el consiste que algoritmo aprende por su cuenta propia **[10]**, además el aprendizaje automático es combinar varios componentes del algoritmo, como una función de costo, una función de activación y una técnica de regularización, para formar un algoritmo de aprendizaje profundo completo **[11]**.

Las Arquitecturas de Redes Neuronales según Martin del Brio y Serrano **[12]** lo denomina arquitectura a la topología, estructura o patrón de conexionado de una red neuronal. En general y de modo análogo a lo que sucede en el cerebro, las neuronas se suelen agrupar en unidades estructurales que denominaremos capas y el conjunto de una o más capas conforma la red neuronal y al inicio las arquitecturas tendrán muchas predicciones incorrectas **[8]**. Las CNN son un tipo especial de Redes Neuronales Artificiales, que esta réplica del funcionamiento del

C&T Riqchary
Revista de Investigación en Ciencia y Tecnología
ISSN: 2810-8124 (en línea) / ISSN: 2706-543x

Vol. 5 Núm. 1 (2023)
Publicado: 15/10/23
DOI: https://doi.org/10.57166/riqchary/v5.n1.2023.104

cerebro humano es uno de los "modelos de moda" **[13]** y en este trabajo se definirán las antes mencionadas: VGG16 es un modelo de aprendizaje profundo presentado por Karen Simonyan y Andrew Zisserman en el artículo "Redes convolucionales muy profundas para el reconocimiento de imágenes a gran escala" **[14]**, ResNet50 es según **[15]** ResNet50 es un modelo de Red Neuronal Convolucional propuesto en el 2016 por He et al. El cual, no posee una profundidad de capas fija y es dependiente del número de módulos que se utilicen, además permite un tamaño de entrada máximo de 224 x 224 y MobileNet es otro tipo de CNN que "reduce significativamente el número de parámetros en comparación con la red con convoluciones regulares con la misma profundidad en las redes" **[16]**, más adelante cada arquitectura en estudio se definirán con más detalle. y la métrica a utilizar es la matriz de confusión para calcular así el accuracy de cada arquitectura analizada.

Con respecto al diagnóstico, es el arte de aproximarse al conocimiento de un ser humano, de su situación de salud física y emocional, mediante la comunicación verbal y no verbal y la exploración física de su organismo, en el marco de una relación respetuosa, cordial y la alerta, que tiene como fin la solución o alivio de los padecimientos o los riesgos descuidos descubiertos" **[17]** y se tiene las características del término diagnóstico como: clima adecuado en relación médico paciente, médico alerta con la vista aguda, conocimiento adecuado en el campo laboral y desarrollo de aptitudes, además, que se puede medir con la inferencia deductiva, descripciones teóricas, inferencia inductiva, inferencia análoga y inferencia estadística. Según la OMS **[18]** la neumonía es la principal causa de muerte infantil en todo el mundo. Se estima que esta mató a unos 920.136 niños menores de cinco años en 2015, lo que representa el 15% de todas las muertes entre niños en todo el mundo. La neumonía afecta a los niños y sus familias" y las características son: Streptococcus pneumoniae la causa más común, haemophilus influenzae de tipo b (Hib) es la segunda causa de neumonía bacteriana, el virus sincitial respiratorio es la causa común de neumonía viral, pneumocystis jiroveci es la principal causa de esta en niños menores de seis meses con VIH/SIDA causando una cuarta parte de las muertes en bebés VIH positivos, diagnóstico de streptococcus pneumoniae es la causa más común de neumonía bacteriana en niños, en consecuencia el diagnóstico de la neumonía es un componente fundamental de toda estrategia para reducir la mortalidad infantil y para el diagnóstico de neumonía suele basarse en la clínica y en la radiología **[18]** y se debe intentar una prueba microbiológica para orientar el tratamiento y el diagnóstico nosológico" **[19]**, luego el diagnóstico expone lo siguiente: Una radiografía de tórax analiza si hay inflamación en los pulmones. Se suele usar RX de tórax para detectar neumonía" **[20]**, Los análisis de sangre, como un hemograma completo (CBC), se usan para saber si el sistema inmunitario está luchando contra una infección **[20]**, La oximetría de pulso mide la cantidad de oxígeno en la sangre, este puede impedir que los pulmones dejen pasar suficiente oxígeno a la sangre **[20]** y un examen de RX para que el médico interprete dicho RX, el radiólogo buscará puntos blancos en los pulmones (llamados infiltrados) **[20]** y también las mediciones del diagnóstico es la presencia de taquipnea constituye uno de los indicadores más confiables de gravedad de la neumonía en los análisis univariados y multivariados, esto es sin embargo, en la práctica clínica, se recomienda que la frecuencia respiratoria mayor de 20 respiraciones/min sea considerada un signo de gravedad confiable en los pacientes con neumonía **[21]**.

2 LAS ARQUITECTURAS DE REDES NEURONALES

2.1 VGG16

Según Pintos y Simonyan [22] tiene como entrada una imagen RGB de tamaño fijo de 224 × 224 píxeles [23]. "Redes convolucionales muy profundas para el reconocimiento de imágenes a gran escala" [14] la imagen se pasa a través de una serie de capas convolucionales (conv.) donde se utilizan filtros con un campo receptivo muy pequeño: 3 × 3 (que es el tamaño más pequeño para capturar la noción de izquierda/derecha, arriba/abajo, centro). VGG16 proviene del hecho de que el modelo tiene 16 capas en total. La Arquitectura se caracteriza por tener una estructura simple y uniforme, utilizando solo capas convolucionales de tamaño 3x3 con stride 1 y capas de pooling de tamaño 2x2 con stride 2. No se utilizan técnicas más avanzadas, como convoluciones dilatadas o convoluciones separables en profundidad. VGG16 es una Red Neuronal Convolucional profunda que utiliza convoluciones 3x3 y capas de pooling 2x2 para extraer características de una imagen. Consta de 13 capas convolucionales, 5 capas de pooling y 3 capas totalmente conectadas, seguidas de una capa de salida. Aunque esta arquitectura es más profunda y tiene más parámetros que algunas de las arquitecturas más modernas, su simplicidad y uniformidad han demostrado un buen rendimiento en muchas tareas de visión computacional.

Revista de investigación en ciencia y tecnología
Vol. 5 Núm. 1 (2023) - publicado:15/10/2023
DOI https://doi.org/10.57166/riqchary/v5.n1.2023.104

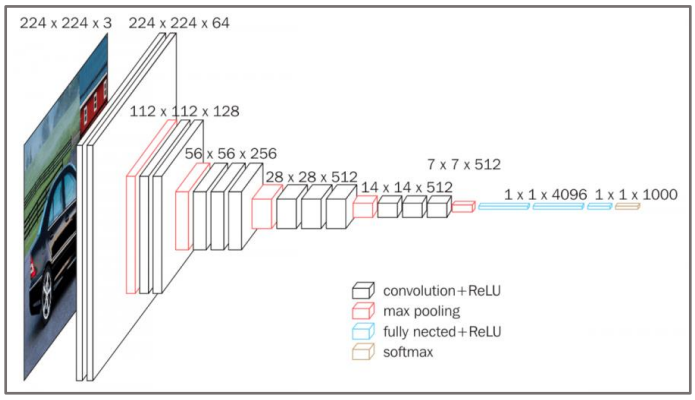


Fig. 1. Arquitectura de Red Neuronal Convolucional VGG16 [14]

2.2 RESNET50

ResNet50 es una Red Neuronal Residual con 50 capas, su estructura es simple y efectiva, es beneficiosa por su amplia aplicación [24] como se observa en la figura 02. La red contiene un bloque de estructura residual, donde solo unos pocos dentro del mapeo de identidad, necesitan hacer una coincidencia de cotas y usar una capa convolucional de 1x1 para agregar cotas. Con ResNet50 [25], conseguimos resolver el problema de desvanecimiento del gradiente, debido a que la red es demasiado profunda, además con esta red los gradientes pueden pasar directamente a través de las conexiones de salto hacia atrás desde las capas posteriores hasta los filtros iniciales, demostrando un rendimiento sobresaliente en tareas de clasificación de imágenes y detección de objetos. ResNet-50 ha influido en el diseño de numerosas arquitecturas posteriores y consta de más 23 millones de parámetros que claramente es menor a VGG16 y consiste en una serie de bloques residuales que aprenden características a diferentes niveles de abstracción. Al final, se utilizan capas de pooling, capas completamente conectadas y una capa de salida para realizar tareas específicas como clasificación de imágenes.

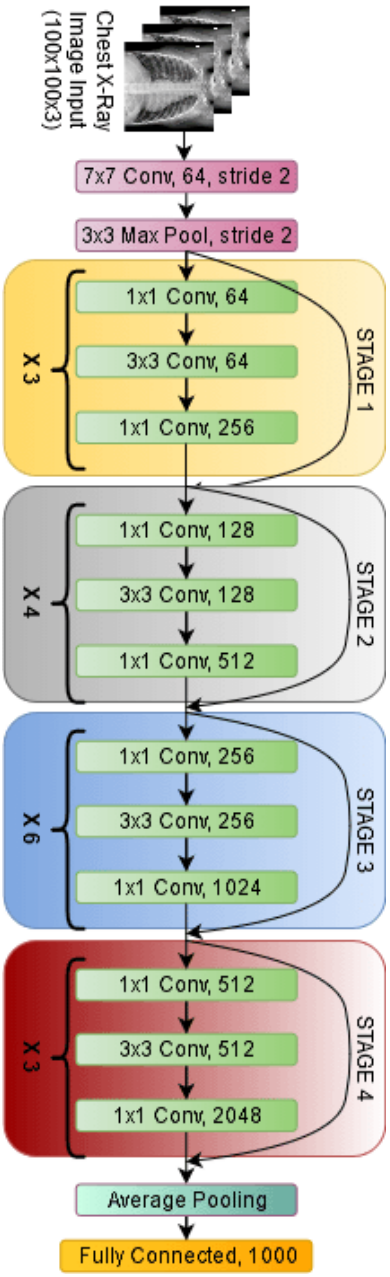


Fig. 2. Arquitectura de Red Neuronal Convolucional ResNet50 [25]

2.3 MOBILENET

Para [26] MobileNet se basa en convoluciones separables en profundidad, que es una forma de convoluciones factorizadas que descomponen una convolución estándar en una convolución en profundidad y una convolución 1 × 1 llamada convolución de puntos. MobileNet "Reduce significativamente el número de parámetros en comparación con la red con convoluciones regulares con la misma profundidad en las redes" [27]. La convolución en profundidad aplica un único filtro a cada canal de entrada. Luego, la convolución de puntos aplica una convolución 1 × 1 para combinar las salidas de la convolución en profundidad y se basa en la convolución separable, que separa los filtros espaciales y de profundidad en dos etapas, así ayuda a reducir la carga computacional y el número de parámetros de la red, lo que la hace más eficiente en términos de recursos y tiempo de ejecución.

C&T Riqchary
Revista de Investigación en Ciencia y Tecnología
ISSN: 2810-8124 (en línea) / ISSN: 2706-543x

Vol. 5 Núm. 1 (2023)
Publicado: 15/10/23
DOI: https://doi.org/10.57166/riqchary/v5.n1.2023.104

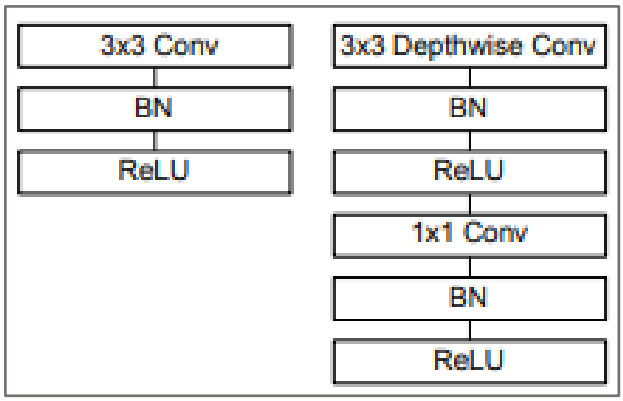


Fig. 3. Arquitectura de Red Neuronal Convolucional MobileNet [28]

3 MÉTRICAS

3.1 Matriz

Para [29] la matriz de confusión es una de las métricas más intuitivas y sencillas para estimar la exactitud y otros, que es una tabla que describe el rendimiento de la arquitectura [30] y a partir de ella podemos calcular la métrica del accuracy [31].



Fig. 4. Matriz de confusión

3.2 Accuracy

Accuracy o exactitud es una de las métricas más utilizadas y confiables para conocer el rendimiento de las CNN [30], para Mora [29] es el cálculo en el porcentaje de muestras que se clasifican correctamente y se muestra la ecuación respectiva:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

4 IMÁGENES

4.1 Imágenes de RX

Según [20] "Una radiografía de tórax analiza si hay inflamación en los pulmones. Se suele usar una radiografía de tórax para detectar neumonía" para interpretar los RX, el radiólogo buscará puntos blancos en los pulmones (llamados infiltrados) que son característicos de una infección, así como se muestra en la figura 5.



Fig. 5. Radiografía de un pulmón con neumonía [32]

5 METODOLOGÍA

La investigación es aplicada con nivel experimental y diseño factorial por bloques, para Hernández Sampieri y otros [33], los diseños factoriales manipulan dos o más variables independientes e incluyen dos o más niveles o modalidades de presencia en cada una de las variables independientes.

5.1 Población y Muestra

La población en este trabajo de investigación ha sido adquirida de la base de datos de Kaggle [32], entre datos de entrenamiento, validación y prueba todos suman la cantidad de la población total es 5856 y se trabajó con la carpeta de prueba (test) correspondiente a 746 datos que son imágenes de RX de pulmones sanos y de imágenes de RX de pulmones con la enfermedad de la Neumonía. Mediante el **Muestreo Sistemático** K = 4 se obtuvo n = 187.

5.2 Procedimiento

Primero se realizó búsquedas de CNN que tengan características de entradas semejantes, con la finalidad de que los resultados en el procesamiento de imágenes equivalente al tamaño de pixeles de 224x224. Segundo se evaluó y se ejecutó las arquitecturas con los parámetros originales y también variando otros parámetros para poder comparar los diferentes resultados para luego determinar el mejor entre las tres, como se detalla en la sección 6.2. Tercera aplicación de instrumentos de recolección de datos que se obtuvo de la data de la base de datos de kaggle [32] y se normalizaron las imágenes recolectadas para tener el mismo tamaño y renombrar, con el propósito que no exista error a la hora del entrenamiento y testeo. Cuarto Análisis de los datos, el cual se utilizaron los datos para la prueba de las CNN.

5.3 Técnicas e instrumentos

Registro de Base datos, CNN como: VGG16, ResNet50 y MobileNet, herramientas software iloveimg [34] para redimensionar a un solo tamaño y ReNamer [35] para renombrar en bloque las imágenes, el uso de Python con Google Colab con diferentes librerías como tensorflow, matplotlib, numpy, os, cv2, random y otros.

6 RESULTADOS

6.1 Distribución de las imágenes

En la tabla 1 presenta la distribución de las imágenes donde se estableció la división de datos basándonos en trabajos similares sugiere la proporción [36] para los datos del entrenamiento un 80% del total, y para los datos de validación el 20%, la normalización de las imágenes de RX utilizando un software iloveimg [34] para redimensionar las imágenes a un solo tamaño de 224x224 y para renombrar se utilizó el software ReNamer [35].

TABLA 1
Distribución de las imágenes en directorios

Directorio Principal	Directorio secundario	Tipo	Cantidad
TRAIN	NORMAL	JPG	1104
	PNEUMONIA	JPG	2984
TEST	NORMAL	JPG	203
	PNEUMONIA	JPG	543
VAL	NORMAL	JPG	276
	PNEUMONIA	JPG	746
TOTAL		JPG	5856

6.2 Análisis y ejecución de las Arquitecturas de Redes Neuronales Convolucionales

Se emplean funciones de pérdida específicas en el análisis de las tres CNN. Estas funciones son "categorical_crossentropy" (loss = "categorical_crossentropy", optimizer = "adam", métrica = "accuracy") y "binary_crossentropy" (loss = "binary_crossentropy", optimizer = "adam", métrica = "accuracy"). La elección de estas funciones se basa en la naturaleza de este trabajo, que implica la clasificación en dos categorías: radiografías de pulmones normales y radiografías con neumonía. Como resultado, se llevarán a cabo dos conjuntos de experimentos para cada CNN. Uno de ellos utilizará los valores predeterminados a 50 épocas, mientras que el otro variará la función de pérdida a "binary_crossentropy" y se entrenará a 25 épocas, manteniendo constante el optimizador Adam [37] y la métrica Accuracy [30].

6.2.1 Experimento con VGG16

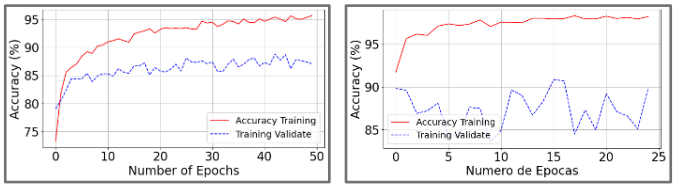


Fig. 6. Validación de la Arquitectura VGG16 de accuracy. **Izquierda:** Training accuracy con los parámetros predeterminados con 50 épocas de entrenamiento ambas curvas son similares dando un resultado acepatable. **Derecha:** Training accuracy con los parámetros modificados con 25 épocas de entrenamiento, las curvas son muy diferentes dando un resultado no aceptable.

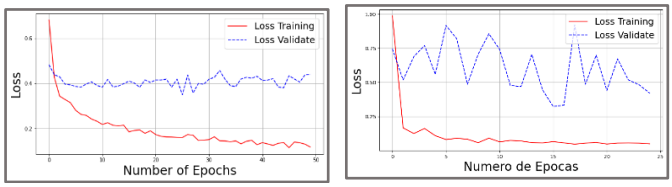


Fig. 7. Validación de la Arquitectura VGG16 de la función de pérdida. **Izquierda:** Training loss con los parámetros predeterminados con 50 épocas de entrenamiento ambas curvas con comportamiento similar y acepatable. **Derecha:** Training loss con los parámetros modificados con 25 épocas de entrenamiento, el comportamiento de la curva de validación es diferente cuyo resultado no es aceptable.

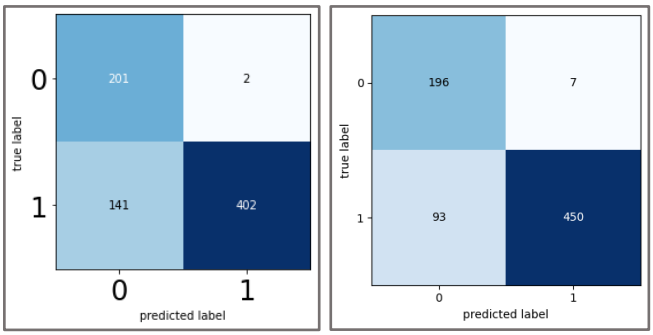


Fig. 8. Matriz de confusión de la Arquitectura VGG16 con 50 y 25 épocas de entrenamiento respectivo. **Izquierda:** es el resultado de una matriz de confusión con los parámetros predeterminados. **Derecha:** es el resultado de una matriz de confusión con los parámetros modificados.

A continuación, con los resultados de ambas matrices, calculamos el accuracy del experimento realizado con los valores originales de la CNN VGG16 con un resultado de 80.83% y también con los valores modificados de la función de pérdida con un resultado de 86.60%, como se muestran con la fórmula en ambos resultados.

$$Accuracy_{Predeterminado} = \frac{201 + 402}{201 + 402 + 2 + 141} = 80.83\%$$

$$Accuracy_{Modificado} = \frac{196 + 450}{196 + 450 + 7 + 93} = 86.60\%$$

6.2.2 Experimento con ResNet50

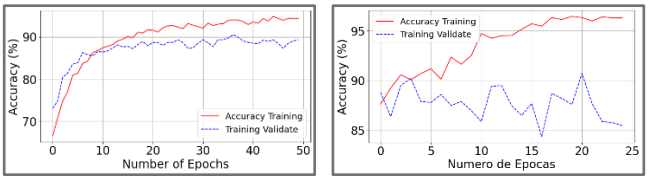


Fig. 9. Validación de la Arquitectura ResNet50 de accuracy. **Izquierda:** Training accuracy con los parámetros predeterminados con 50 epocas de entrenamiento cuyas curvas son idénticas y aceptable. **Derecha:** Training accuracy con los parámetros modificados con 25 épocas de entrenamiento cuyas curvas son diferentes y no es aceptable.

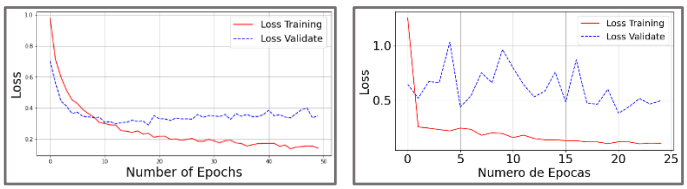


Fig. 10. Validación de la Arquitectura ResNet50 de la función de pérdida. **Izquierda:** Training loss con los parámetros predeterminados con 50 épocas de entrenamiento cuyas curvas con comportamiento similar y aceptable. **Derecha:** Training loss con los parámetros modificados con 25 épocas de entrenamiento cuya curva de validación es diferente a la del entrenamiento y no es aceptable.

C&T Riqchary
Revista de Investigación en Ciencia y Tecnología
ISSN: 2810-8124 (en línea) / ISSN: 2706-543x

Vol. 5 Núm. 1 (2023)
Publicado: 15/10/23
DOI: https://doi.org/10.57166/riqchary/v5.n1.2023.104

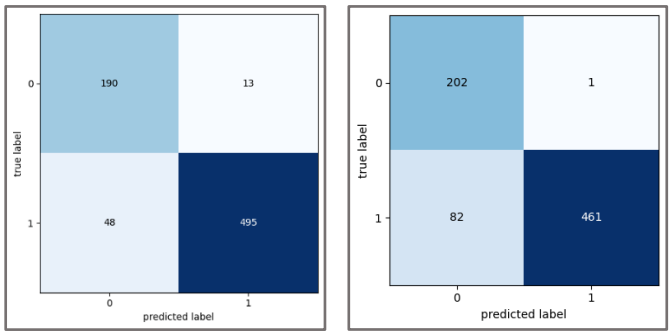


Fig. 11. Matriz de confusión de la Arquitectura ResNet50 con 50 y 25 épocas de entrenamiento respectivo. **Izquierda:** es el resultado de una matriz de confusión con los parámetros predeterminados. **Derecha:** es el resultado de una matriz de confusión con los parámetros modificados.

A continuación, con los resultados de ambas matrices, calculamos el accuracy tanto del experimento realizado con los valores originales de la CNN ResNet50 con un resultado de 91.82% de accuracy y también con los valores modificados de la función de pérdida con resultado de 88.87% de accuracy, tal como se muestran con la fórmula para ambos resultados.

$$Accuracy_{Predeterminado} = \frac{190 + 495}{190 + 495 + 13 + 48} = 91.82\%$$

$$Accuracy_{Modificado} = \frac{202 + 461}{202 + 461 + 1 + 82} = 88.87\%$$

6.2.3 Experimento con MobileNet

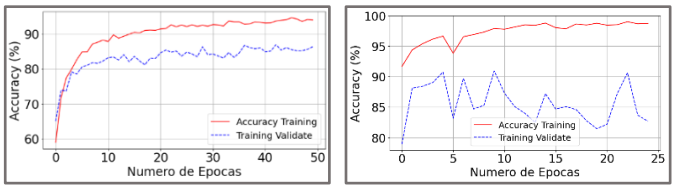


Fig. 12. Validación de la Arquitectura MobileNet de accuracy. **Izquierda:** Training accuracy con los parámetros predeterminados con 50 épocas de entrenamiento cuyas curvas con comportamiento similar y aceptable. **Derecha:** Training accuracy con los parámetros modificados con 25 épocas de entrenamiento y la curva de validación es diferente con respecto al de entrenamiento y no es aceptable.

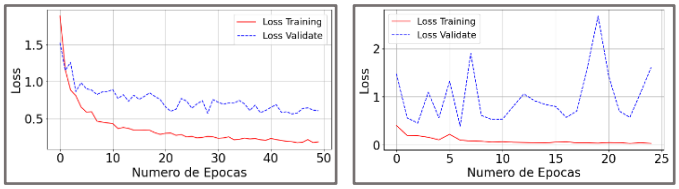


Fig. 13. Validación de la Arquitectura MobileNet de la función de pérdida. **Izquierda:** Training loss con los parámetros predeterminados con 50 épocas de entrenamiento con curvas con comportamiento semejante y aceptable. **Derecha:** Training loss con los parámetros modificados con 25 épocas de entrenamiento con la curva de validación diferente al de entrenamiento y no es eceptable.

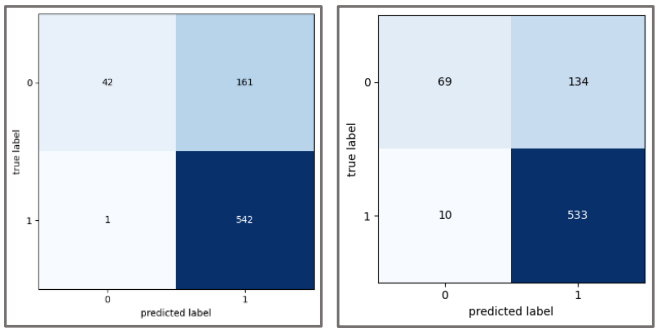


Fig. 14. Matriz de confusión de la Arquitectura MobileNet con 50 y 25 épocas de entrenamiento respectivo. **Izquierda:** es el resultado de una matriz de confusión con los parámetros predeterminados. **Derecha:** es el resultado de una matriz de confusión con los parámetros modificados.

A continuación, con los resultados de ambas matrices, calculamos el accuracy del experimento realizado con los valores originales de la CNN VGG16 con un resultado de 78.28% y también con los valores modificados de la función de perdida con resultado de 80.70%, tal como se muestran con la fórmula para ambos casos.

$$Accuracy_{Predeterminado} = \frac{42 + 542}{42 + 542 + 161 + 1} = 78.28\%$$

$$Accuracy_{Modificado} = \frac{69 + 533}{69 + 533 + 134 + 10} = 80.70\%$$

6.3 Resultados de las Arquitecturas VGG16, RESNET50 Y MOBILENET

En esta sección se presenta la matriz de confusión de las Arquitecturas de Redes Neuronales Convolucionales VGG16, ResNet50 y MobileNet, Matriz de observación y Métricas de evaluación, las cuales fueron calculadas a partir de la ecuación explicada en la sección 3.2

En la Figura 15 se observa la matriz de confusión obtenida al finalizar el proceso de validación de las 2 clases. En esta matriz se muestra el cálculo de los verdaderos positivos y negativos, falso positivos y negativos, para cada una de las dos clases donde se tiene que "0" representa a los datos de las imágenes de RX de personas sanas y "1" representa a los datos de las imágenes de RX de personas con Neumonía.

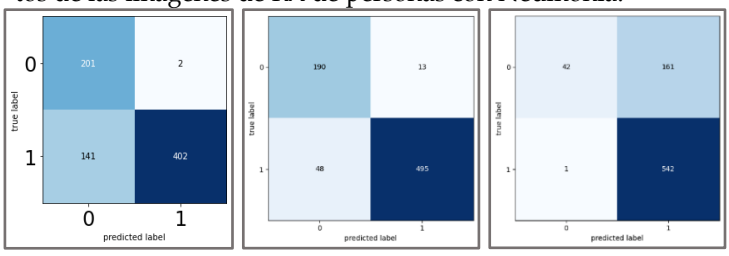


Fig. 15. Matriz de confusión de la Arquitectura con 50 épocas de entrenamiento. **Izquierda:** es el resultado de una matriz de confusión de VGG16 con los parámetros predeterminados. **Medio:** es el resultado de una matriz de confusión de ResNet50 con los parámetros predeterminados. **Derecha:** es el resultado de una matriz de confusión de MobileNet con los parámetros predeterminados.

A continuación, empleando los resultados obtenidos de las tres matrices, aplicamos la fórmula para calcular la precisión (accuracy) de las redes neuronales convolucionales VGG16, ResNet50 y MobileNet en los experimentos realizados utilizando sus valores originales o por defecto.

$$Accuracy_{VGG16} = \frac{201 + 402}{201 + 402 + 2 + 141} = 80.83\%$$

$$Accuracy_{ResNet50} = \frac{190 + 495}{190 + 495 + 13 + 48} = 91.82\%$$

$$Accuracy_{MobileNet} = \frac{42 + 542}{42 + 533 + 161 + 1} = 78.28\%$$

En el marco de esta investigación, se han obtenido resultados específicos que se detallan en la Tabla 2. En dicha tabla, se puede observar que la elección óptima para la Arquitectura de Redes Neuronales Convolucionales es la RESNET50. Por lo tanto, se sugiere emplear RESNET50 en la aplicación de detección de neumonía. Esto se recomienda con el propósito de proporcionar a los profesionales de la salud una herramienta esencial que les permita ofrecer diagnósticos más precisos en un tiempo más breve, evitando así largos períodos de espera para los resultados de los exámenes por parte de los pacientes.

TABLA 2
Resultado de las arquitecturas analizadas con accuracy

N°	Parámetros de entranamiento	Arquitectura	Accuracy
1	(loss='categorical_crossentropy', optimizer='adam', metrics=['accuracy']	VGG16	80.83%
		ResNet50	91.82%
		MobileNet	78.28%
2	(loss='binary_crossentropy', optimizer='adam', metrics=['accuracy']	VGG16	86.60%
		ResNet50	88.87%
		MobileNet	80.70%

Entonces, se determina que la Arquitectura de Red Neuronal Convolucional RESNET50 corresponde a la mayor puntuación de Accuracy, siendo la Arquitectura de Red Neuronal Convolucional RESNET50 con una predicción del 91.82%, como se muestra en la siguiente Figura 16.

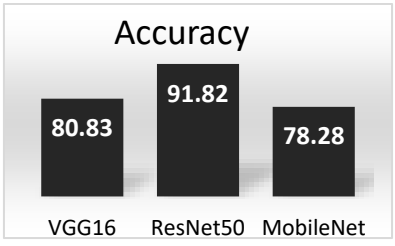


Fig. 16. Tasa de Reconocimiento de imágenes de RX de las Arquitecturas VGG16, ResNet50 y MobileNet

Estos resultados señalan claramente que, en lo que respecta a la predicción en el reconocimiento de imágenes de radiografías de neumonía, la Arquitectura ResNet50 es indiscutiblemente la que presenta el mayor porcentaje de acierto. En segunda posición se encuentra la Arquitectura VGG16, aunque con una diferencia considerable respecto a la primera. Por último, la Arquitectura MobileNet ocupa el tercer lugar y se caracteriza por tener un porcentaje muy bajo de acierto, lo que la hace no recomendable para este tipo de investigación, tal como se ilustra en la Figura 16.

6.4 Análisis estadístico

Se utilizó la prueba de Friedman al no poderce normalizar los datos, el cual consiste en una prueba no paramétrica desarrollado por el economista Milton Friedman. Para varias muestras relacionadas. Equivalente a la prueba ANOVA para medidas repetidas es la no paramétrica, el método consiste en ordenar los datos por filas o bloques, reemplazándolos por su respectivo orden [37] que se explica en la sección 6.5 de contrastación de hipótesis.

6.5 Contrastación de hipótesis

Resultado de las CNN VGG16, ResNet50 y MobileNet para determinar el que tiene mayor porcentaje de accuracy en el reconocimiento de Neumonía con imágenes RX.

Planteamiento de la Hipótesis

Ho: μ1 = μ2 = μ3 [No hay diferencia entre las arquitecturas VGG16, ResNet50 o MobileNet en la precisión de la detección de la Neumonía]

Ha: μ1 ≠ μ2 ≠ μ3 [Alguna de las arquitecturas VGG16, ResNet50 o MobileNet es diferente, en la precisión de la detección de la Neumonía]

Nivel de significancia: α = 0.05 y la confiabilidad del 95%

Estadístico

Tras intentar normalizar los datos y no lograrlo, se recurrió a la prueba no paramétrica de Friedman debido a que se obtuvo un valor de p < 0.01 y en la Figura 17, se puede observar claramente que la curva de color azul presenta una diferencia significativa en comparación con la pendiente de color rojo. Esto justifica la elección de la prueba de Friedman en estas circunstancias particulares.

Prueba de normalidad

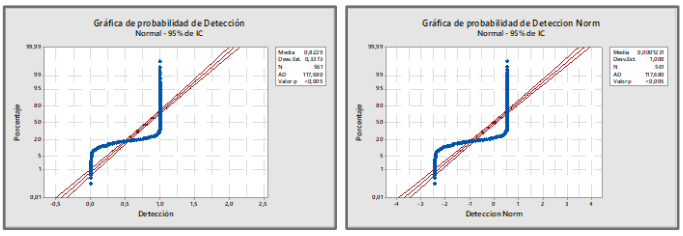


Fig. 17. La primera figura corresponde de accuracy de las tres arquitecturas sin normalizar y lo segundo corresponde datos que no cumplen con la normalización como p < 0.01, para los datos no normalizados corresponde a datos no paramétricos que utiliza la prueba de Friedman

C&T Riqchary
Revista de Investigación en Ciencia y Tecnología
ISSN: 2810-8124 (en línea) / ISSN: 2706-543x

Vol. 5 Núm. 1 (2023)
Publicado: 15/10/23
DOI: https://doi.org/10.57166/riqchary/v5.n1.2023.104

Fórmula de la prueba de Friedman

$$X^2 = \frac{12}{Nk(k+1)}\sum_{j=1}^{k} R_j^2 - 3N(k+1)$$

Donde: **k** el número de tratamientos, **N** es el número total de observaciones y **Rj** = es la suma de los rangos de los tratamientos en el j-ésimo bloque

TABLA 3
Resultado de estadístico

Estadísticos de prueba	Valores
N	187
Chi-cuadrado	11,797
gl	2
Sig. asintótica	,003

Región critica

Se rechaza la hipótesis nula $\boldsymbol{H_0}$ sí: $\boldsymbol{FrC > X^2}$

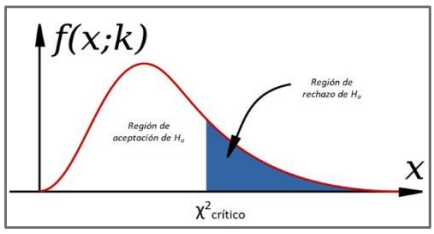


Fig. 18. Región de rechazo de $\boldsymbol{H_0}$

Decisión:
FrC = 11.797 > X^2 = 5.99, por tanto, rechazamos la Ho y aceptamos la Ha donde una de las arquitecturas VGG16, ResNet50 ó MobileNet es diferente, en la precisión de la detección de la Neumonía.

TABLA 4
Resultados del promedio de cada Arquitectura

Arquitecturas	Rango promedio
RESNET50	2,11
VGG16	1,80
MOBILENET	2,09

La Arquitectura ResNet50 lidera en términos de promedio, lo que confirma que es la superior entre las tres Arquitecturas de Redes Neuronales Convolucionales analizadas en este estudio de investigación, destacándose especialmente en la precisión del diagnóstico de la neumonía.

6.6 Discusiones

Tras realizar los experimentos en nuestro trabajo y determinar la mejor arquitectura con accuracy más alta en porcentaje siendo la CNN ResNet50 la que obtuvo este reconocimiento, antes de eso los datos no pudieron ser normalizados motivo, por el cual se utilizó la prueba no paramétrica de Friedman para así llegar al resultado esperado por tanto para trabajos similares corresponde realizar con esta prueba mencionada y con los resultados obtenidos respecto al accuracy de VGG16 se obtuvo resultados con parámetro original/modificado 80.83%/86.60% respectivamente, ResNet50 con resultados con parámetro original/modificado 91.82%/88.87% respectivamente y MobileNet de resultados con parámetro original/modificado 78.28%/80.70% y para realizar la comparación con los trabajos de Maldonado y otros [1] que obtuvieron resultados de accuracy de la implementación propia de 96.19%, InceptionV3 de 86.46% y VGG16 de 96.32%, Godoy y otros [2] llegaron a resultados de accuracy para resnet50 96.61%, VGG16 86.26%, Caya y otros [3] llegaron a los resultados de accuracy de la implementación propia, ResNet50 e InceptionV3, con 96.59%, 97.63% y 98.48%, Cabrejos y otros [4] utilizaron CNN como: AlexNet, VGG y ResNet-50 llegando al resultado de accuracy de 95.22%, 94.53 y 94.62% y finalmente Moura y otros [5] presenta CNN como MobileNetV1, MobileNetV2 e InceptionV3 con accuracy de 95.4%, 94.2% y 92.8%. Nuestros resultados, aunque menores en términos de porcentaje en comparación con algunos de estos estudios, presentan una ventaja al obtener una mayor precisión en dos de los resultados en comparación con dos trabajos previos. Es importante destacar que, en nuestro enfoque, realizamos dos experimentos para cada CNN, el primero con parámetros originales con 50 épocas, y el segundo modificando la función de pérdida de "categorical_crossentropy" a "binary_crossentropy", dado que nuestro análisis involucra dos clases, manteniendo la métrica de **accuracy** y el optimizador **adam** a 25 épocas.

7 CONCLUSIONES Y RECOMENDACIONES

7.1 Conclusiones

Aplicando los mismo parámetros de entrada para el preprocesamiento y conjunto de datos en las tres arquitecturas de CNN es posible realizar la comparación con la métrica de evaluación denominado "Accuracy" [13], obteniendo los siguientes resultados según el orden de análisis de la investigación: 80.83% de Accuracy para VGG16, un 91.82% de Accuracy para ResNet50 y por último con un 78.28% de Accuracy para MobileNet y según los resultados obtenidos muestran que la Arquitectura de CNN ResNet50 es más efectiva que sobresale con el resto de sus similares, por lo que se determina que es la mejor Arquitectura de Red Neuronal Convolucional.

7.2 Recomendaciones

Después de culminar el trabajo de investigación se recomienda utilizar una BD de la región de origen junto con la arquitectura de CNN ResNet50 para el desarrollo de aplicativos móviles o web como herramienta de apoyo a los profesionales en la salud, es por los altos valores obtenidos en la experimentación.

REFERENCIAS

[1] L. F. Maldonado y E. O. Moreano, «Diseño y aplicación de redes neuronales convolucionales para el diagnóstico de enfermedades pulmonares, a partir de radiografías de tórax frontal,» 2021.

[2] V. L. Godoy, «Algoritmo de diagnóstico preliminar de neumonía a partir de

imágenes radiográficas del tórax,» Lima, 2020.

[3] J. C. Caya, «Evaluación de Modelos de Redes Neuronales Convolucionales aplicado a Radiografías de Tórax, para apoyar al Proceso de Diagnóstico de Neumonía asociada al COVID-19,» Lima, 2020.

[4] V. M. Cabrejos, «Las Redes Neuronales Convolucionales y la mejora en el diagnóstico de Neumonía – área de Radiología,» Lima, 2022.

[5] J. V. Moura Sousa, W. Rosa de Almedia, A. Andrade Saraiva, D. B. Sousa Santos, P. M. Cunha Pimentel, Lopex de s y L. Lopez de Sousa, «Clasificación de imágenes de Neumonía en dispositivos móviles con Red Neuronal Cuantizada,» Brasil, 2020.

[6] P. R y G. A, «Introducción a la inteligencia artificial: sistemas expertos, redes neuronales artificiales y computación evolutiva,» 2001.

[7] G. Dalvinder Singh , «A Critical Conceptual Analysis of Definitions of Artificial Intelligence as Applicable to Computer Engineering,» IOSR Journal of Computer Engineering (IOSR-JCE), vol. 16, nº 2, pp. 9-13, 2014. https://doi.org/10.9790/0661-16210913

[8] A. D. Recuenco Cabrera y W. A. Reyes Alva, «Inteligencia artificial: Camino a un nuevo esquema del mundo,» Trijillo, 2020.

[9] M. A. Boden , «Inteligencia artificial,» España, 2017.

[10] A. Gerón, Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow, vol. Segunda , O'reilly, 2019.

[11] I. Goodfellow, Y. Bengio y A. Courville, «Deep Learning».

[12] B. Martin del Brio y C. Serrano Cinca, «Fundamentos de las redes neuronales artificiales: hardware y software,» Zaragoza, 1995. https://doi.org/10.54886/scire.v1i1.1036

[13] L. J. Sandoval, «Algoritmos de aprendizaje automático para análisis y predicción de datos,» 2018.

[14] Neuro, «VGG16 – Convolutional Network for Classification and Detection,» 28 Agosto 2022. [En línea]. Available: https://neurohive.io/en/popular-networks/vgg16/.

[15] A. Gómez, S. Tabik, J. Luengo y F. Herrera, «Redes Neuronales Convolucionales para Una Clasificacion Precisa de Imágenes de Corales. In XVIII Conferencia de la Asociación Española para la Inteligencia Artificial,» 2019.

[16] Abhijeet Pujara, «Image Classification With MobileNet,» medium, 10 12 2022. [En línea]. Available: https://builtin.com/machine-learning/mobilenet.

[17] Á. M. Botero y J. A. Chacon, El proceso diagnóstico, una aproximacion al raciocinio clínico, Caldas: Editorial Universidad de Caldas, 2007.

[18] OMS-Neumonia, «OMS-Neumonia,» 11 Noviembre 2021. [En línea]. Available: https://www.who.int/es/news-room/fact-sheets/detail/pneumonia.

[19] C. J. Álvarez, Neumonías: concepto, clasificación y diagnóstico dife-rencial.

[20] NIH, «National heart, lung, and blood institute,» 04 Agosto 2022. [En línea]. Available: https://www.nhlbi.nih.gov/.

[21] F. Saldías y J. Pavié, «Evaluación de la gravedad en la neumonía del adulto adquirida en la comunidad,» Revista chilena de enfermedades respiratorias, Santiago, 2005. https://doi.org/10.4067/S0716-10182005000400006

[22] N. Urbano Pintos, H. A. Lacomi y M. B. Lovarato, «B-VGG16: Red Neuronal de Convolución cuantizada binariamente para la clasificación de imágenes,» elektron ambar, Argentina, 2022. https://doi.org/10.37537/rev.elektron.6.2.169.2022

[23] K. Simonyan y A. Zisserman, «Very deep convolutional networks for large-scale image recognition,» 2015.

[24] É. Moreno Puig, «Reconocimiento de enfermedades a largo plazo,» 2022.

[25] K. He, X. Zhang, S. Ren y J. Sun, «Deep Residual Learning for Image Recognition,» arXiv, 10 Diciembre 2015. https://doi.org/10.1109/CVPR.2016.90

[26] A. G. Howard, Z. Menglong , C. Bo , K. Dmitry , W. Weijun , W. Tobias , A. Marco y A. Hartwig , «MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications,» 2017.

[27] Abhijeet Pujara, «medium,» 10 12 2022. [En línea]. Available: https://medium.com/analytics-vidhya/image-classification-with-mobilenet-cc6fbb2cd470.

[28] P. Sirawan y O. Surinta, «Food Image Classification with Improved MobileNet Architecture and Data Augmentation,» 2020.

[29] E. A. Huerta Mora, V. Gonzales Huitrón, H. Rodriguez Range y L. E. Amabilis Sosa, «Detección de enfermedades foliares con arquitecturas de redes neuronales convolucionales,» RINDERESU, Mexico, 2020.

[30] W. A. Lozada Portilla, M. J. Suarez Barón y E. Avendaño Fernandez, «Aplicación de redes neuronales convolucionales para la detección del tizón tardío Phytophthora infestans en papa Solanum tuberosum,» 2021.

[31] J. C. Valero Gómez, A. P. Zuñiga Incalla y J. C. Clares Perca, «Detección de la tuberculosis con algoritmos de Deep Learning en imágenes de radiografías del tórax,» VIVE, Moquegua, 2021. https://doi.org/10.33996/revista-vive.v4i12.119

[32] Kaggle, «Kaggle,» 2023. [En línea]. Available: https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia. [Último acceso: 22 Agosto 2022].

[33] R. Hernández, C. Fernández y M. d. P. Baptista, Metodología de la investigación, 6ta Edición ed., 2000.

[34] Iloveimg, «Iloveimg,» [En línea]. Available: https://www.iloveimg.com/es/redimensionar-imagen#resize-options,pixels. [Último acceso: 07 Julio 2023].

[35] softonic, «softonic,» 23 Mayo 2023. [En línea]. Available: https://renamer-free.softonic.com/. [Último acceso: 07 Julio 2023].

[36] A. Rácz, D. Bajusz y K. Héberger, «Effect of Dataset Size and Train/Test Split Ratios in QSAR/QSPR Multiclass Classification,» 2021. https://doi.org/10.3390/molecules26041111

[37] D. P. Kingma y B. Jimmy , «Adam: A method for stochastic optimization,» arXiv, 22 Diciembre 2014.

[38] A. Quispe A, K. M. Calla V, J. S. Yangali V, J. L. Rodríguez L y I. I. Pumacayo P, «Estadística no paramétrica aplicada a la investigación científica con software SPSS, MINITAB Y EXCEL,» EIDEC, Colombia, 2019.

Biografías

Sergio Elías Enciso Ortiz, Técnico Profesional en Electrónica en el Instituto Superior Tecnológico IDAT y bachiller en Ingeniería Informática y Sistemas de la Universidad Nacional Micaela Bastidas de Apurímac.

Ecler Mamani Vilca, Universidad Nacional Micaela Bastidas de Apurímac - Perú, Dr. en Ciencias de la Computación, desarrollador de aplicaciones multimedia.

Erech Ordoñez Ramos, catedrático Asociado en la Universidad Nacional Micaela Bastidas de Apurímac – Perú, Director de la Carrera Profesional de Ingeniería Informática y Sistemas.

C&T Riqchary
Revista de Investigación en Ciencia y Tecnología
ISSN: 2810-8124 (en línea) / ISSN: 2706-543x

Vol. 5 Núm. 1 (2023)
Publicado: 15/10/23
DOI: https://doi.org/10.57166/riqchary/v5.n1.2023.107

Análisis del afirmado con la adición de polímero acrilamida y cemento, Ocobamba, Chincheros, Apurímac 2023

Analysis of the affirmed with the addition of acrylamide polymer and cement, Ocobamba, Chincheros, Apurimac 2023

Isael Gomez-Villa[A], Sleyther Arturo De La Cruz-Vega[B]

ORCID:0000-0002-4808-9986[A], ORCID:0000-0003-0254-301X[B]

(Recepción: 07/08/2023 y aceptación 18/09/2023)

Resumen— Durante mucho tiempo en las redes viales con poca transitabilidad de vehículos se viene construyendo caminos no pavimentados donde la carpeta superior de este es la capa de afirmado, este material es de tipo granular que están expuestas a agentes externos que debilitan de forma directa la resistencia de la capa. El objetivo es determinar el comportamiento de capa del afirmado con la adición de polímero acrilamida y cemento, Ocobamba, Chincheros, Apurímac 2023. La metodología es de tipo aplicada, diseño experimental, enfoque cuantitativo la población y muestra que se evalúa es la cantera Chalhuany de 625 m2 que será utilizado para la estabilización con aditivo químico polímero acrilamida y cemento. En el ensayo realizado de CBR al 100% con una penetración de carga 0.1" (2.5 mm) con la dosificación de 0.3gr polímero acrilamida /litro de agua y 1.5% de cemento fue la mejor, mientras que con respecto al en el ensayo de corte directo con la dosificación de 0.3gr polímero acrilamida /litro de agua y 0.5% de cemento fue la mejor, por último, en el ensayo de permeabilidad usando el aditivo polímero acrilamida y cemento lo hacen más impermeables al material.

Palabras clave: polímero acrilamida, CBR, corte directo, permeabilidad

Abstract— For a long time, unpaved roads have been built in road networks with little trafficability for vehicles where the upper folder of this is the affirmed layer, this material is of a granular type that is exposed to external agents that directly weaken the layer resistance. The objective is to determine the layer behavior of the affirmed with the addition of acrylamide polymer and cement, Ocobamba, Chincheros, Apurímac 2023. The methodology is of the applied type, experimental design, quantitative approach, the population and sample that is evaluated is the Chalhuany quarry of 625 m2 that will be used for stabilization with acrylamide polymer chemical additive and cement. In the 100% CBR test with a 0.1" (2.5 mm) load penetration with the dosage of 0.3gr acrylamide polymer / liter of water and 1.5% cement was the best, while with respect to the test of direct cut with the dosage of 0.3gr acrylamide polymer / liter of water and 0.5% cement was the best, finally, in the permeability test using the acrylamide polymer additive and cement make it more impermeable to the material.

Keywords: acrylamide polymer, CBR, direct shear, permeability

A. Isael Gomez_Villa[A], Universidad Cesar Vallejo, igomezv@ucvvirtual.edu.pe
B. Sleyther Arturo De La Cruz_Vega[B], Universidad Cesar Vallejo, sdelacruzv@ucv.edu.pe

1 INTRODUCCIÓN

El deterioro prematuro de los afirmados en los caminos vecinales de bajo volumen de transito es muy frecuente por la baja resistencia del suelo, donde con lleva a una multitud de problemas para las conexiones viales entre comunidades, en muchos de los casos no solo se da por la baja calidad de los materiales, también tiene que ver las precipitaciones, las que ocasionan el desgaste de la capa de rodadura [1].
La falta de materiales de buena calidad cerca de los proyectos se ha vuelto un problema inminente, las principales causas son el desconocimiento de los aditivos estabilizadores químicos, por parte el desconocimiento de las autoridades municipales, así mismo la muy poca investigación de estudios de suelos que se realiza en las zonas rurales para proyectos de infraestructura vial [2].
El progresivo deterioro de la capa de rodadura, da como respuesta una disminución de los estándares de calidad de las características mecánicas y funcionales de los materiales y un aumento excesivo en los costos de mantenimiento y conservación vial, lo que este suceso ocasiona en corto periodo de tiempo es perjudicar la transitabilidad entre las zonas rurales y la zonas urbanas, así como también a los turistas, ya que mayormente las zonas turísticas están ubicadas en zonas rurales donde se hace muy complicado al acceso al desplazamiento, lo que hace un costo mayor al momento de desplazarse y a la vez se hace muy extendido el tiempo de viaje y de transporte de carga.
La propuesta del investigador es establecer los mecanismos adecuados para mejorar la trabajabilidad de la capa de afirmado empleando las metodologías constructivas viales desde un punto de vista normativo y técnico, usando los distintos modificadores como los aditivos de ceras, polímeros, aceites, entre otros. usando en esta investigación el polímero acrilamida dando una capa de rodadura con resistencia de (CBR >100%) y a su vez haciéndolo más impermeable a las precipitaciones fluviales.
El suelo es considerado como el material más antiguo en el sector de la construcción, complejo y utilizado por los ingenieros en la realización de ensayos de propiedades fisicoquímicas, mecánicas, resistencia, compresibilidad y permeabilidad, por último, la estabilidad volumétrica y durabilidad, estos ensayos son de gran importancia en la ingeniería para la realización de cualquier proyecto de ingeniera [3].
El suelo es un sustrato de una fracción de tierra o la acumulación de diferentes partículas mineralógicas en su estado no consolidado, está conformado por diferentes fases como: gaseosa, sólida y liquida, sobre esta se llevan a cabo distintas obras de ingeniería [4].
Los suelos pobres se caracterizan por presentar una resistencia muy baja y un alto contenido de humedad, siempre han generado problemas para los ingenieros de una forma peculiar durante o después de la construcción de la obra, para hacer del suelo un material adecuado se deben modificarse o diseñarse para adaptarse a los requisitos de un proyecto de infraestructura en particular [5].
La clasificación se lleva a cabo a través del ensayo análisis granulométrico, dicho ensayo se aplica en una muestra de suelos donde se hace referencia a la distribución de los diferentes tamaños de las partículas a través de una sucesión de tamices, cuyas ranuras son cuadradas y esta se ordena de mayor a menor [6].
El análisis granulométrico consiste en separar y clasificar por tamaños los granos que componen el suelo utilizando para ello mallas o tamices con diferentes aberturas, que está establecido en la norma de ASTM S136 que indica las diferentes medidas de las mallas como la malla No.200 (0.074mm), No.100 (0.149 mm), No.60 (0.25mm), No.50 (0.297 mm), No.30 (0.59mm), No.16 (1.19 mm), No.10 (2.00mm), No.4 (4.76 mm), 3/8" (9.5mm), 1/2" (12.7 mm), ¾" (19.1mm), 1" (25.4 mm), 1 ½" (38.1mm), 2" (50.8 mm) [7].
La compactación de pavimentos es un proceso de suma importancia en la construcción de nuevos pavimentos. Para la conformación de la estructura de nuevos pavimentos, se requiere que en su totalidad de las capas deben estar compactado [8].
La compactación de suelos se da cuando un esfuerzo vertical es aplicado a un suelo, donde esta causa la densificación a medida que el aire se desplaza de los poros entre los granos del suelo. La prueba de compactación de Proctor es un método de laboratorio de determinar la humedad óptima en la que un tipo de suelo dado se volverá más denso y alcanzará su máximo densidad seca [9].
Los ensayos para determinar las propiedades físicas y mecánicas; propiedades físicas se tiene: MTC E 110 - límite líquido, MTC E 111 - límite plástico, MTC E 107 - análisis granulométrico por tamizado, NTP 339.134 - clasificación de suelos método SUCS, NTP 339.135 - clasificación de suelos método AASHTO y las propiedades mecánicas se tiene: MTC E 115 - proctor modificado, MTC E 132 - california bearing ratio (CBR) [10].
La caracterización física indica que para determinar esta característica se debe evaluar el límite de Atterberg, contenido de humedad como también gravedad específica y por último la granulometría por tamizado y lavado sobre la malla N° 200 [11].
La estabilización química es una técnica que surgió como alternativa para modificar las propiedades de los suelos, esta técnica en mucho de sus aplicaciones mejora las propiedades físicas y mecánicas del suelo tales como la resistencia, plasticidad, compresibilidad, permeabilidad, resistencia a la compresión, resistencia al corte [12]. Esta estabilización se da de manera frecuente con los aditivos de polímeros, cemento, cal

C&T Riqchary
Revista de Investigación en Ciencia y Tecnología
ISSN: 2810-8124 (en línea) / ISSN: 2706-543x

Vol. 5 Núm. 1 (2023)
Publicado: 15/10/23
DOI: https://doi.org/10.57166/riqchary/v5.n1.2023.107

en la estabilización de los suelos, tienen como ventaja principal el tiempo de curado y fraguado, el costo de esta estabilización química hace más costosa en comparación con otras estabilizaciones [13].

La estabilización mecánica se realiza aplicando el proceso de compactación, este proceso de compactación tiende a mejorar las propiedades del suelo, causando el aumento de la capacidad de carga y de la resistencia, también ayuda en la reducción de la compresibilidad y aptitud para absorber agua [14].

Los Aditivos son sustancias de muy alto peso molecular que se caracterizan por ser reductores de agua, que actúan al ser adsorbidos en la superficie de las partículas de cemento, produciendo la dispersión de las partículas y reduciendo la cantidad de agua necesaria para un determinado nivel de consistencia. Según su composición, pueden estar basados en condensados de melamina-formaldehído, sulfonados, condensados de naftaleno formaldehído sulfonados, lignosulfonatos modificados (MLS) o polímeros acrílicos [15].

El ensayo California Bearing Ratio según Sharafi, Shekarbeigi fue introducida por Porter en 1926 y luego fue ampliada en 1929 por la administración de carreteras del estado de California, EE. UU. La prueba CBR es el método más común para determinar la resistencia relativa del suelo para la construcción de carreteras. El CBR de un suelo es la relación entre la fuerza requerida para sumergir un pistón con una figura, velocidad y profundidad definidas en un suelo probado y la fuerza requerida para sumergir el mismo pistón con la misma velocidad y profundidad en el material estándar, como se ve en la siguiente relación [16].

El corte directo es un ensayo de laboratorio donde al momento de su aplicación se obtiene el esfuerzo de cedencia al cortante y el esfuerzo de corte último, estos dos valores se utilizan para el análisis y diseño de elementos estructurales sujetos a cargas de flexión, el esfuerzo ultimo de corte si llega a ser superado por el esfuerzo de corte, produce las fallas en el elemento estructural, provocando esfuerzos en planos trasversales y longitudinales [17].

La permeabilidad es la capacidad que tiene una roca que consiente el paso de un fluido de agua o cualquier otro y la estimación de su coeficiente puede obtenerse a partir de su medición in situ y en laboratorio, se considera una técnica muy usada para determinar el coeficiente de permeabilidad en los suelos finos, arcillas, limos y hasta arenas finas. [18].

Este trabajo de investigación se tiene como objetivo, determinar el comportamiento de capa del afirmado con la adición de polímero acrilamida y cemento, Ocobamba, Chincheros, Apurímac 2023, como primer objetivo específico es determinar la variación del CBR de la capa de afirmado con la adición de polímero acrilamida y cemento, Ocobamba, Chincheros, Apurímac 2023, como segundo objetivo específico es determinar la variación del corte directo del afirmado con la adición del polímero acrilamida y cemento, Ocobamba, Chincheros, Apurímac 2023, como tercer objetivo específico es determinar la variación de la permeabilidad del afirmado con la adición del polímero acrilamida y cemento, Ocobamba, Chincheros, Apurímac 2023.

2 MATERIALES Y METODOS

Población:

Se evalúo la cantera de 625 m2 que será utilizado para la estabilización con aditivo químico polímero acrilamida y cemento para la carretera Ocobamba - Rocchacc, Chicheros, Apurímac.

Muestra:

Se tiene una cantidad determinada para el trabajo la muestra fue la cantidad del material de afirmado que obtuvo de las cuatro esquinas de la cantera cada lado tiene una medida de 25 m2, a este material extraído de la cantera se llegó a aplicar el agente estabilizante polímero acrilamida y un porcentaje de cemento en sus diferentes porcentajes, para la carretera del tramo Ocobamba – Rocchacc. Chincheros, Andahuaylas.

Muestreo:

Es así que de acuerdo con lo manifestado en este trabajo de investigación de llevo a cabo el muestreo probabilístico (aleatorio)

Procedimientos:

Segundo: La extracción y transporte del material de afirmado para su estabilización desde la cantera de estudio hasta el laboratorio donde se analiza el material sin adición de aditivo y con aditivo para la carretera Ocobamba – Rochacc.

Sexto: Se realizó el ensayo de CBR con una cantidad de material de afirmado con la aplicación de adiciones del aditivo químico polímero acrilamida de 0.3 gr/litro de agua con la adición de cemento del (0.5 %, 1.0 %, 1.5 %).

Séptimo: Se realizó el ensayo de corte directo con una cantidad de material de afirmado con la aplicación de adiciones del aditivo químico polímero acrilamida de 0.3 gr/litro de agua con la adición de cemento del (0.5 %, 1.0 %, 1.5 %).

Octavo: Se realizó el ensayo de permeabilidad con una cantidad de material de afirmado con la aplicación de adiciones del aditivo químico polímero acrilamida de 0.3 gr/litro de agua con la adición de cemento del (0.5 %, 1.0 %, 1.5 %).



Fig. 1. Extracción del material de la cantera Chalhuany



Fig. 2. Realización del ensayo de CBR



Fig. 3. Realización del ensayo permeabilidad

3 RESULTADOS Y DISCUCIONES

El aditivo empleado en este trabajo de investigación es el polímero acrilamida y cemento que se aplicó directamente en una fracción de material granular natural extraída de la cantera Chalhuany, se da una dosificación de polímero acrilamida una cantidad de 0.3 gr/litro de agua con la adición de cemento del (0.5 %, 1.0 %, 1.5 %) con la una finalidad de mejorar la resistencia del material de afirmado se obtiene los resultados en los siguientes ensayos a continuación:

TABLA 1
Resultados de CBR de muestra patrón y polímero Acrilamida y cemento

Espécimen	Polímero Acrilamida (gr/10 lt de agua)	Cemento (% en peso cemento)	CBR al 100% (MDS)	CBR al 95% (MDS)
Esp. 0	0.00	0.0%	37%	46%
Esp. 1	3.00	0.5%	72%	96%
Esp. 2	3.00	1.0%	80%	98%
Esp. 3	3.00	1.5%	104%	117%

En la tabla 1 se manifiesta los resultados que han sido obtenidos en el laboratorio de mecánica de suelos, con material de afirmado para fines de pavimentación donde se muestra el espécimen patrón y los especímenes con adiciones químicas con las proporciones de polímero acrilamida 0.3gr/litro de agua y cemento con (0.5%, 1%, 1.5%) dado de la siguiente forma; el espécimen 0 (Esp. 0) como espécimen patrón con 0gr de polímero acrilamida y 0.0% de cemento, espécimen 1 (Esp. 1) con 3gr de polímero acrilamida /10lt de agua y 0.5% de cemento, espécimen 2 (Esp. 2) con 3gr de polímero acrilamida /10lt de agua y 1.0% de cemento, espécimen 3 (Esp. 3) con 3gr de polímero acrilamida /10lt de agua y 1.5% de cemento, donde con los resultados del ensayo de CBR al 100% MDS con penetración de carga de 0.1" (2.5 mm) fueron: (Esp. 0) con un 37%, (Esp. 1) con un 72%, (Esp. 2) con un 80%, (Esp. 3) con un 104%. donde también se tiene los resultados del ensayo de CBR al 100% MDS con penetración de carga de 0.2" (5.08 mm) fueron: (Esp. 0) con un 37%, (Esp. 1) con un 72%, (Esp. 2) con un 80%, (Esp. 3) con un 104%.

Revista de Investigación en Ciencia y Tecnología
ISSN: 2810-8124 (en línea) / ISSN: 2706-543x

Vol. 5 Núm. 1 (2023)
Publicado: 15/10/23
DOI: https://doi.org/10.57166/riqchary/v5.n1.2023.107

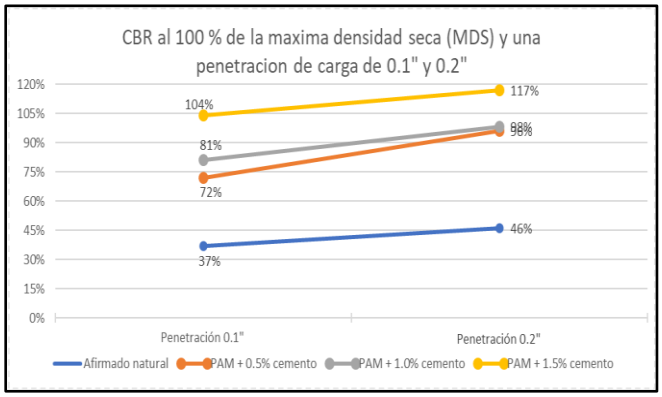


Fig. 5. Resultado de CBR al 100% con una penetración de carga de 0.1" y 0.2" en una (MDS

Para este ensayo de corte directo se hizo imposible extraer un espécimen inalterado, por el cual se usó un material afirmado suelto para poder compactarlo y así moldearlo en un recipiente de forma cuadrangular, donde el espécimen patrón y los especímenes con adiciones en sus distintas proporciones de polímero acrilamida 0.3gr/litro de agua y cemento con (0.5%, 1%, 1.5%) dados de esta forma; el espécimen (Esp. 0) como espécimen patrón 0gr de polímero acrilamida y 0.0% de cemento, (Esp. 1) con 3gr de polímero acrilamida /10lt de agua y 0.5% de cemento, (Esp. 2) con 3gr de polímero acrilamida /10lt de agua y 1.0% de cemento, (Esp. 3) con 3gr de polímero acrilamida /10lt de agua y cemento de 1.5% cada uno de estos especímenes fueron sometidos a distintos esfuerzos para obtener 2 parámetros de resistencia que son la cohesión que tiene como símbolo (C) y ángulo de fricción que tiene como símbolo (Ø), se menciona además que estos parámetros de resistencia singularmente se usan para determinar la capacidad admisible del suelo que se usa con fines de cimentación, dando como resultados lo que se expresa en la tabla 2, donde indica lo siguiente; el esp.0 tiene como C = 0.056 y Ø = 28.20, el esp.1 tiene como C = 0.101 y Ø = 29.40, el esp.2 tiene como C = 0.050 y Ø = 30.50, el esp.3 tiene como C = 0.040 y Ø = 31.40.

TABLA 2
Resultado del ensayo de corte directo con polímero acrilamida y cemento

Espécimen	Polímero Acrilamida (gr/10 lt de agua)	Cemento (% en peso cemento)	Parámetros de resistencia Cohesión (Kg/cm2)	Angulo de fricción (°)
Esp. 0	0.00	0.0%	0.056	28.20
Esp. 1	3.00	0.5%	0.101	29.40
Esp. 2	3.00	1.0%	0.050	30.50
Esp. 3	3.00	1.5%	0.040	31.40

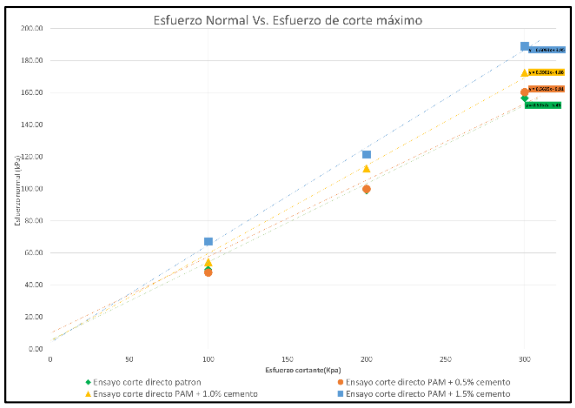


Fig. 6. Esfuerzo normal vs esfuerzo de corte máximo

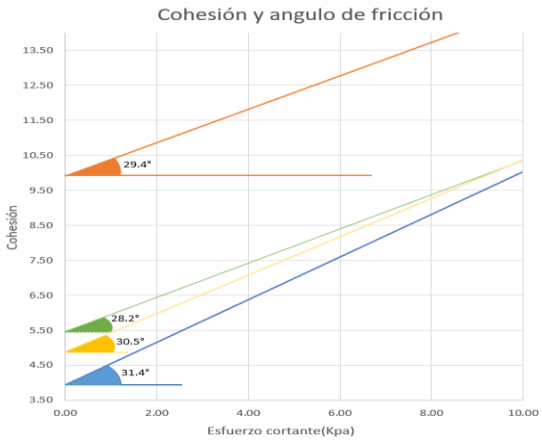


Fig. 7. Cohesión y ángulo de fricción

Para el ensayo de permeabilidad se usó un material de afirmado con la fracción fina la cual pasa el tamiz N°40 donde se moldeo de forma cilíndrica bajo criterio de la norma europea EN 772-11:2011, estos especímenes fueron moldeados de forma cilíndricas fueron realizados con las dosificaciones el espécimen (Esp. 0) como espécimen patrón 0gr de polímero acrilamida y 0.0% de cemento, (Esp. 1) con 3gr de polímero acrilamida /10lt de agua y 0.5% de cemento, (Esp. 2) con 3gr de polímero acrilamida /10lt de agua y 1.0% de cemento, (Esp. 3) con 3gr de polímero acrilamida /10lt de agua y cemento de 1.5%, este ensayo fue realizado con la única finalidad de medir la facilidad del agua con la que atraviesa el molde esto es llamado coeficiente de permeabilidad como se muestra en la tabla 3 para (esp. 0) se tuvo 6.28E-06, 6.37E-06, 6.37E-06, para el (esp. 1) se tuvo 9.06E-06, 8.27E-06, 6.93E-06, para el (esp. 2) se tuvo 1.39E-05, 1.21E-05, 1.15E-05, para el (esp. 3) se tuvo 2.48E-05, 2.94E-05, 1.62E-05.

TABLA 3
Resultado del ensayo de permeabilidad con polímero acrilamida y cemento

N	Polímero PAM	Cemento	suelo (gr)	Cwi	Ks mm/min	Promedio Ks
	0.00	0.00	1000	2653	6.28E-06	
Esp. 0	0.00	0.00	1000	2617	6.37E-06	6.34E-06
	0.00	0.00	1000	2617	6.37E-06	
	3.00	0.50	994.5	1839	9.06E-06	
Esp. 1	3.00	0.50	994.5	2016	8.27E-06	8.09E-06
	3.00	0.50	994.5	2405	6.93E-06	
	3.00	1.00	989.5	1203	1.39E-05	
Esp. 2	3.00	1.00	989.5	1379	1.21E-05	1.25E-05
	3.00	1.00	989.5	1450	1.15E-05	
	3.00	1.50	984.5	672	2.48E-05	
Esp. 3	3.00	1.50	984.5	566	2.94E-05	2.35E-05

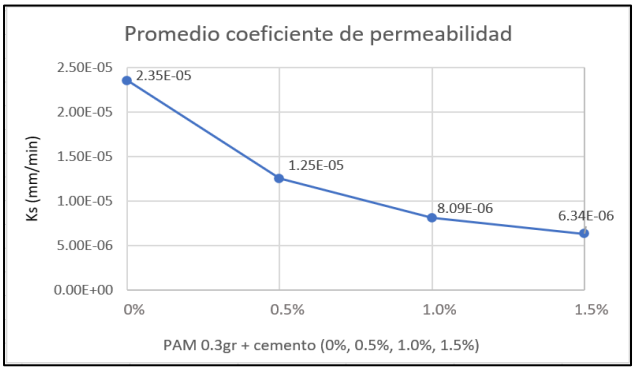


Fig. 8. Promedio coeficiente de permeabilidad

Los ensayos realizados con material de afirmado proveniente de la cantera Chalhuany siguiendo las normas técnicas del MTC de la entidad máxima de carreteras del ministerio de transporte y comunicaciones, donde los resultados obtenidos mejoraron en gran escala la resistencia del material de afirmado que se realizó con el ensayo de CBR al 100% con la adición de 0.3 gr polímero acrilamida y 1.5 % de cemento dando como valor referencial 104% en comparación con muestra patrón que fue 37%, en la presente investigación se dio un CBR a un 100% (MDS) con un 0.1" de penetración, entonces de acuerdo a los resultados adquiridos se puede manifestar del fig. 1 que con la adición del aditivo polímero acrilamida se genera una recta de tendencia ascendentes. estos resultados obtenidos concuerdan con la tesis peruana donde se usó el polímero acrilamida y cemento donde se tuvo como objetivo comparar suelos granulares en sus características de resistencia y durabilidad entre dos aditivos que son Proes y Polycom y ambas con dosificaciones de cemento Portland Tipo I. donde se dio el ensayo de CBR a un 100% (MDS) con un 0.1" de penetración con adiciones de 0.3 gr polímero/litros de agua y con 1.5 % de cemento dando como valor referencial 124% en comparación con muestra patrón que fue 63%, en la presente investigación se dio un CBR a un 100% (MDS) con un 0.1" de penetración con adiciones de 0.3 gr polímero/litros de agua y con 1.5 % de cemento dando como valor referencial 104% en comparación con la muestra patrón que fue 37%, esto tiene un valor aproximado de coherencia, ya que en ambas investigaciones el polímero mencionado muestra un resultado ascendente [19].

Los ensayos realizados con respecto al ensayo de corte de directo se hacen con fines de cimentación, en esta investigación se usó el aditivo polímero acrilamida en 0.3gr/litro de agua y 0.5% de cemento dando una mejora las características del suelo en la cohesión y en el ángulo de fricción de C = 0.056 y Ø = 28.20 con respecto al espécimen 1 patrón que muestra C = 0.101 y Ø = 29.40 como se muestra el grafico 2 con esto se demuestra que la mejora que hay entre el suelo granular de afirmado natural y suelo granular de afirmado estabilizado con aditivo. estos resultados obtenidos concuerdan con la tesis colombiana donde se hiso el uso de nuevas tecnologías como fluidos estabilizantes de suelos que son llamados polímeros, los suelos estabilizados con aditivos químicos buscan generar una disminución de ph y brinda una cohesión entre las partículas garantizando una construcción segura, se debe contar con la capacitación de expertos y del fabricante para usar el tipo de polímero en el suelo adecuado, uno de los parámetros a tener en cuenta son el medio ambiente, es así que los polímeros a comparación con las bentonitas son eco amigables siendo así reutilizables bajo un proceso adecuado y no requiere de un tratamiento complicado para ser desechados [20].

Los resultados obtenidos del ensayo del Coeficiente de Permeabilidad que se llevó acabo aplicando la norma europea EN 196-1:2018 de las muestras cilíndricas, realizados en el laboratorio de Geocon S.A.C – Andahuaylas, muestran que las muestras cilíndricas patrón tienen un Coeficiente de Permeabilidad promedio de 3 especímenes dando un valor de referencial 6.34E-06, mientras que los 3 especímenes con aditivo de polímero acrilamida 0.3gr/litro de agua y cemento 1.5% que se realizaron de forma cilíndricas son experimentales que muestran según la tabla 3 tienen un Coeficiente de Permeabilidad promedio de 2.35E-05, de manera que el Coeficiente es de forma ascendente con respecto a los del patrón, estos resultados obtenidos concuerdan con la tesis peruana que manifiesta que en el ensayo del Coeficiente de Permeabilidad de las muestras ha alcanzado un avance de 65.85%, calificando dentro del rango de 1.04 E^{-04} a 1.36 E^{-}

C&T Riqchary
Revista de Investigación en Ciencia y Tecnología
ISSN: 2810-8124 (en línea) / ISSN: 2706-543x

Vol. 5 Núm. 1 (2023)
Publicado: 15/10/23
DOI: https://doi.org/10.57166/riqchary/v5.n1.2023.107

º ⁶ , siendo favorable ya que estos compuestos forman una matriz que contribuye a la resistencia del suelo, que es transformado de su alta expansividad, de un estado natural indeseable a más granular, a un material relativamente impermeable, durable, resistente a los agrietamientos [21].

4 CONCLUSIONES

En el ensayo realizado de CBR al 100% con una penetración de carga 0.1" (2.5 mm) se llegó a la conclusión que en los valores obtenidos entre el espécimen 0 de 37% y los especímenes (1,2,3) dieron un valor de 72%, 80%, 104%, los valores estabilizados con el aditivo polímero acrilamida y cemento que son estos últimos especímenes mencionados han llegado a mejorar indispensablemente y a su vez han superado el valor mínimo establecido por la norma, la dosificación optima obtenido es de 0.3gr polímero acrilamida /litro de agua y 1.5% de cemento tipo portland que dio un valor de CBR al 100% con una penetración de carga 0.1" (2.5 mm) 104% que con respecto al espécimen patrón que se obtuvo un valor de 37% donde se aprecia una mejora de un 141%.

En el ensayo de corte directo se llegó a la conclusión que en los valores obtenidos de la cohesión y el ángulo de fricción entre y el espécimen (0), C = 0.056 y Ø = 28.20 y los especímenes que son estabilizados con polímero acrilamida y cemento tipo portland que son: espécimen (1) (C = 0.101 y Ø = 29.40), espécimen (2) (C = 0.050 y Ø = 30.50), espécimen (3) (C = 0.040 y Ø = 31.40), la característica de la cohesión muestra ascendencia con el espécimen (1) y en las demás especímenes (2,3) se nota una descendencia esto se explica el extraño comportamiento del suelo, por otra parte en la característica del ángulo de fricción el suelo granular muestra un resultado ascendentes, se puede manifestar que la dosificación optima se da con el espécimen (1) mostrando 0.3gr de polímero acrilamida y 0.5% de cemento tipo portland que dieron un valor de (C = 0.101 y Ø = 29.40) la cohesión y el ángulo de fricción mejoraron simultáneamente.

Se llegó a realizar el ensayo de permeabilidad llegando a la conclusión que el espécimen (0) sin aditivo alguno muestra un valor promedio que fue 6.34E-06 y los especímenes que son estabilizados con polímero acrilamida y cemento tipo portland que son el espécimen (1) un valor promedio de 8.09E-06, el espécimen (2) un valor promedio de 1.25E-05, espécimen (3) un valor promedio de 2.35E-05, estos especímenes (1,2,3) estabilizados con aditivos demuestran que se hacen más impermeables en comparación del espécimen (0) que es de tipo patrón.

5 REFERENCIAS

[1] S. Villanueva, «Propuesta de estabilización de carreteras de bajo volumen de tránsito en la sierra, sobre los 2000 m.s.n.m, utilizando poliacrilamida aniónica, organosilano y un sulfonatado. caso: Poncos – Kochayoc, departamento de Ancash», Universidad Ricardo Palma, Lima Peru, 2017. https://doi.org/10.23850/22565035.2530

[2] E. Pacuy, «Estabilización de subrasante con adición de ceniza de hojas de maíz (Zea Mays), Yungay, Ancash», Universidad Cesar Vallejo, Huaraz Peru, 2022. https://doi.org/10.1002/14356007.b07_613.pub3

[3] J. Rivera, A. Aguirre, R. Mejia, A. Orobio, Estabilización química de suelos - materiales convenciales y activados alcalinamnete. Informador tecnico, 84(2), 202-226, 2020 https://doi.org/10.23850/22565035.2530

[4] S. Nortcliff, H. Hulpke, C. Bannick, K. Terytze, G. Knnop, M. Bredemeier & H. Schulter, Definition, Function and Utilization of Soil. Ullman´s Encyclopedia of Industrial Chemistry, 1(33), 399-419, 2012 https://doi.org/10.1002/14356007.b07_613.pub3

[5] J. Jijo, Strength benefit of sawdust/wood ash amendment in cement stabilization of an expansive soil, Revista Facultad de Ingeniería, 28(50), 44-61, 2019. https://doi.org/10.19053/01211129.v28.n50.2019.8790

[6] O. Palacio, Á Chávez, Y. Velásquez, evaluación y comparación del análisis granulométrico obtenido de agregados naturales y reciclados, Tecnura, 21(53), 96-106, 2017. https://doi.org/10.14483/22487638.8195

[7] M. Payán, estabilización de sahcab con cemento Pórtland Tipo I para la construcción de pisos en viviendas rurales, Ingenieria, 6(2), 33-45, 2002.

[8] M. Zvonari´c, I. Bariši´c, M. Gali´c & K. Minažek, Influence of Laboratory Compaction Method on Compaction and Strength Characteristics of Unbound and Cement-Bound Mixtures. Applied Sciences, 11(11), 1-12, 2021. https://doi.org/10.3390/app11114750

[9] R. Grosh, Effect of soil moisture in the analysis of undrained shear strength of compacted clayey soil, Journal of Civil Engineering and Construction Technology, 4(1), 23-31, 2013.

[10] R. Linarez, M. Aguilar, E. Rojas, estabilización de suelos arcillosos a nivel de subrasante con adición de bolsas de polietileo fundido, Revista de investigacion cientifica UNTRM, 3(2), 33-40. 2020). https://doi.org/10.25127/ucni.v3i2.612

[11] Y. Valencia, J. Patiño, M. Álvarez & Ó. Echeverri, cambio en las propiedades geotécnicas de un suelo sometido a ignición en laboratorio, Revista Ingenierías Universidad de Medellín, 17(32), 85-107, 2018. https://doi.org/10.22395/rium.v17n32a5

[12] E. Llano, D. Rios & G. Restrepo, evaluación de tecnologías para la estabilización de suelos viales empleando intemperismo acelerado una estrategia de análisis de impactos sobre la biodiversidad, Tecno Logicas, 23(49), 185-199, 2020. https://doi.org/10.22430/22565337.1624

[13] J. Alarcon, M. Jimenez & R. Benitez, Stabilization Of Soils Through The Use Of Oily Sludge.Revista ingenieria de construccion RIC, 35(1), 5-20, 2020.

[14] C. Higuera, J. Gómez, Ó. Pardo, caracterización de un suelo arcilloso tratado con hidróxido de calcio. Facultad de ingenieria, 21(32), 21-40. 2012.

[15] M. Sosa, Y. Villagrán, J. Peralta & C. Zega, Efficiency Of Cement-Admixture Systems In Mortars With Binary And Ternary Portland Cements. DYNA, 85(204), 134-142, 2018. https://doi.org/10.15446/dyna.v85n204.66468

[16] H. Sharafi & M. Shekarbeigi, Experimental Evaluation Of The Behavior Of Sandy Soil–Cement Mixture. Revista INGENIERIA UC, 26(3), 258-272, 2018.

[17] E. Marquez, A. Meza, diseño de dispositivo para ensayos de corte directo en concreto. Conciencia tecnologica, 1(59), 1-18, 2020.

[18] C. Camacho, N. Zabala, Y. Flores, estimación del coeficiente de permeabilidad del suelo en la ciudad de Macas -. Dominio de las ciencias, 6(3), 78-94, 2020.
[19] D. Guerra, D. Quispe, Estudio comparativo entre los aditivos proes y polycom en ensayos de resistencia y durabilidad para afirmados estabilizados, Ayacucho, Universidad Cesar Vallejo, Lima Perú, 2021.
[20] S. Mora, J. Tiusara, Comportamiento de los polímeros como agente estabilizante en los suelos para la construcción de cimentaciones, Universidad Católica de Colombia, Bogotá, Colombia, 2020.
[21] M. Tolentino, Permeabilidad del suelo con adición del 10% de ceniza de concha de abanico, carretera Cambio Puente – Cascajal, Universidad San Pedro, Chimbote Perú, 2018.

C&T Riqchary
Revista de Investigación en Ciencia y Tecnología
ISSN: 2810-8124 (en línea) / ISSN: 2706-543x

Vol. 5 Núm. 1 (2023)
Publicado: 15/10/23
DOI: https://doi.org/10.57166/riqchary/v5.n1.2023.109

Tutoría virtual en la formación académica y humana de estudiantes de la UNAMBA

Virtual tutoring in the academic and human formation of UNAMBA students

Francisco Cari-Incahuanaco[A], Yhon Fuentes-Huamán[B], y Maritza Condori-Quispe[C]

ORCID:0000-0002-2807-0495[A], ORCID:0000-0003-4228-1982[B] y ORCID:0009-0000-6890-5453[C]

(Recepción: 15/07/2023 y aceptación 10/08/2023)

Resumen— El tiempo ha sido testigo primigenio en la evolución del conocimiento humano, día a día aparecen, fruto del ingenio del hombre, una cantidad impresionante de novedades tecnológicas que rompen todo tipo de barreras y aumentan el interés por diseñar nuevas herramientas virtuales que conduzcan hacia el aprendizaje multidisciplinario[1]. Esta investigación demuestra la contribución de la tutoría virtual en la formación académica y humana de estudiantes de la Universidad Nacional Micaela Bastidas de Apurímac; la tutoría virtual es un proceso de acompañamiento durante la formación de los estudiantes universitarios, mediante la atención personalizada a estudiantes por parte de los docentes tutores, apoyándose conceptualmente en teorías del aprendizaje que posibilitan un mejor acercamiento a los estudiantes, cuya labor del docente tutor es ayudar al estudiante a aprender para culminar con éxito su carrera profesional, además es muy importante mencionar la existencia de una relación mutua entre tutor y tutelados, en esta interrelación la comunicación virtual debe ser fluida y armoniosa entre ambos actores. Los resultados que se obtuvieron en las pruebas de hipótesis específicas, demuestran que la tutoría virtual contribuye significativamente en la formación académica de los estudiantes de la UNAMBA; así como también las encuestas demuestran que la tutoría virtual contribuye en la formación humana de los estudiantes de la UNAMBA.

Palabras clave: Tutoría virtual, tutor, e-learning, formación humana.

Abstract— Time has been a primordial witness in the evolution of human knowledge; day by day, as a result of human ingenuity, an impressive amount of technological novelties appear, breaking all kinds of barriers and increasing the interest in designing new virtual tools that lead to multidisciplinary learning [1]. This research determines the contribution of virtual tutoring in the academic and human formation of students of the Universidad Nacional Micaela Bastidas de Apurímac; virtual tutoring is a process of accompaniment during the training of tutored students, through personalized attention to students by tutor teachers, conceptually supported by learning theories that enable a better approach to students, whose work of the tutor teacher is to help students learn to successfully complete their professional career, it is also very important to mention the existence of a mutual relationship between tutor and tutored, in this interrelationship virtual communication should be fluid and harmonious between the two actors. The results obtained in the specific hypothesis tests show that virtual tutoring contributes significantly to the academic formation of UNAMBA students; thus, the surveys show that virtual tutoring contributes to the human formation of UNAM-BA students.

Keywords: Virtual tutoring, virtual tutor, human formation, TIC.

A. *Francisco Cari_Incahuanaco, Departamento Académico de Informática y Sistemas, UNAMBA,* fcari@unamba.edu.pe
B. *Yhon Fuentes_Human, Departamento de Ingeniería, UTEA,* yfuentes@unamba.edu.pe
C. *Maritza Condori Quispe, Departamento Académico de Humanidades, UNAMBA,* mcondoriq@gmail.com

1 INTRODUCCIÓN

La pandemia causada por el brote de la COVID-19 obligó a las instituciones de educación superior a transitar de forma emergente a la enseñanza remota [2]. El proceso de acompañamiento en el desarrollo personal, social, psicológico, emocional y cognitivo de las personas resulta de vital importancia, debido a que para lograr los objetivos y metas en las diferentes esferas de la vida el hombre necesita de alguien que lo oriente, guie y acompañe en las acciones y procesos a realizar [3]. La tutoría virtual consiste en la comunicación establecida en tiempo real entre el tutor y los tutelados en asignaturas que desarrolla su proceso de formación académica y humana a través del uso de las Tecnologías de la Información y Comunicación (TIC). Mediante esta tecnología permite dar seguimiento y orientación a los estudiantes, así como el fortalecimiento de sus capacidades innovadoras en el desarrollo de las sesiones didácticas. Mientras Toledo indica que la tutoría a través de las aulas virtuales, es una experiencia de formación y autoformación de profesionales[1]. Según Contreras Navarro, se concluye que la tutoría virtual en la educación peruana es considerada como un soporte importante dentro

de la estrategia aprendo en casa que contribuye de manera efectiva en la mejora de los aprendizajes y la continuidad de la enseñanza [4].

El principal objetivo de la Tutoría Académica es acompañar a los estudiantes de pregrado durante su etapa universitaria, como acción inherente y complementaria de la labor docente, creando espacios para que los estudiantes puedan informarse, dialogar y reflexionar. Cuenta con una normatividad a través de un reglamento y el apoyo de docentes tutores y asesores académicos. En la UNAMBA la tutoría académica no se desarrolla adecuadamente, los estudiantes participan irregularmente en las horas programadas, no se realiza un control y seguimiento adecuado de las actividades realizadas, los informes se presentan sólo para cumplir y el historial se queda en el olvido. Para demostrar la eficiencia de esta labor tutorial se realizó una comparación entre las clases presenciales antes de la pandemia y las clases virtualmente durante la pandemia.

Por último, el tener una buena formación humana impulsará a los estudiantes a dar una respuesta a la necesidad de acompañamiento de personas que se encuentran en soledad o que se estén aisladas de la sociedad y así mejorar la calidad de vida de las personas que se encuentran en situaciones desfavorables y potenciar así las relaciones sociales y de comunicación en el ambiente que se desarrollen.

2 MATERIALES Y MÉTODOS

El presente trabajo se ubica, bajo el criterio del enfoque investigativo propuesto por Hernández, Fernández y Baptista dentro de la investigación cuantitativa. Según propósito de la investigación, el presente estudio reúne las condiciones metodológicas de una investigación aplicada, en razón, que se utiliza los conocimientos de la tecnología informática con el propósito de aplicarla en el proceso de la tutoría académica. Nivel descriptivo-correlacional [5]. El diseño de investigación es no experimental de tipo transeccional, porque la investigación se centra en analizar cuál es el nivel o estado de una variable en un momento dado.
La población a utilizar durante el desarrollo del estudio está compuesta por 570 asignaturas y 3870 estudiantes de la UNAMBA, matriculados en el semestre académico 2019-II según servicios académicos. Se utilizó el muestreo no probabilístico en su tipificación de muestreo por conveniencia, en razón de no existir control alguno sobre la población estudiantil, se tomó 66 asignaturas de diferentes áreas de estudio para su análisis. Técnicas e instrumentos de investigación se realizó al inicio y al final de la investigación utilizando los instrumentos propuestos que a continuación se detalla:

- Registro de notas del contenido conceptual
- Registro de notas actitudinal
- Encuesta a estudiantes para determinar la interrelación humana

Entrevista: Se realizó a los estudiantes de la UNAMBA durante el proceso de desarrollo de la tutoría, con la finalidad de obtener los conocimientos necesarios sobre las diferentes actividades que componen la tutoría académica.
Encuesta: se aplicó las encuestas a los estudiantes de la UNAMBA para obtener información sobre la tutoría que se realiza con los docentes tutores.
Procesamiento y análisis de datos: La información obtenida con las diferentes técnicas de investigación se procedió a procesar en una hoja de cálculo Excel y en el paquete estadístico SPSS, para luego ser interpretados de acuerdo a los resultados obtenidos.

3 RESULTADOS

3.1 Hipótesis nula y alterna de la formación académica

Se formula la hipótesis nula, que la tutoria virtual no contribuye en la formación académica de los estudiantes de la UNAMBA, y la hipotesis alterna se formula que la tutorial virtual contribuye en la formación académica de los estudiantes de la UNAMBA, con un nivel de significancia de $\alpha = 0.05$, es decir, con una confianza del 95%. En este análisis se ve por conveniente utilizar el estadístico Z para la diferencia de medias, ya que la muestra es mayor que 30, en donde se obtiene el siguiente resultado: $Z_c = -1.69$.

Según la regla de decision: Si $Z_c < Z_t$, se rechaza la H_0, caso contrario se acepta la H_a, entonces, como $Z_c = -1.69 < Z_t = -1.64$, se rechaza la H_0, se acepta la hipótesis H_a; es decir, La tutoria virtual contribuye en la formación académica de los estudiantes de la UNAMBA, por lo que sí existe diferencia entre los promedios obtenidos en el promedio de porcentaje de notas en los semestres académicos 2019-II y 2021-I.

Hipótesis nula y alterna para formación humana
Se formula la hipotesis nula, la tutoria virtual no contribuye en la formación humana de estudiantes de la UNAMBA y la hipotesis alterna, sí contribuye significativamente en la formación humana de los estudiantes de la UNAMBA, con un nivel de significancia α=0.05 es decir con una confianza del 95%. En este análisis se ve por conveniente utilizar el estadístico Z para la diferencia de medias ya que la muestra es mayor a 30, en donde se obtiene el siguiente resultado: $Z_c = -3.17$

Según la regla de decisión: Si $Z_c < Z_t$, se rechaza la H_0, caso contrario se acepta la H_a, entonces como $Z_c = -3.97 < Z_t = -1.64$, se rechaza la H_0, se acepta la hipótesis H_a; es decir, la tutorial virtual contribuye significativamente en la formación humana de los estudiantes de la UNAMBA, por lo que sí existe diferencia entre los promedios obtenidos en los dos semestres académicos 2019-II y 2021-II. Para esta decisión se utilizó el porcentaje de notas actitudinales.

3.2 Decisión para la interrelación docente-estudiante

Para la toma de decisión de interrelación docente-estudiante, se realiza de acuerdo al cuadro de distribución de frecuencias obtenida en la tabla 1, se concluye que efectivamente la tutorial virtual contribuye en la interrelación entre docente y estudiante de la UNAMBA, por lo que es muy importante mantener la relación mutua entre docentes y estudiantes y esta relación facilitará la comunicación eficaz entre docentes y estudiante, así contribuyendo eficazmente en la formación académica y humana del estudiante tutelado.

C&T Riqchary
Revista de Investigación en Ciencia y Tecnología
ISSN: 2810-8124 (en línea) / ISSN: 2706-543x

Vol. 5 Núm. 1 (2023)
Publicado: 16/10/23
DOI: https://doi.org/10.57166/riqchary/v5.n1.2023.109

TABLA 1
Resultados obtenidos de la encuesta interrelación docente-estudiante

Algunas preguntas sobre la encuesta de interrelación docente-estudiante	Si		No	
	Cant.	%	Cant.	%
¿El tutor procura mantener comunicación contigo?	41	62%	25	38%
¿El tutor muestra disponibilidad para atender tus problemas?	51	77%	15	23%
¿El tutor muestra respeto por tus ideas, opiniones y creencias?	56	85%	10	15%
¿El tutor muestra interés y capacidad para escuchar, tratar y proporcionar información para atender los problemas académicos y personales que se te presentan?	55	83%	11	17%
¿El tutor utiliza estrategias que te permiten mejorar tu rendimiento académico?	51	77%	15	23%

4 DISCUSIONES Y CONCLUSIONES

4.1 Discusiones

Según Espinoza en sus conclusiones indica que los autores consultados concuerdan en señalar que las funciones y roles del tutor virtual en los nuevos entornos tecnológicos educativos cobran matices de actuación diferentes a los que se asumen en la tutoría en la modalidad presencial, aunque básicamente su papel es el de acompañar al alumno en su proceso formativo, animándole al aprendizaje y evacuando sus dudas; demanda de novedosas exigencias y competencias comunicativas, organizativas y pedagógicas, así como la de fungir como moderador del proceso [6]. El objetivo del presente estudio fue determinar la relación que existe entre la gestión de la educación remota y competencias de tutor virtual en la Institución Educativa 1279 - Huaycán, Lima 2021[7]

Actualmente la tutoría académica en la UNAMBA es inadecuada y se realiza aun de forma tradicional, poca contribución en el asesoramiento académico de estudiantes en su formación profesional. Además, la información recabada con los informes finales de tutoría se queda en el olvido, nadie realiza el seguimiento a estudiantes especiales con problemas académicos, como por ejemplo, con matrículas condicionadas. Los encargados o comisiones de tutoría de cada escuela no realizan el adecuado seguimiento de las actividades realizadas en las diferentes escuelas, por lo que no se cumplen a cabalidad las actividades tutoriales de parte de los tutores. Para demostrar esta investigación se comparó dos semestres académicos diferentes; la primera hipótesis se demostró que la tutorial virtual contribuye en la formación académica de los estudiantes en la UNAMBA, por lo que sí existe diferencia entre los porcentajes de promedios obtenidos en los semestres académicos 2019-II y 2021-I, en esta investigación se hizo una comparación de ambos semestres académicos, uno cuando las clases se realizaban en forma presencial y en el semestre académico 2021-I las clases continuaban virtualmente; de la misma forma, la segunda hipótesis se comprueba con el porcentaje de notas del contenido actitudinal; la tercera hipótesis se determina con las encuestas realizadas a los estudiantes sobre la tutoría virtual, llegando a la conclusión que efectivamente la tutorial virtual contribuye en la interrelación entre docente y estudiante de la UNAMBA, por lo que es muy importante mantener la relación mutua entre docentes y estudiantes, esta interrelación facilitará la comunicación eficaz entre los actores, así contribuyendo en la formación académica y humana de los futuros profesionales.

A partir de los resultados obtenidos, aceptamos la hipótesis alterna general que establece que la tutorial virtual contribuye significativamente en la formación académica y humana en los estudiantes de la UNAMBA. Por lo que es necesario implantar la tutoría virtual a nivel de la universidad. Por otra parte, en la investigación realizada por Huyta y Bravo, en su estudio, se concluyó que, existe relación positiva, alta y significativa entre la gestión de la educación remota y las competencias del tutor virtual. (Rho Spearman = 0,822 y sig 0,000). Es decir, un incremento en las tareas de gestión con énfasis en la ERDE, están en relación directa con una mejora en las competencias del tutor virtual y viceversa. Asimismo, existe relación positiva, alta y significativa entre la gestión de la educación remota y la competencia pedagógica (Rho Spearman = 0,866, sig 0,000), del mismo modo, existe relación positiva, alta y significativa entre la gestión de la educación remota y la competencia social ética, Rho Spearman 0,714 sig 0,000. Finalmente, existe relación positiva, moderada y significativa entre la gestión de la educación remota y la competencia técnica, Rho Spearman = 0,535 sig 0,000.

En caso de la UNAMBA ambos actores tienen una relación muy favorable que contribuyen en el desarrollo de las actividades académicas. El Desarrollo de la tutoría virtual es fundamental en la UNAMBA, tal como indica la profesora Gilly Salmon "El papel esencial del tutor virtual es promover la interacción humana y la comunicación a través de la transmisión y la construcción de conocimientos y habilidades". De acuerdo a diversos autores, Salmon ha contribuido en gran medida a la comprensión del rol moderador del tutor y sus cualidades y habilidades. En efecto, basándose en una investigación, realizada a lo largo de varios años, la profesora Gilly Salmon, de la Open University, ha establecido un modelo para la moderación de los foros de discusión. El modelo contempla 5 etapas, en las cuales el tutor despliega un conjunto de competencias: acceso y motivación, la socialización, compartir información, construcción de conocimiento y desarrollo. Por lo que se recomienda a la universidad tomar en cuenta el desarrollo de la tutoría virtual como una metodología de apoyo a las clases presenciales y en consecuencia en la formación académica y humana de los estudiantes de la UNAMBA, tal como indica Jorge Mesa [8]. En sus conclusiones Espinoza y Ricaldi resumen, después de un exhaustivo análisis de la literatura especializada los autores coinciden en señalar que los tutores virtuales debe conocer los múltiples beneficios de los EVA, entre otros se mencionan: mejor y mayor acceso a la información, rapidez al acceso e intercambio de información, se rompen con las barreras de espacio y tiempo, interactividad, flexibilidad, educación personalizada, el aprendizaje se individualiza atendiendo al ritmo y estilo de

cada participante y contribuyen al desarrollo formativo del alumno [6].

4.2 Conclusiones

En esta investigación se realizó la comparación de dos semestres académicos, uno cuando las clases se realizaban presencialmente antes de la pandemia y la otra cuando las clases se desarrollaban virtualmente durante la pandemia, en la hipotesis se concluye que sí existe diferencia significativa entre ambos semestres académicos en el porcentaje de promedio de notas y se concluye que la tutoria virtual contribuye en la formación academica y humana de estudiantes de la UNAMBA, y por otro lado según las encuestas realizadas sobre la interrelación docente-estudiante se concluye que efectivamente la tutorial virtual contribuye en la interrelación entre docente y estudiante de la UNAMBA, por lo que es muy importante mantener la relación mutua entre docentes y estudiantes, esta relación facilitará la comunicación eficaz entre los actores, así contribuyendo eficazmente en la formación humana del estudiante tutelado. Según Valderrama Nuñez, se esboza la necesidad de investigar sobre la educación universitaria postpandemia en estudiantes que han vivenciado distintas modalidades (presenciales, virtuales y mixtas), considerando los aspectos pedagógicos, emocionales y relacionales del proceso de enseñanza-aprendizaje, ya que los continuos cambios experimentados impulsa a conocer con mayor profundidad sus implicancias en el alumnado y el profesorado universitario, de tal manera de hacer los ajustes para que la experiencia formativa sea significativa y valorada por todos/as los/as involucrados/as, y así obtener el mayor provecho de cada una de esas modalidades [9]. La reconstrucción y sistematización de las prácticas de tutoría virtual transitó por tres momentos, permitiendo determinar, como aspectos psicopedagógicos esenciales para su desarrollo exitoso, la preparación docente, el diagnóstico, la comunicación asertiva, la coherencia de las influencias y la planificación, ejecución, control y valoración de las acciones tutorials [10].

Agradecimientos

A la Universidad Nacional Micael Bastidas de Apurímac por habernos permitido realizar esta investigación.

Referencias

[1] A. Toledo Chambilla, "LA TUTORÍA A TRAVÉS DE LAS AULAS VIRTUALES: UNA EXPERIENCIA DE FORMACIÓN Y AUTOFORMACIÓN.", Bolivia, sep. 2019. Consultado: el 5 de septiembre de 2023. [En línea]. Disponible en: https://dialnet.unirioja.es/servlet/articulo?codigo=9001484

[2] L. Balseca Córdova, V. Orellana Navarrete, I. C. Rodríguez Ordóñez, y D. A. Salas Bustos, "Formación de tutoras y tutores virtuales. Una experiencia educativa desde el punto de vista de los capacitadores", Revista Andina de Educación, vol. 4, núm. 2, pp. 1–10, may 2021, https://doi.org/10.32719/26312816.2021.4.2.1

[3] J. Luis Cueva Delgado Liliana de la Caridad Molerio Rosa Misleydi Ramírez Serrano, J. Luis Cueva Delgado, L. de la Caridad Molerio Rosa, y M. Ramírez Serrano, "Las tutorías virtuales en el proceso universitario Las tutorías virtuales en el proceso universitario Virtual tutorials in the university process", 2019. [En línea]. Disponible en: https://orcid.org/0000-0001-8760-6854

[4] D. A. Contreras Navarro y Bejarano Álvarez Patricia Mónica, "La tutoría virtual en las clases remota de la estrategia aprendo en casa", Ciencia Latina Revista Científica Multidisciplinar, vol. 6, núm. 2, pp. 1748–1757, abr. 2022, https://doi.org/10.37811/cl_rcm.v6i2.1993

[5] R. Hernández Sampieri, C. Fernández Collado, y P. Baptista Lucio, Metodología de la Investigación, 6ta ed. México, 2014.

[6] E. E. Espinoza Freire y M. L. Ricaldi Echevarria, "El tutor en los ambientes virtuales de aprendizaje", Universidad y Sociedad, pp. 201–210, 2018.

[7] E. Huayta-Campos y G. Bravo-Huaynates, "Educación remota y competencias del tutor virtual en la pandemia COVID-19 - Lima", 593 Digital Publisher CEIT, vol. 7, núm. 4, pp. 167–187, jul. 2022, https://doi.org/10.33386/593dp.2022.4.1095

[8] J. Mesa Vázquez, M. Parra Álvarez, y G. E. C. Marcillo, "VIRTUAL TUTORING ENVIRONMENT: AN ALTERNATIVE TO STREAMLINE SCIENTIFIC RESEARCH PROCESSES", 2021.

[9] C. M. Valderrama Nuñez, K. Roa Hernández, C. Landeros Díaz, J. Reyes Macias, y M. Quilabrán Meneses, "Tutorías virtuales postpandemia: implicancias en los aprendizajes de estudiantes de terapia ocupacional que cursan sus prácticas profesionales", Revisión Dic, vol. 12, 2023, https://doi.org/10.1590/2526-8910.ctoao265234603

[10] A. Rodríguez Figueroa, M. A. Pérez Egües, y L. del C. Torres Zerquera, "La tutoría virtual en la carrera Licenciatura en Educación. Pedagogía Psicología, experiencias en la Universidad de Cienfuegos", pp. 1–19, dic. 2022, Consultado: el 3 de septiembre de 2023. [En línea]. Disponible en: https://revistamapa.org/index.php/es/article/view/342/507

Biografías

Francisco Cari Incahuanaco, Maestro en Investigación y Docencia Universitaria, Ingeniero Estadístico e Informático con Segunda Especialidad en Ingeniería de Sistemas. Experiencia en docencia universitaria, actual Docente Asociado a Tiempo Completo en la Universidad Nacional Micaela Bastidas de Apurímac, desarrollando las asignaturas de Estructura de Datos y Metodología de Investigación Científica.

Yhon Fuentes Huamán, Magíster Scientiae en informática con mención en: gerencia de tecnologías de información y Comunicaciones, docente en la Universidad Nacional Micaela Bastidas de Apurímac, fue Director del Departamento de Ingeniería de la UTEA y Docente de la Escuela de Ingeniería Ambiental y RR.NN. Desarrollador de proyectos de software informáticos en empresas privadas y públicas, Becado por FONDECYT en la Universidad Nacional de San Agustín de Arequipa en el Doctorado de Ciencias de la Computación.

Maritza Condori Quispe, Lic. En Educación con estudios de posgrado en docencia universitaria, docente en la Universidad Nacional Micael Bastidas de Apurímac y docente de educación básica regular en área de ciencia y tecnología.

C&T Riqchary
Revista de Investigación en Ciencia y Tecnología
ISSN: 2810-8124 (en línea) / ISSN: 2706-543x

Vol. 5 Núm. 1 (2023)
Publicado: 15/10/23
DOI: https://doi.org/10.57166/riqchary/v5.n1.2023.1010

Metodología para el diagnóstico y la gestión de mantenimiento en las subestaciones eléctrica de los grupos de generación distribuida

Methodology for the diagnosis and maintenance management in the electrical substations of the distributed generation groups

Carlos Manuel Ruano González

ORCID: 0000-0002-9647-8759

(Recepción: 20/08/2023 y aceptación 03/10/2023)

Resumen— La generación distribuida juega un papel importante en la disponibilidad en los Sistema Eléctrico de Potencias. Por su naturaleza, cambia la concepción de las técnicas de mantenimiento respecto a las tradicionalmente usadas en las subestaciones convencionales, por lo que la gestión para su realización debe estar bien concebida y fundamentada para lograr su eficiencia operacional. Partiendo de esta consideración, se dirige el análisis hacia el diagnóstico y mantenimiento de las celdas de la subestación del grupo de generación con fuel oíl HYUNDAI de 40 MW ubicado en la cuidad de Camaguey Cuba. Actualmente no se cuenta con un proceder para realizar esta actividad, y se emplean procedimientos aplicados a las subestaciones de exteriores. Después de estudios realizados ante las frecuentes averías que han comprometido el servicio y provocados daños a los equipos que componen las celdas, se propone una metodología para el diagnóstico y la gestión del mantenimiento a este tipo de subestaciones, lográndose reducir las averías y reducir los costos por concepto de mantenimiento.

Palabra clave: celdas, metodología, mantenimiento, subestación

Abstract— The distributed generation has great importance for the availability of the National Electrical of Power. For its nature, must change the conception of the maintenance procedures in relation to the traditionally used in the conventional substations, which is why the step for its realization should be well conceived and well justified. Since this consideration, the analysis is directed toward the diagnosis and maintenance of the cells of the substation of the group of generation with fuel oil HYUNDAI of 40 MW. At the present, there is not a procedure to accomplish this activity, and actions applied to the substations of exteriors are used. After of studies developed in front of frequent damages that have compromised the service and provoked damages to the equipment's of the cells, a methodology for the diagnosis and the maintenance management to this type of substations is proposed.

Keywords: cells, methodology, maintenance, substation

Carlos M. Ruano_González Universidad de Camagüey – Cuba, cmruanog@gmail.com

1 INTRODUCCIÓN

La generación distribuida juega un papel importante como complemento de la generación base y mientras no se modifique la "Matriz energética", dando paso a la generación con fuentes renovables de energías, se hace necesario mantenerla en buen estado técnico.

Las subestaciones eléctricas de interiores, asociadas a la generación distribuida, independientemente que de forma conceptual cumplen el mismo objetivo que las subestaciones eléctricas de exteriores, tienen sus particularidades en su diseño que hace que la concepción de su diagnóstico y mantenimiento, deban tener otras consideraciones respecto a las subestaciones de exteriores tradicionalmente empleadas. Por esto es de suma importancia contar una metodología para el diagnóstico y la gestión del mantenimiento, adecuada para evitar las frecuentes averías que comprometen el servicio de energía eléctrica a los usuarios y precaver daños a los equipos primarios que componen las celdas eléctricas.

El mantenimiento según la IEEE no es más que el acto de preservar o mantener las condiciones de un equipo para su correcta operación, o sea, la combinación de todas las acciones técnicas y administrativas que incluyen acciones de supervisión, que puedan mantener o restaurar el estado para el cual cumple con las funciones afines a su desempeño un equipo o sistema determinado (IEEE Std 3007.2, 2010).

El objetivo de la gestión del mantenimiento a las subestaciones eléctricas de media tensión de interiores, es asegurar la máxima confiabilidad y disponibilidad de las celdas a bajos costos para que cumplan con sus funciones operativas dentro de un contexto operacional, previniendo o corrigiendo, cuando sea necesario, condiciones que pongan en riesgo la operación de las mismas y de la instalación a la que están asociadas, así como a los equipos e instalaciones vecinas.

Varios autores generalmente dividen el mantenimiento en dos categorías principales, el Mantenimiento no Planificado (correctivo, reparación de averías) y el Mantenimiento Planificado (preventivo, predictivo, proactivo). Sin embargo, otros autores prestigiosos clasifican al mantenimiento en correctivo, preventivo y predictivo. También se puede ver en más detalles otras acepciones de las filosofías del mantenimiento enfocado a la confiabilidad.

Por otro lado, de acuerdo a investigaciones más recientes

realizada por grupos de trabajo de Grandes Redes Eléctricas por su sigla en inglés (CIGRE), se plantea que dentro de los esquemas de mantenimientos más utilizados están el mantenimiento preventivo en un 55%, este generalmente consiste en reacondicionar o sustituir a intervalos regulares un equipo o sus componentes, independientemente de su estado en ese momento. Esos intervalos son establecidos por los fabricantes basados en experiencias anteriores, sin embargo, las investigaciones demuestran que la efectividad de estos métodos están fuertemente vinculados a la frecuencia de las inspecciones de mantenimiento, y en el caso de las celdas eléctricas muchas de las fallas que ocurren no se detectan simplemente por ello es necesario usar otras técnicas no invasivas que ayudan a determinar su estado), el mantenimiento basado en la condición representa el 8% y un 37% la combinación de ambos.

El diseño de estudio se desarrolló bajo un enfoque cualitativo descriptivo utilizando la metodología de estudio de caso en un grupo de Generación Distribuida con fuel oíl HYUNDAI de 40 MW, el mismo está compuesto por cuatro baterías conformadas con cuatro motores del tipo HSR77198P de 2,5 MW cada uno, cada batería tiene una generación de 10 MW para un total de 40 MW.

La subestación eléctrica de salida del grupo o Shelter, como es conocida pertenece al grupo de generación distribuida ubicado en el municipio de Nuevitas, provincia de Camaguey Cuba, la conforman 11 celdas eléctricas con la siguiente distribución:

- Cuatro transformadores tipo KMO – 0542 de salida de las baterías de 12,5 MVA y 34.5/6,6 kV.
- Cuatro líneas eléctricas de salida, una de enlace de barra y dos salidas de conexión con una subestación de 110 kV de salida de planta de una Central Termoeléctrica.
- Cada celda de alta tensión la componen, generalmente, un interruptor de potencia, un desconectivo de barra, un desconectivo de tierra, un juego de pararrayos, transformadores de potencial, un transformador de corriente por fase y los conductores, todos a un nivel de tensión a 34,5 kV.

Esta subestación es atendida por la Empresa de Construcciones de la Industria Eléctrica (ECIE), siendo responsabilidad de esta entidad; la realización de los mantenimientos, la atención a averías y daños a los equipos primarios, así como la sustitución de algunos en casos en que sea necesario.

Independientemente de ser subestaciones eléctricas de interiores y encontrarse en un espacio reducido de cabinas metálicas y climatizadas, siempre existe penetración de partículas contaminantes como polvos en suspensión y hollín que, ante la condensación de la humedad, provocan, una condición anormal de funcionamiento, llegando a descargas de tensión entre las partes energizadas o activas de los equipos y hasta a severas fallas por contaminación con un apreciable índice de falla. Por lo antes expuesto, se ha prestado una especial atención a las técnicas de diagnóstico y mantenimiento de esta subestación. A medida que las tecnologías avanzaban también se necesitaba un nuevo enfoque en el mantenimiento, por ello los especialistas crearon lo que se conoce ahora como el Mantenimiento Predictivo o Basado en la Condición, el mismo consiste en inspeccionar los equipos a intervalos regulares y tomar acción para evitar las fallas o las consecuencias de las mismas según condición. Incluye tanto las inspecciones objetivas (con instrumentos) y subjetivas (con los sentidos), como la reparación del defecto (falla potencial), estas técnicas del mantenimiento han tenido un cierto avance en algunos equipos primarios de subestaciones como son los transformadores de potencias, pero no el resto de los elementos que componen los campos en las subestaciones de exteriores y las celdas eléctricas de las subestaciones de interiores.

El estudio del arte en la gestión de los mantenimientos a las subestaciones eléctricas de media tensión de interiores, se refiere a la aplicación de técnicas y herramientas avanzadas en el ámbito de la ingeniería eléctrica para mejorar la eficiencia, calidad y seguridad de las operaciones de mantenimiento. Se observa que en ninguno de los estudios, se ha realizado un análisis profundo de la gestión del mantenimiento para evitar de alguna forma las diferentes fallas y averías registradas en las celdas eléctricas, por lo que a pesar del aumento gradual de las diferentes investigaciones científicas y tesis relacionadas con el diagnóstico y monitoreo de las condiciones técnicas de los equipos primarios de subestaciones, se denota un vacío en los estudios realizados en el mundo referente a la gestión de mantenimiento de las subestaciones de media tensión de interiores.

2 METODOLOGÍA

2.1 Métodos de investigación

- **Métodos teóricos**

Histórico: Para conocer los métodos más comunes desarrollados en la realización de la gestión de los mantenimientos y diagnósticos, desde su surgimiento, evolución y desarrollo.

Análisis-Síntesis: Para analizar los componentes primarios de las celdas eléctricas de modo integral y comprender su estructura. En la síntesis deben unirse armónicamente las partes previamente analizadas, a fin de revelar sus relaciones y características generales.

Sistémico: Para establecer la relación entre los factores que intervienen en el desgaste o envejecimiento de los equipos primarios.

- **Método estadístico**

Para la obtención de los patrones de comportamiento de las variables de estado para validarlo.

2.1.1 Monitoreo

El monitoreo y diagnóstico de los equipos de subestaciones han atraído gran atención desde hace mucho tiempo, este interés se ha disparado en los últimos años debido a los cambios estructurales en los negocios de la electricidad. En el competitivo mercado actual, los productores deben seguir las estrategias que coincidan en lo más posible con los intereses de los consumidores, ya que el desarrollo en el monitoreo y diagnóstico debe estar directamente relacionado con las necesidades de las empresas eléctricas y para reducir

C&T Riqchary
Revista de Investigación en Ciencia y Tecnología
ISSN: 2810-8124 (en línea) / ISSN: 2706-543x

Vol. 5 Núm. 1 (2023)
Publicado: 15/10/23
DOI: https://doi.org/10.57166/riqchary/v5.n1.2023.1010

los costos operacionales. El monitoreo y el diagnóstico se aprecia como una posible vía de optimización de los activos existentes [1], con vistas a reducir los costos de mantenimiento, prevenir las salidas por emergencia, y lograr que los equipos trabajen un plazo mayor y de una forma más segura [2].

El monitoreo es el proceso continuo y sistemático mediante el cual se verifica la eficiencia y la eficacia de un producto mediante la identificación de sus logros y debilidades y en consecuencia se recomiendan medidas correctivas para optimizar los resultados esperados [3]. Es, por tanto, condición para la rectificación o profundización de la ejecución y para asegurar la retroalimentación entre los objetivos y presupuestos teóricos y las lecciones aprendidas a partir de la práctica. Asimismo, es el responsable de preparar y aportar la información que hace posible sistematizar resultados y procesos, por tanto, es un insumo básico para el diagnóstico. Para que el monitoreo sea exitoso requiere del establecimiento de un sistema de información, identificando a los usuarios de la información, los tipos de información prioritaria, vinculando las necesidades y las fuentes de información, estableciendo métodos apropiados para efectuar la recopilación de datos e identificar los recursos necesarios [4].

2.1.2 Diagnóstico

Por otro lado, el diagnóstico es el proceso integral y continuo de investigación y análisis de los cambios más o menos permanentes que se materializan en mediano y largo plazo, como una consecuencia directa o indirecta del quehacer de un equipo en el contexto, la población y las organizaciones participantes. Por ello, constituye una herramienta para la transformación, que arroja luz sobre las alternativas para la mejora permanente de las intervenciones presentes y futuras, o sea, transfiere buenas prácticas [5].
Desde esta concepción, el monitoreo y el diagnóstico tienen que ser coherentes con su objeto de análisis.
El monitoreo es la medición de uno o varios parámetros previamente establecidos que se define como la recolección de datos, e incluye el desarrollo de sensores, técnicas de medición y hardware para la adquisición y almacenamiento de los mismos. El diagnóstico es la ciencia encargada de determinar por los síntomas, el estado y el carácter de una posible rotura, envejecimiento, etc. de un equipo, sistema o proceso a partir del análisis y procesamiento de sus datos de pruebas, de fallas e históricos [6]. El diagnóstico tiene como objetivo detectar las desviaciones de los parámetros de control, analizar sus tendencias, determinar sus causas de manera que permita al personal proponer las medidas correctoras que permitan la intervención oportuna del equipo para lograr máxima disponibilidad y confiabilidad de los mismos a través del mantenimiento [7].

2.2 Generalidades del Mantenimiento

El mantenimiento según la IEEE no es más que el acto de preservar o mantener las condiciones de un equipo para su correcta operación, o sea, la combinación de todas las acciones técnicas y administrativas que incluyen acciones de supervisión, que puedan mantener o restaurar el estado para el cual cumple con las funciones afines a su desempeño un equipo o sistema determinado (IEEE Std 3007.2, 2010).

Varios autores generalmente dividen el mantenimiento en dos categorías principales, el Mantenimiento no Planificado (correctivo, reparación de averías) y el Mantenimiento Planificado (preventivo, predictivo, proactivo). Sin embargo, otros autores prestigiosos clasifican al mantenimiento en correctivo, preventivo y predictivo.

Por otro lado, de acuerdo a investigaciones más recientes, se plantea que dentro de los esquemas de mantenimientos más utilizados están el mantenimiento preventivo en un 55%, este generalmente consiste en reacondicionar o sustituir a intervalos regulares un equipo o sus componentes, independientemente de su estado en ese momento. Esos intervalos son establecidos por los fabricantes basados en experiencias anteriores, sin embargo, las investigaciones demuestran que la efectividad de estos métodos están fuertemente vinculados a la frecuencia de las inspecciones de mantenimiento, y en el caso de las celdas eléctricas muchas de las fallas que ocurren no se detectan simplemente por ello es necesario usar otras técnicas no invasivas que ayudan a determinar su estado), el mantenimiento basado en la condición representa el 8% y un 37% la combinación de ambos.

Se aplican de forma general los siguientes tipos de mantenimiento:
- Preventivo
- Predictivo
- Correctivo

Para cada uno de ellos, se describen a continuación sus principales características y definiciones:

a) Mantenimiento preventivo: Es un conjunto de actividades (inspecciones, pruebas, ajustes, reparaciones, toma de muestras, etc.) regularmente programadas y aplicadas a los equipos de una instalación determinada, con el objeto de minimizar su "degradación" o pérdida de vida útil.
b) Mantenimiento predictivo: Es una técnica que consiste en monitorear regularmente (más seguido que en el mantenimiento preventivo) los parámetros "claves" de un equipo en operación, con la finalidad de detectar y/o corregir a tiempo un problema potencial antes de que se produzca la falla del equipo.
c) Mantenimiento correctivo: Cuando se detecta que el desempeño de algún equipo comienza a disminuir y se presume las posibles causas de este bajo desempeño, se debe atacar a dichas causas y realizar un mantenimiento del equipo para evitar que el problema se agrave. El mantenimiento realizado en este tipo de circunstancias se lo conoce como mantenimiento correctivo.

Estos conceptos han ido evolucionando y desarrollando y en cualquier caso se pueden distinguir cuatro generaciones en la evolución del concepto de mantenimiento:

a) 1ª Generación: La más larga, desde la revolución industrial hasta después de la 2ª Guerra Mundial, aunque todavía

impera en muchas industrias. El Mantenimiento se ocupa sólo de arreglar las averías. Es el Mantenimiento correctivo.

b) 2ª Generación: Entre la 2ª Guerra Mundial y finales de los años 70 se descubre la relación entre edad de los equipos y la probabilidad de fallo. Se comienza a hacer sustituciones preventivas. Es el Mantenimiento preventivo.

c) 3ª Generación: Surge a principios de los años 80. Se empieza a realizar estudios CAUSA-EFECTO para averiguar el origen de los problemas. Es el mantenimiento predictivo o detección precoz de síntomas incipientes para actuar antes de que las consecuencias sean inadmisibles. Se comienza a hacer partícipe a producción en las tareas de detección de fallos.

d) 4ª Generación: Aparece en los primeros años 90. El mantenimiento se contempla como una parte del concepto de "calidad total". Se entiende que mediante una adecuada gestión del mantenimiento es posible aumentar la disponibilidad al tiempo que se reducen los costos. Es el Mantenimiento Basado en el Riesgo (MBR). Se concibe el mantenimiento como un proceso de la empresa al que contribuyen también otros departamentos. Se identifica el mantenimiento como fuente de beneficios, frente al antiguo concepto de mantenimiento como "mal necesario". La posibilidad de que una máquina falle y las consecuencias asociadas para la empresa es un riesgo que hay que gestionar, teniendo como objetivo la disponibilidad necesaria en cada caso al mínimo coste.

Sobre la base de los resultados obtenidos de pruebas realizadas al equipo eléctrico, el personal responsable del mantenimiento, tiene los argumentos suficientes para para tomar la decisión de mantener energizado o retirar de servicio un equipo en operación que requiera mantenimiento, para el mantenimiento del equipo, es conveniente considerar los aspectos siguientes:

a) Archivo histórico y análisis de resultados obtenidos en inspecciones y pruebas. Es necesario además considerar las condiciones operativas de los equipos, así como las recomendaciones de los fabricantes.
b) Establecer las necesidades de mantenimiento para cada equipo.
c) Formular las actividades de los programas de mantenimiento.
d) Determinar actividades con prioridad de mantenimiento para cada equipo en particular.
e) Se debe contar con personal especializado y competente para realizar las actividades de mantenimiento al equipo y establecer métodos para su control.

Se requiere un cambio de mentalidad en las personas y se utilizan herramientas como:

a) Ingeniería del riesgo (determinar consecuencias de fallos que pueden ser aceptables o no).
b) Análisis de fiabilidad (identificar tareas preventivas factibles y rentables).
c) Mejora de la mantenibilidad (reducir tiempos y costes de mantenimiento).

Control y evaluación del mantenimiento a subestaciones eléctricas mediante indicadores de clase mundial

Para conocer cuán eficiente es la aplicación de la política de mantenimiento planificada, se necesita controlar y evaluar la gestión del mantenimiento, esto permite actuar de forma rápida y precisa sobre los factores débiles en el mantenimiento. Deben realizarse informes concisos y específicos, formados por tablas de indicadores y sus respectivos gráficos para algunos de estos. Estos informes deben permitir un análisis fácil y adecuado a cada nivel de gestión, sobre la base de "indicadores de clase mundial", llamados así por ser utilizados y obtenidos en casi todos los países de la misma forma, de los seis indicadores de clase mundial, cuatro se refieren al análisis de la gestión de equipos y dos a la gestión de costos. En este trabajo solo se tuvo en cuenta tres indicadores del primer grupo, el indicador "tiempo medio para la falla" no se analiza, por referirse a equipos reparados. Los tres indicadores analizados son:

a) **Tiempo medio para reparaciones (TMPR):** relación entre el tiempo total de intervención correctiva en un conjunto de elementos con falla y el número total de fallas detectadas en esos elementos, en el período observado. Tiene la siguiente expresión:

$$TMPR = \frac{\sum HTMN}{NTMC}\ [h] \qquad (1)$$

Donde:
HTMN: tiempo total de mantenimiento (h).
NTMC: número total de fallas detectadas en el período observado.

b) **Disponibilidad (DISP):** porcentaje del tiempo total que el equipo estuvo a disposición del órgano de operación para desempeñar su actividad.

$$DISP = \frac{HROP}{HROP + HTMN} \times 100 \qquad (2)$$

c) **Tiempo promedio entre fallos (TMEF):** indica el intervalo de tiempo más probable entre un arranque y la aparición de un fallo; es decir, es el tiempo medio transcurrido hasta la llegada del evento "fallo". Mientras mayor sea su valor, mayor es la confiabilidad del componente o equipo. Tiene la siguiente expresión:

$$TMEF = \frac{NOIT \times HROP}{NTMC}\ [h] \qquad (3)$$

Donde:
NOIT: total de elementos analizados.
HROP: tiempo total de operación (h).

$$TMEF = \frac{7 \times 2160}{2} = 7560\ [h]$$

Sobre esta línea de trabajo se dirige el análisis a las celdas que conforman la subestación del grupo de generación distribuida. Teniendo en cuenta los indicadores mencionados anteriormente y las bases de datos generadas de la gestión del

C&T Riqchary
Revista de Investigación en Ciencia y Tecnología
ISSN: 2810-8124 (en línea) / ISSN: 2706-543x

Vol. 5 Núm. 1 (2023)
Publicado: 15/10/23
DOI: https://doi.org/10.57166/riqchary/v5.n1.2023.1010

mantenimiento aplicada a la subestación de referencia, se obtiene que el tiempo promedio entre fallos tomando como ejemplo una celda que la conforman siete elementos primarios tiene como promedio dos fallas en el trimestre, es:

2.2.1 Metodología existente

La metodología existente son varias, ya que diferentes empresas están encargadas de mantener y atender las averías en los grupos de Generación Distribuida, que, al no existir un procedimiento específico para este tipo de subestación eléctrica, se aplica el procedimiento para subestaciones de exteriores, ejecutándose las actividades de mantenimiento en los tiempos programados en esta normativa [2].
Según este documento, el mantenimiento a las subestaciones eléctricas de exteriores se ejecuta una vez al año a los equipos primarios, así como acciones de diagnóstico con termografía infrarroja una vez a la semana. La medición termográfica no se realiza en las subestaciones de grupos electrógenos por lo que solo se ejecuta su mantenimiento anualmente sin acciones de diagnóstico. En la tabla 1 se aprecia un ejemplo de la programación de las celdas.

TABLA 1
Programación anual de los mantenimientos por celdas de acuerdo al modelo existente.

No. Celdas	Equipos	Año y mes de ejecución del mantenimiento		
		2020	2021	2022
1	Interruptor	Junio	Junio	Junio
	Desc. de barra	Junio	Junio	Junio
	Desc. a tierra	Junio	Junio	Junio
	TC	Junio	Junio	Junio
	Pararrayos	Junio	Junio	Junio
	Conductores	Junio	Junio	Junio

3 RESULTADO Y DISCUSIÓN

Propuesta de metodología para el mantenimiento a subestaciones eléctricas de 34,5 kV utilizadas en los grupos de generación distribuida.
Las instalaciones de los grupos electrógenos para la generación distribuida, aun en estos momentos juegan un rol fundamental como complemento de la generación base de los Sistema Eléctricos de potencias, hasta que no se logre modificar la matriz energética en lo concerniente a la generación de electricidad con fuentes de energía renovables y teniendo en cuenta la metodología de los mantenimientos a las subestaciones eléctricas de estas instalaciones no tienen con un proceder y se aplican los estándares establecidos en el procedimiento de las subestaciones eléctricas de exteriores y al no existir ninguna actividad de diagnóstico [3], este tipo de subestación eléctrica presentan una situación específica por su configuración, que son conformadas por celdas dentro de un local climatizado a una temperatura de trabajo entre 21 y 23 °C, la cual no siempre se cumple y no existe una total hermeticidad, que provoca la acumulación de sólidos en suspensión y hollín que se depositan en las superficies provocando conducción entre las partes activas de los elementos y tierra.

Existiendo saltos de corrientes indeseados que conllevan a fallas o salidas fuera de servicio de las celdas eléctricas, por lo que se necesita disminuir la frecuencia de verificación de sus parámetros ósea que el tiempo entre un mantenimiento o control de sus parámetros puede variar, donde se busca una relación de tiempo mínimo para realizar la intervención antes que ocurra la avería [4], esto apoyado con un sistema de diagnóstico termografico una vez a la semana como se realiza en las subestaciones eléctricas de exteriores que permite el monitoreo de la condición.

Esta propuesta está basada en la experiencia de la explotación, mantenimientos y atención de las averías en una subestación eléctrica de un grupo electrógeno de 40 MW de la Generación Distribuida, que de acuerdo a los estudios y análisis realizados debidos a las frecuentes fallas presentadas en las celdas que componen la subestación eléctrica conllevan a implementar una técnica de diagnóstico de termografía infrarroja que es posible ejecutar, ya que se cuenta con la tecnología y la realización de mantenimientos en un menor plazo de tiempo.

Los indicadores de gestión del mantenimiento una vez aplicada la nueva metodología, se puede apreciar que alcanza una mayor confiabilidad al disminuir los eventos de fallos o detectar precozmente su ocurrencia, la cual permite una mejor planificación del mantenimiento. En este caso los eventos de fallo en el periodo observado, ósea en un trimestre no ocurren fallos, si no su periodicidad es de una falla en el semestre, obteniendo de la ecuación de tiempo promedio entre fallos referenciados a tres meses tenemos un valor de:

$$TMEF = 30240\ [h]$$

Realizando una comparación con el tiempo promedio de la anterior metodología vemos una mejora sustancial. El tiempo promedio entre fallo en los años anteriores Tabla 2 (elaboración propia). Con la anterior metodología era de aproximadamente de $7560\,[h]$ y ahora con la nueva metodología es de $30240\,[h]$, valor muy superior [5].

TABLA 2
Comparación del tiempo promedio de fallo, en los últimos tres años

Años	2022	2021	2020
Tiempo promedio de fallo	7555	7855	7560

4 CONCLUCIONES

En el análisis que se realizó en la subestación eléctrica del grupo electrógeno por las continuas averías producidas por saltos de corrientes en los elementos que componen las celdas eléctricas, entre la parte activas de los equipos y tierra que provocan salidas de servicio, con sus inconvenientes como son falta de servicio al cliente, reparación de elementos dañados que trae consigo gastos de recursos materiales y de capital humano y al no existir un procedimiento o metodología para la realización de los mantenimientos, se aplica el utilizado en las subestaciones de alta tensión de exteriores, pero el diseño de estas subestaciones conformadas por celdas dentro de un local climatizado tiene sus especificidades, además de no realizarse ningún tipo de diagnóstico, es necesario establecer una metodología para su concepción,

Por lo que es necesario utilizar como método de diagnóstico la termografía infrarroja una vez por semana, para la cual se cuenta con los medios técnicos necesarios como la cámara infrarroja InfRec modelo R300R-D y de esta forma determinar tempranamente los posibles saltos de corrientes entre la parte activa y tierra.

Se demuestra que la metodologías utilizadas en la gestión de los mantenimientos para las subestaciones de interiores utilizadas en los grupos electrógeno de generación distribuida y en específico las normativas para realizar el mantenimientos no se adecuan para este tipo de subestación, por lo que se hace necesario implantar una metodología, que aunque no tenga en cuenta el monitoreo de la condición de la subestación o de las celdas en específico, ya que no se cuenta con la tecnología necesaria, se pueda identificar a tiempo la posible falla, realizando un eficiente diagnóstico termográfico e interviniendo oportunamente los elementos que la componen.

5 REFERENCIAS

[1] Casaña-Medel, J. C., Macías-Socarras, I., y Morales-Tamayo, Y. "El mantenimiento a partir de los indicadores de clases mundial en la fábrica Lácteos Bayamo" Revista Ciencias Técnicas Agropecuarias. 2021, vol. 30, n. 3, p. 115–132 ISSN 2071 – 0054. [Consultado el 25 de enero de 2022]. Disponible en: http://scielo.sld.cu

[2] PEREA LOZANO, Brenda Yasnier; LOPEZ SUAREZ, Henry Norholey. "Implementación de mantenimiento preventivo y predictivo en la empresa EQUIACERO SA" Revista UNILIBRE 2020, vol. 8, n. 3, p. 85–112

[3] MONTANÉ – GARCÍA, Jorge Juan; DORRBERCKER – DRAKE, Santiago Alfredo; HERNÁNDEZ – AREU, Orestes. "El mantenimiento a los transformadores de potencia; su análisis en el caso de una central termoeléctrica cubana". Ingeniería Energética Vol. XXXII, 2/2011 Abril – Julio p- 56- 64 ISSN 1815 – 5901.

[4] RUANO GONZÁLEZ, Carlos. "Metodología para el mantenimiento de interruptores empleados en las maniobras de reactores y bancos de reactores". XIX Simposio de Ingeniería Eléctrica (SIE-2021). III Convención Científica Internacional 2021. Universidad Central de las Villas "Martha Abreu" Santa Clara Noviembre 2021.

[5] Revista Mantenimiento Mundial. "Indicadores de Clase Mundial", vol. 3, núm 15[en línea]. [ref. de 16. de agosto 2019]. Disponible en Web: http://www.mantenimientomundial.com.

[6] VÉLEZ- QUIROZ, Alcira Magdalena; FILGUEIRAS- SAINZ DE ROZAS, Miriam Lourdes; VILARAGUT- LLANES, Miriam; Revista Ingeniería Energética. 2022. Vol. 43, núm. 3, septiembre/diciembre. ISSN: 1815-5901.

[7] ABB. "Celdas Metal Clad-Metal Enclosed". 2020.

[8] ROQUE-LOPEZ. Alexander. "Transformadores Eléctricos". Universidad Nacional Autonoma De Honduras. Tegucigalpa 2019.

[9] ZAPATA de la CRUZ, Luis. "Filosofía y generación de valores a través de la gestión de mantenimiento", 20 Congreso Peruano Ingeniería de Mantenimiento, Lima. 2020.

[10] ITESA, soluciones de energía. "Celdas de Media Tensión: ¿Qué tipos hay?", [en línea] 2022. Disponible en Web: http://www.itesa.com.pe

[11] Sánchez Durán, R. "El futuro y el mantenimiento energético". Universidad de Sevilla (España). Tesis Doctoral. https://dialnet.unirioja.es/servlet/tesis?codigo=283233, PP.234. 2020

[12] Moreno Muñoz. Alberto, Máster en Ingeniería del mantenimiento (MUEA), " Implementación del mantenimiento TPM, Técnicas Lean y 4.0", Mayo 2022.

[13] DEUS AGUILERA. CARLOS, CABRERA GÓMEZ. JESÚS, Una mirada a la evolución natural del mantenimiento, Nota técnica 382, CubaMan, Diciembre de 2022.

[14] [14] Fujitsu, Observatorio de Industria 4.0. España, El Mantenimiento Predictivo en la Industria 4.0, Club Excelencia en Gestión: www.clubexcelencia.org. 2022

[15] M. en C. Luna Pérez. Miguel Ángel, Doctorado en Ingeniería de Sistemas, Instituto Politécnico Nacional, SEPI ESIME Zacatenco, Diseño de una metodología de mantenimiento predictivo para asegurar procesos de producción de la industria 4.0, South Florida Journal of Development, Miami, v.2, n.1, p. 1009-1017, jan./feb. 2021. ISSN 2675-5459.

6 BIBLIOGRAFÍA

Carlos Manuel Ruano González Ingeniero Electricista desde 1985 en la Universidad de Camagüey. Cuba. Desde el año 1990 trabaja en el Sistema Eléctrico de Potencia en las subestaciones de media y alta tensión, ha participado en varios eventos internacionales de la especialidad, es master en ciencia y actualmente está matriculado en un programa doctoral, trabaja en la actualidad como especialista de campo, tiene vasta experiencia en la operación y control de los sistemas eléctricos de potencia, así como en la operación y mantenimiento de subestaciones eléctricas y líneas de transmisión eléctricas.

C&T Riqchary
Revista de Investigación en Ciencia y Tecnología
ISSN: 2810-8124 (en línea) / ISSN: 2706-543x

Vol. 5 Núm. 1 (2023)
Publicado: 15/11/23
DOI: https://doi.org/10.57166/riqchary/v5.n1.2023.111

Método de Box-Jenkins para el pronóstico de consumo de energía eléctrica del 2021-2023

Box-Jenkins method for electricity consumption forecast for 2021-2023

Dennis Uriel Añasco-Chata[A], Percy Huata-Panca[B], Ecler Mamani-Vilca[C], Adolfo Carlos Jimenez-Chura[D], y Pablo Cesar Tapia-Catacora[E]

ORCID 0009-0006-7254-5837[A], ORCID 0000-0002-1624-5526[B], ORCID 0000-0002-5205-3660[E], ORCID 0000-0003-3125-8197[D] y ORCID 0000-0002-8608-4203[E]

(Recepción: 23/09/2023 y aceptación 06/10/2023)

Resumen— El trabajo fue determinar un modelo de serie de tiempo basado en el método de Box-Jenkins, que se ajuste apropiadamente a la serie original en el pronóstico del consumo de energía eléctrica de la región de Puno – Perú, para los próximos tres años (2021-2023), esto permitiría a la empresa Electro Puno contar con el modelo que respalde la toma de decisiones , que se ajusten a sus objetivos corporativos en un horizonte de tiempo a mediano plazo, garantizando la distribución y la comercialización de energía eléctrica. El estudio fue cuantitativo con diseño no experimental, de corte longitudinal de tendencia (trend), la población fue, 84 reportes de consumo mensual de energía eléctrica representados en (MWh/mes), considerando un periodo comercial de 7 años, se aplicó las tres fases de la metodología Box-Jenkins como: identificación, estimación y validación del modelo, llegando a determinar un modelo multiplicativo estacional ARIMA (0,1,3)×(2,0,0)_12 que se ajustó perfectamente a la serie en estudio, siendo de “alta precisión” con un valor MAPE del 2.79% , finalmente en la última fase de la metodología se realizó los pronósticos mensuales de consumo de energía eléctrica, con una desviación en promedio por cada pronóstico de un valor MAE =625.80 MWh.

Palabras Clave: Box-Jenkins, Pronóstico, Energía eléctrica

Abstract— The work was to determine a time series model based on the Box-Jenkins method, which is appropriately adjusted to the original series in the forecast of electricity consumption in the region of Puno Peru, for the next three years (2021-2023), this would allow the company Electro Puno to have the model to support decision making, which conform to its corporate objectives in a medium-term time horizon, ensuring the distribution and marketing of electricity. The study was quantitative with non-experimental design, trend longitudinal cut, the population was 84 reports of monthly electricity consumption represented in (MWh/month), considering a commercial period of 7 years, the three phases of the Box-Jenkins methodology were applied as: identification, estimation and validation of the model, arriving to determine a multiplicative seasonal ARIMA (0,1,3)×(2,0,0,0)_12 model that adjusted perfectly to the series under study, being of "high precision" with a MAPE value of 2. 79% , finally in the last phase of the methodology, monthly forecasts of electricity consumption were made, with an average deviation for each forecast of a MAE value =625.80 MWh.

Keywords: Box-Jenkins, Forecasting, Electric Power

A. *Dennis Uriel Añasco-Chata, Facultad de Ingeniería Estadística e Informática-UNA-Puno,* rusbel92@hotmail.com
B. *Percy Huata-Panca, Facultad de Ingeniería Estadística e Informática-UNA-Puno,* phuata@unap.edu.pe
C. *Ecler Mamani-Vilca, Departamento de Ing. Informática y Sistemas- UNAMBA-Apurímac,* eclervirtual@unamba.edu.pe
D. *Adolfo Carlos Jimenez-Chura, Escuela Profesional de Ingeniería de Sistemas -UNA-Puno,* ajimenez@unap.edu.pe
E. *Pablo Cesar Tapia-Catacora, Escuela Profesional de Ingeniería de Sistemas -UNA-Puno,* pctapia@unap.edu.pe

1 INTRODUCCIÓN

La energía eléctrica es un elemento fundamental en el desarrollo de una sociedad, según el Banco Mundial [1]. La energía es esencial para hacer posibles las inversiones, las innovaciones y las nuevas industrias, que son los motores del empleo, incluido el crecimiento, y la prosperidad compartida de economías enteras [1]. Sin embargo, casi 675 millones de personas todavía viven sin electricidad en todo el mundo, y alrededor de 2.3 millones de personas cocinan o calientan sus hogares con combustibles contaminantes que dañan su salud y el medio ambiente. Las crisis energéticas afectan a la mayoría de los países, pero los países en vías desarrollo enfrentan las mayores cargas, especialmente los países importadores de energía [2] . Desde su aparición de la energía eléctrica en la humanidad, su demanda se encuentra en constante crecimiento el crecimiento de la vivienda y la población puede aumentar la demanda de electricidad, ya que más viviendas y edificios requieren electricidad para iluminación, calefacción y refrigeración [3].

Es fundamental tener información acerca del comporta-

miento futuro del consumo de energía eléctrica, resulta imperativo desarrollar un modelo matemático adecuado y confiable, que nos proporcione resultados certeros sobre el consumo de energía en los próximos tres años. Este modelo nos permitirá contar con estimaciones confiables y realistas, lo cual resulta esencial para poder planificar y tomar decisiones frente a una necesidad a futuro. En ese marco, la pregunta es ¿Cuál es el modelo determinado con la metodología Box - Jenkins para el pronóstico de consumo de energía eléctrica para los años 2021 – 2023 en la región de Puno?, como hipótesis principal se ha planteado que el modelo ARIMA proporciona un mejor ajuste para validar y pronosticar el consumo de energía eléctrica de la región de Puno, y como objetivo general es determinar un modelo de series de tiempo que pronostique el consumo de energía eléctrica a mediano plazo para la región de Puno, empleando modelos estocásticos univariados ARIMA conjuntamente con la metodología de Box – Jenkins.

Los reportes realizados por ELECTRO S.A.A en el año 2021 señalan que en la región de Puno desde enero del 2014 a diciembre 2020 el número de clientes se ha incrementado en 66,176 usuarios durante 7 años, teniendo a diciembre de 2020 un total de 403,141 usuarios; anualmente el incremento de clientes en promedio es de 11,029 representado por una tasa promedio de crecimiento del 3% anual, el reporte también revela que el consumo promedio anual de un cliente residencial en el año 2014 era de 46.14 Kwh, en cambio para el año 2020 el consumo promedio anual se elevó a 95.76 Kwh, reflejando un incremente del 52% durante 7 años, así mismo el sector industrial como la minería ha registrado incrementos durante los 7 años obteniendo en el año 2020 un incremento del 82.82 MW con respecto a los años anteriores hasta el año 2014.

En la región de Puno, el incremento en la demanda de energía eléctrica en la zona urbana y rural se fundamenta en el crecimiento de la población y el crecimiento poblacional, que hacen que la utilización de energía sea constante en los hogares, en el caso del sector industrial la demanda de bienes y servicios, la creación de nuevas empresas en el sector implica mayor consumo de energía, por ende el servicio de suministro de energía eléctrica debe ser continua y de calidad [4] .

Así también se encontraron diversos trabajos anteriores como: En Bagua Vázques y Gamonal [5] realizaron el pronóstico de consumo mensual de energía eléctrica, de la provincia de Bagua Grande, la investigación cuenta con información histórica de consumos de energía eléctrica de la Empresa Municipal de Servicios de Utcubamba SAC. de los periodos 2009 a 2015, Información que fue sometida a técnicas y pruebas estadísticas para identificar el modelo ARIMA, se concluye que el mejor modelo para el pronóstico de consumo de energía eléctrica es el modelo ARIMA(4,2,0), con un coeficiente de determinación de R^2=0,996. En Cusco Marcelo Barreto y Villareal Escate [6] logran determinan el modelo multiplicativo estacional ARIMA(0,1,1)×(1,0,2)_12 predice eficientemente la demanda mensual de energía eléctrica residencial de la región de Cusco, el modelo es respaldado por la métrica de avaluación MAPE con un valor porcentual del 2.296%, además señala que la aplicación de la metodología de Box- Jenkins se ajusta perfectamente a las series energía eléctrica históricas. En Puno, distrito de Ayaviri [4] realizó el modelo de pronóstico para el consumo de energía eléctrica para el distrito de Ayaviri, el objetivo fue determinar un modelo univariante que describa y pronostique el consumo de energía eléctrica, indicando que el modelo permite describir y predecir el comportamiento de consumo de energía eléctrica, es el modelo multiplicativo estacional ARIMA(3,1,1)×(1,0,1)_12.

Para entender ampliaremos algunas definiciones:

La palabra "pronóstico" proviene del griego "prognostikon" que significa "enunciado de lo que es probable que ocurra en el futuro". En el campo de la estadística los pronósticos son una estimación cualitativa o cuantitativa de una o varias variables que conforman un evento futuro, tomando como base información actual o del pasado [7].

La Metodología Box Y Jenkins es un modelo de series de tiempo, cuya función representa formalmente los procesos aleatorios (procesos estocásticos) generadas por las series reales, las misma que se pueden realizar por medio de modelos lineales de series de tiempo, que buscan describir los posibles modelos teóricos los cuales permitan explicar y detectar el comportamiento de la serie y de su proceso generador, tomando en cuenta la dependencia que existe entre los datos [8]. La principal característica es examinar el modelo más adecuado analizando errores recientes de pronóstico para seleccionar el ajuste apropiado, cumpliendo la condición de estacionariedad. Los procesos aleatorios estacionarios lineales se clasifican en tres tipos: modelos Autorregresivos(AR), modelos de Medias Móviles (MA) y modelos Mixtos ARMA, en los modelos con procesos no estacionarios se encuentran los modelos ARIMA y los modelos estacionales ARIMA [9].

En la década de los 70' en el año de 1973 George Edward Pelham Box y Gwilym Mirion Jenkins publican el libro "Análisis de Series Temporales Pronóstico y Control", años después se conocería como "metodología de Box-Jenkins" o técnicamente conocida como "metodología ARIMA", donde combinan los enfoques de modelos Autorregresivos y el modelo de Medias Móviles (estudiados inicialmente por Yule), desarrollando un proceso iterativo que consta de cuatro fases: Identificación del modelo, estimación de los parámetros o coeficientes , validación de los modelos seleccionados y predicción de valores, este proceso iterativo sirve como herramienta para la selección de un modelo que mejor se ajuste a la serie temporal obteniendo como resultado pronósticos con mayor grado de exactitud [10], resumimos dichas fases en el diagrama de la Fig 1.

C&T Riqchary
Revista de Investigación en Ciencia y Tecnología
ISSN: 2810-8124 (en línea) / ISSN: 2706-543x

Vol. 5 Núm. 1 (2023)
Publicado: 05/11/23
DOI: https://doi.org/10.57166/riqchary/v5.n1.2023.111

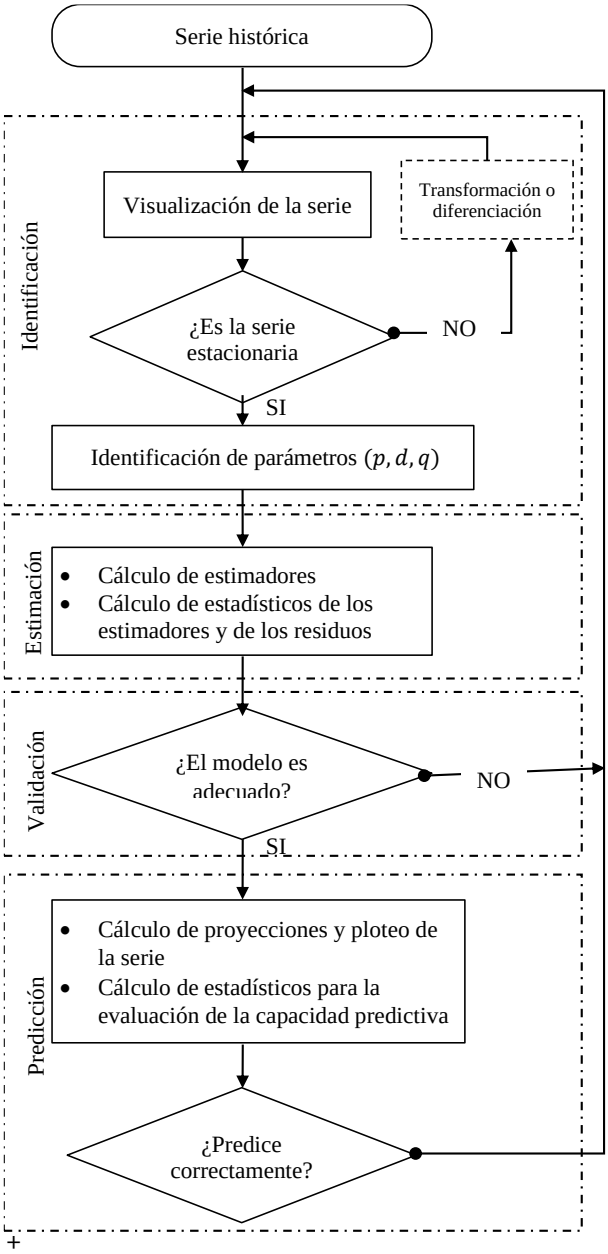


Fig. 1. Diagrama de flujo de la metodología Box–Jenkins.

2 MATERIALES Y METODOS

La investigación se desarrolló en la Región de Puno al Suroeste de la República del Perú; situada en la meseta del Collao a una altitud de 3,827 m s.n.m, cuenta con una extensión territorial de 71 999.00 Km2(representa el 6,0 % del territorio nacional) siendo la quinta Región más grande de nuestro territorio Nacional, esta subdividida en 13 Provincias y 110 distritos, cuenta con una población estimada al año 2017 de 1,172,697 habitantes de acuerdo al Censo Nacional de Población y Vivienda 2017 [11]. EL ámbito de estudio fue en la empresa ELECTRO PUNO S.A.A quien en amparo a la Ley de Concesiones Eléctricas D.L N° 25844 y su reglamento D.S 009-93-EM; con fecha 7 de diciembre del año 2000 el ministerio de energía y minas mediante Resolución Suprema N° 106-2000-EM y sus modificaciones, otorga a la empresa la concesión dentro de la región de Puno, a cargo de la operación de la Central Hidroeléctrica de Chijisia localizada en la Provincia de Sandia [12].

El estudio es cuantitativo, cuya finalidad es " explicar, predecir, verificar y controlar los fenómenos" [13], con diseño no experimental, de corte longitudinal de tendencia (trend), este tipo de diseño se basa en la recolección de los datos a través del tiempo en puntos o periodos [14]. Para el procesamiento y análisis de información de la serie de consumo de energía eléctrica, se usó el software libre R con las siguientes librerías: forecast, tseries, ggplot2, tidyverse, tidyr, dplyr, readxl, readxl, astsa, flextable, nortest, lmtest siendo las dos primeras las utilizadas para la estimación. La población fue, 84 reportes de consumo mensual de energía eléctrica representados en (MWh/mes), considerando un periodo comercial de 7 años que comprende desde enero-2014 a diciembre-2020 [13] ver Tabla 1. Para la obtención de la data se aplicó la técnica de revisión documental, teniendo acceso al software SIELSE COMERCIAL propiedad de ELECTRO-PUNO S.A.A, proporcionado los resúmenes de facturación mensualmente por cada periodo comercial.

TABLA 1
Consumo Mensual de Energía Eléctrica de la Región Puno 2014-2020

Mes / Años	2014	2015	2016	2017	2018	2019	2020
Enero	21095	22285	22848	23627	23988	23793	24580
Febrero	20942	21949	22186	24361	23144	24267	24468
Marzo	18513	21441	21772	21084	22499	21659	22807
Abril	21410	23165	23485	24222	24682	24810	21163
Mayo	21643	24112	22888	23518	23841	23968	17639
Junio	22841	23873	24323	25249	25628	25850	19225
Julio	22394	23359	23787	24985	25141	25366	20316
Agosto	22669	23130	23943	24144	24583	25400	22847
Setiembre	22351	24061	24200	24505	24982	25788	22759
Octubre	21594	23338	24154	24463	24134	24933	23911
Noviemb.	22746	23411	23592	24611	24519	24615	24465
Diciembre	22021	22770	23974	24150	23916	24138	24973

Datos mensuales del consumo de energía eléctrica en Megavatios por hora (MWh) de la Región de Puno, período 2014 a 2020 tomados de Electro Puno.

3 RESULTADOS

3.1 Identificación

En el comportamiento de la serie consumo de energía eléc-

trica se identificaron a cuatro componentes de la serie (tendencia, ciclo, aleatoriedad y estacionalidad) utilizando métodos de descomposición; mediante un análisis de normalidad de datos, correlogramas y el test aumentado de Dickey – Fuller, se determinará la estacionariedad de la serie, de ser el caso se estabilizará la serie en media y varianza para la identificación de posibles modelos. El análisis al comportamiento de la serie en estudio, se afirma que la serie consumo de energía eléctrica presenta una tendencia creciente positiva (cumpliendo el requisito) Figura 11, donde se observa la distribución de los valores individuales de la serie consumo de energía eléctrica; de el consumo máximo de energía fue de $Max = 25{,}850.55\ MWh$, seguido de un consumo mínimo de energía eléctrica de $Min = 17{,}639.05\ MWh$; el consumo promedio obtenido de la serie es de $\bar{\chi} = 23{,}356.93\ MWh$ (Línea horizontal verde) con una desviación estándar de $\sigma = 1{,}570.98$ y una mediana de $\hat{\chi} = 23{,}817.16$ (Línea horizontal azul), de acuerdo a los datos presentados afirmamos que los valores de la serie sigue una distribución asimétrica negativa hacia la izquierda tal como se muestra en la Figura 2, donde los valores se distribuyen hacia la izquierda con una tendencia variante hacia arriba y abajo de 1,570.98 MWh con respecto a su promedio, todo indica que los valores de la serie consumo los valores individuales de la serie no giran en torno a su promedio y varianza, por lo expuesto la serie es "no estacionaria".

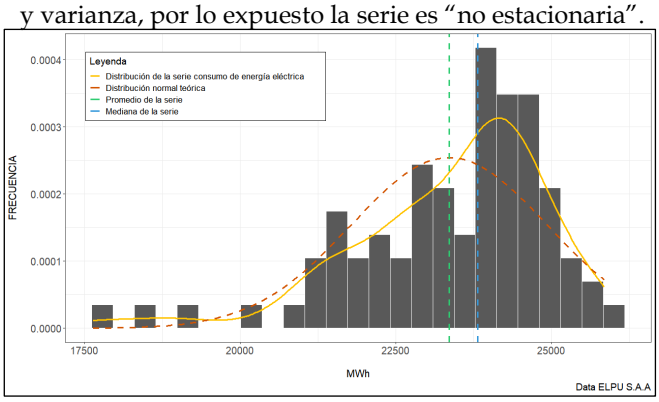


Fig. 2. Histograma de la Serie Consumo de energía eléctrica de la Región Puno, frente al histograma teórico de la serie.

Por el análisis visual se determinó "No estacionaria" para validar usaremos el Test Aumentado de Dickey – Fuller, este comprobará a través de un contraste de hipótesis si la serie presenta una tendencia significativa. El resultado con un nivel de di significancia ∝=0.05 la decisión basados en al Tabla 3 fue, se dispone de evidencia suficiente para aceptar la hipótesis nula. De acuerdo a los resultados se confirma que la serie presenta estacionalidad, bajo estos resultados se deberá diferenciar o transformar la serie a una serie estacionaria.

TABLA 2
Test Aumentado de Dickey – Fuller

Statistic	P.value	Parameter	Method	Alternative
-2.9	0.1917	4.0	Augmented Dickey-Fuller Test	stationary

Signif. codes: 0 <= '***' < 0.001 < '**' < 0.01 < '*' < 0.05

Vázquez [5] en su estudio indican que la serie consumo de energía eléctrica de la provincia de Bagua Grande, presenta un comportamiento no estacionario, afirmando que los valores de la serie varían en un 45.8% con respecto a su promedio, para comprobar la estacionariedad de la serie los investigadores aplican la prueba estadística de Levene, obteniendo como resultado un valor de significancia del $11.006 > 0.05$ concluyendo que la serie es no estacionaria, hallazgo que concuerda con nuestro estudio.

3.2 Segunda etapa de la fase identificación

Diferenciación de la serie después de haberse aplicado la primera diferencia ($d = 1$), en el proceso de diferenciación la serie sufre la pérdida de un valor inicial y los valores de la serie cambian en su totalidad, esto debió a la aplicación de la fórmula de diferenciación $y_t = x_t - x_{t_1}$ con el objetivo de estabilizar la serie en media y varianza, para eliminar la tendencia y estacionalidad de la serie.



Fig. 3. En la parte superior serie irregular el grafico inferior la serie tiene un recorrido lineal mensual con valores diferenciados que van de arriba hacia abajo entorno a su media sin mostrar signos de tendencia, claramente luego de haber diferenciado la serie original esta se ha convertido en una serie estacionaria tal como lo exige la metodología de Box – Jenkins validado por Test aumentado de Dickey – Fuller donde se obtuvo P= 0.0107<0.05 con un nivel de significancia del 5% se dispone de evidencia suficiente para rechazar la hipótesis nula. Por lo tanto, la serie transformada consumo de energía eléctrica es estacionaria; se confirma que la serie no presenta tendencia ni estacionalidad, bajo estos resultados es posible la identificación del orden de los parámetros del modelo multiplicativo ARIMA(p,d,q) × (P,D,Q)12.

Este análisis es respaldado por Quispe [4] quien refiere que, luego de haber estudiado la serie original de consumo energía eléctrica de la provincia de Ayaviri, se determinó que es ne-

C&T Riqchary
Revista de Investigación en Ciencia y Tecnología
ISSN: 2810-8124 (en línea) / ISSN: 2706-543x

Vol. 5 Núm. 1 (2023)
Publicado: 05/11/23
DOI: https://doi.org/10.57166/riqchary/v5.n1.2023.111

cesario diferenciar la serie en un periodo $d = 1$, para conseguir estabilizar la serie original en una serie estacionaria. Con los correlogramas mostrados en la **Figura 4** es posible identificar el orden de los parámetros del modelo multiplicativo estacional $ARIMA(p,d,q) \times (P,D,Q)_{12}$, realizando un contraste con el orden de los parámetros teóricos identificados por Box - Jenkins, sin embargo existe en el software r versión 4.2.1 la librería forecast para la identificación de modelos $ARIMA$ automáticamente y esta funcion se denomina "auto.arima()", esta función está desarrollada en base a la metodología Box – Jenkins, siendo una herramienta fundamental para la identificación de modelos $ARIMA$ eficientemente.

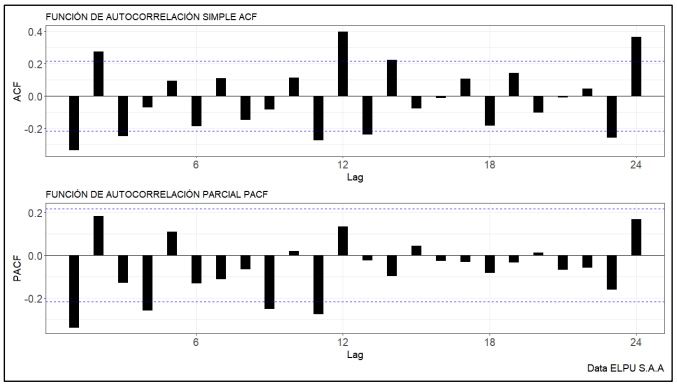


Fig. 4. Función de autocorrelación simple y parcial de la primera diferencia de la serie.

Se presenta la **Fig. 4**, donde se observa la función de autocorrelación simple y parcial de la serie diferenciada en 1 periodo, claramente se evidencia que la serie no presenta tendencia de acuerdo a los rezagos mostrados en el correlograma, tal como se ha confirmado con el test Aumentado de Dickey Fuller. En relación a la función de autocorrelación simple deducimos que los rezagos 1,2 y 3 sobresalen del intervalo del límite de confianza siendo estos rezagos significativos de cero, seguido de rezagos no significativos. En referencia a la función se autocorrelación parcial se identifica a los rezagos 1 y 4, ambos rezagos se encuentran fuera del intervalo del límite de confianza, también seguidos de rezagos no significativos. Por su parte Donado & Huertas realizó una avaluación critica a la función auto.arima() concluyendo que " a pesar de existir métodos automáticos para la identificación de los modelo de series de tiempo, es necesario tener conocimiento teórico de las series temporales y conocimiento técnico de la función auto.arima(), considerando estos factores se reducirá el tiempo de trabajo y se obtendrá resultados confiables para la generación de pronósticos [16]" .

3.3 Tercera etapa de la fase de identificación

Se generaron 71 modelos candidatos, véase en la **¡Error! No se encuentra el origen de la referencia.**, para el pronóstico de consumo de energía eléctrica, en la selección del modelo este debe cumplir con las condiciones del Criterio de Información de Akaike y el criterio de Información de Bayes (BIC) ambos criterios señalan que "mientras más pequeño sea el valor del AIC y BIC, mejor será el modelo para el pronóstico " teniendo en consideración ambos criterios el modelo óptimo es el modelo 1:

$$ARIMA(0,1,3) \times (2,0,0)_{12}$$

TABLA 3
Modelos ARIMA(p,d,q)×(P,D,Q)$_{12}$ identificados

Modelos	$ARIMA(p,d,q) \times (P,D,Q)_{12}$	AIC	BIC
Modelo 1	$ARIMA(0,1,3) \times (2,0,0)_{12}$	**1383.29**	**1397.80**
Modelo 2	$ARIMA(0,1,3) \times (1,0,1)_{12}$	1384.80	1399.32
Modelo 3	$ARIMA(1,1,3) \times (1,0,0)_{12}$	1386.29	1400.81
Modelo 4	$ARIMA(0,1,3) \times (1,0,0)_{12}$	1387.30	1399.40
Modelo 5	$ARIMA(3,1,0) \times (2,0,0)_{12}$	1387.56	1402.07
⋮	⋮	⋮	⋮
Modelo 71	$ARIMA(0,1,0) \times (0,0,1)_{12}$	1406.24	1411.078

71 modelos identificados en el software R con la función auto.arima()

3.4 Estimación

Para la estimación de parámetros según el modelo multiplicativo estacional $ARIMA(0,1,3) \times (2,0,0)_{12}$, para estabilizar la serie en media y varianza se aplicó la diferencia de un periodo en su componente regular $d = 1$, el parámetro seleccionado para la parte regular de orden (q) es una Media Móvil de orden 3 $MA(3)$, del mismo que se deriva los coeficientes $MA(1), MA(2)$ y MA(3), en cuanto a la parte estacional del modelo se identifica el parámetro Autorregresivo AR(2), del cual se deriva los coeficientes $AR(1)$ y $AR(2)$, identificados los coeficientes del modelo se estima la ecuación (1).

$$(1 - \Phi_{12}B^{12} - \Phi_{24}B^{24})(1 - B)y_t = (1 - \omega_1 B - \omega_2 B^2 - \omega_3 B^3)\varepsilon_t \quad (1)$$

Donde:

$y_t =$ *Variable independiente*
$\Phi_P =$ *Coeficiente Autorregresivo estacional*
$B =$ *Operador de retador*
$\varepsilon_t =$ *Error*
$\omega_q =$ *Coeficiente de Media Móvil regular*
$(1 - B) =$ *Diferencia regular*

En la **Tabla 4**, se presenta los estimadores y coeficientes del modelo $ARIMA(0,1,3) \times (2,0,0)_{12}$, donde la probabilidad para los coeficientes $MA(1)$, $MA(3)$, $SAR(1)$ y $SAR(2)$ son significativos a un nivel de confianza del 5% ($p - value < 0.05$); con respecto al coeficiente $MA(2)$ la probabilidad no es significativa, sin embargo el estimador es menor que 1 con un valor de $0.0312 < 1$ considerandose bueno, este coeficiente podría ser considerado para ser excluido del modelo, no obstante el modelo ha sido seleccionado utilizando la técnica de estimación de máxima verosimilitud (MLE) [17], por lo tanto el coeficiente si aporta al modelo para el ajuste de la serie.

$$(1 - \Phi_{12}B^{12} - \Phi_{24}B^{24})(1 - B)y_t = (1 - \omega_1 B - \omega_2 B^2 + \omega_3 B^3)\varepsilon_t \quad (2)$$

TABLA 4
Coeficientes del modelo ARIMA

$ARIMA(0,1,3) \times (2,0,0)_{12}$

Coefficients:

term	estimate	std.error	statistic	p.value
ma1	-0.2258	0.1137	-1.9863	0.046997 *
ma2	0.0312	0.1230	0.2534	0.799962
ma3	-0.5168	0.0943	-5.4817	0.00000004212 ***
sar1	0.2631	0.1240	2.1222	0.033821 *
sar2	0.3696	0.1382	2.6741	0.007492 **

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

sigma^2 = 859900: log likelihood = -685.64

AIC=1383.29 AICc=1384.39 BIC=1397.80

Logrado estimar los parámetros del modelo y remplazando los estimadores de los coeficientes del modelo multiplicativo estacional $ARIMA(0,1,3) \times (2,0,0)_{12}$ en la ecuación **¡Error! No se encuentra el origen de la referencia.**, la estimación del modelo, es la siguiente:

$$(1 - 0.2631B^{12} - 0.3696B^{24})(1 - B)y_t = (1 + 0.2258B - 0.0312B^3 + 0.5168B^3)\varepsilon_t \quad (3)$$

El modelo seleccionado fue validado mediante el análisis de los correlogramas simples (ACF) y Parcial (PACF), donde los residuos del modelo 1 a través de la significancia de los rezagos, así determinar si el modelo se ajusta razonablemente a los datos de la serie original.

La **Fig. 5** muestra los correlogramas simples y parcial del modelo multiplicativo estacional $ARIMA(0,1,3) \times (2,0,0)_{12}$ donde se observa que los rezagos en ambos correlogramas se encuentran dentro del umbral de los límites de confianza al 95%, esto indica que los residuos del modelo seleccionado son puramente aleatorios y que los datos no están correlacionados una de la otra.

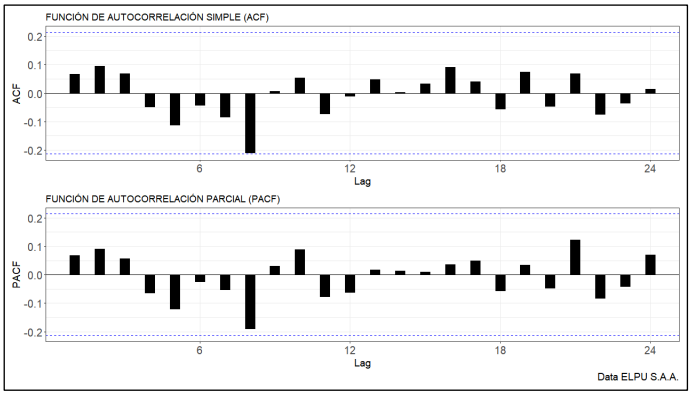


Fig. 5. Autocorrelación Simple y Parcial de los Residuos del Modelo Seleccionado Función de autocorrelación simple y parcial de los residuos del modelo $ARIMA(0,1,3) \times (2,0,0)_{12}$

Por su parte Barahona [18] comprueba la existencia de independencia de su modelo $ARIMA(0,1,1) \times (1,0,0)_{12}$ por el mismo análisis, indicando que no existen retardos significativos en los correlogramas, corroborando que los residuos del modelo son independientes. Vázquez [19] realizan un análisis similar a los residuos mediante las Autocorrelaciones Simples y Parciales del modelo $ARIMA(4,2,0)$, donde los retardos en ambos correlogramas se encuentran dentro de los límites de confianza, el modelo presenta aleatoriedad en su residuo siendo estos de ruido blanco, por lo tanto según refieren los investigadores el modelo es adecuado para la representación de la serie.

Comprobamos estadísticamente a través del planteamiento de una hipótesis el supuesto de independencia mediante el Test de Q de Box – Lung donde se obtuvo P = 0.5316≥0.05 con un nivel de significancia del 5%, se acepta la hipótesis nula. Por lo tanto, los residuos del modelo seleccionado cumplen con el supuesto de independencia esto quiere decir que no existe correlación en los residuos, superado con éxito el proceso de validación del modelo, por tanto, el modelo es óptimo para ser implementado.

En los resultados obtenidos por Quispe [4] en la fase de validación de su modelo multiplicativo estacional $ARIMA(3,1,1) \times (1,0,1)_{12}$ realiza el Test de Q de Ljung-Box para probar la independencia de los datos, obteniendo como resultado una probabilidad superior al $\propto = 0.05$, comprobando su hipótesis nula que los residuos de la serie son independientes, es decir son ruido blanco.

3.5 Pronóstico

La evaluación de la capacidad predictiva se observa en la **Tabla 5** se presenta las métricas para la medición del grado de acierto de un pronóstico, entre el más utilizado es la RMSE (Raíz media de los errores al cuadrado) así como también se usó para comparar modelos seleccionando a aquel modelo que contenga el menor valor estadístico. En este caso, el RMSE del modelo 1 es de 893.57 siendo el valor más bajo entre los cuatro modelos, asegurando que el modelo tiene un mejor ajuste a la serie original por contener el promedio mínimo de errores (diferencia entre la serie original y el modelo ajustado), siendo este un valor aceptable considerando que el modelo seleccionado cumple con el principio de parsimonia apoyado por los criterios de AIC y BIC, ajustándose adecuadamente a la serie original.

TABLA 5
Modelos ARIMA y sus Métricas de Precisión del Pronóstico

Mod.	$ARIMA(p,dk,q) \times (P,D,Q)_{12}$	AIC	BIC	RMSE	MAE	MAPE
1	$ARIMA(0,1,3) \times (2,0,0)_{12}$	1383.29	1397.80	893.576	625.800	2.790
2	$ARIMA(0,1,3) \times (1,0,1)_{12}$	1384.80	1399.32	908.244	624.464	2.787
3	$ARIMA(1,1,3) \times (1,0,0)_{12}$	1386.29	1400.81	924.697	643.741	2.864
4	$ARIMA(0,1,3) \times (1,0,0)_{12}$	1387.30	1399.40	943.120	656.495	2.929
5	$ARIMA(3,1,0) \times (2,0,0)_{12}$	1387.56	1402.07	907.978	649.547	2.880

El valor seleccionado 625.80 en la métrica de evaluación MAE (error absoluto medio) es el indicando. El valor MAPE obte-

C&T Riqchary
Revista de Investigación en Ciencia y Tecnología
ISSN: 2810-8124 (en línea) / ISSN: 2706-543x

Vol. 5 Núm. 1 (2023)
Publicado: 05/11/23
DOI: https://doi.org/10.57166/riqchary/v5.n1.2023.111

nido en modelo en términos porcentuales es del 2.79%, ubicando el valor en la escala de Lewis se traduce que los pronósticos serán de alta precisión, tal como se muestra en la **Tabla 6**. Barahona [18] compara la precisión de su modelo $ARIMA(0,1,1) \times (1,0,0)$ frente al modelo Holt – Winters, encontrado las métricas de evaluación RMSE de 5,3434, una MAE del 4,1358 y un MAPE de 2,4206%, estos indicadores son inferiores a los obtenidos por el método de Holt – Winters, asegurando mediante estas métricas de evaluación que el modelo $ARIMA$ en cuanto a capacidad predictiva es precisa.

TABLA 6
Escala de Lewis(1982) Para el Análisis de Precisión MAPE

Porcentaje MAPE	Precisión del pronóstico
Menos de 10%	Alta precisión
11% a 20%	Buen pronóstico
21% a 50%	Pronóstico razonable
51% a más	Pronóstico inexacto

Se determinó el valor MAPE obtenido por nuestro modelo en términos porcentuales es del 2.79%, ubicando el valor en la escala de Lewis se traduce que los pronósticos son de alta precisión.

Barreto [6] demuestran que el modelo $ARIMA(0,1,1) \times (1,0,2)_{12}$ es mejor en comparación con el método de suavización exponencial Multiplicativo de Winter, afirmación que se apoya bajo los resultados obtenidos mediante el error absoluto porcentual promedio (MAPE), el modelo ARIMA presenta un valor MAPE del 2.296% y el modelo Suavizado exponencial de Holt Winter presenta un valor MAPE del 2.4%, demostrando de esta manera que el modelo $ARIMA$ es quien mejor se ajusta al pronóstico de demanda de energía eléctrica en Cusco, ambas investigaciones coinciden.

Ejecución de pronóstico, se observa en la Fig. **6** se muestra los valores proyectados de consumo de energía eléctrica para la región Puno, para los próximos 3 años aun nivel del 95% de confianza con el modelo multiplicativo estacional $ARIMA(0,1,3) \times (2,0,0)_{12}$ de la ecuación (3) es:

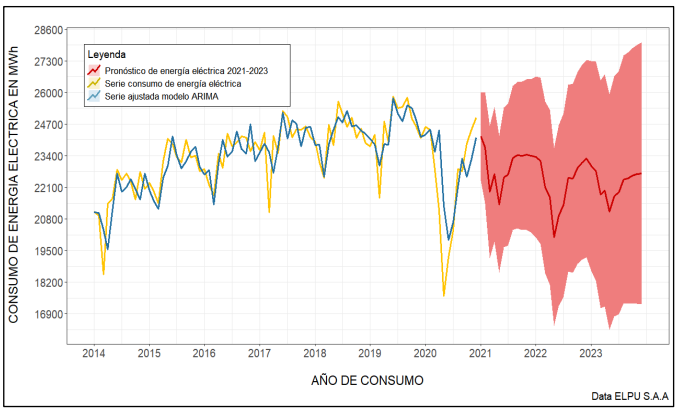


Fig. 6 Pronóstico del consumo de energía eléctrica para los próximos 3 años (2021 al 2023).

4 DISCUSIONES

El modelo determinado es multiplicativo estacional $ARIMA(0,1,3) \times (2,0,0)_{12}$, este modelo es respaldado por los criterios de $AIC = 1383.29$ y $BIC = 1397.80$ valores que son inferiores frente a otros modelos, demostrando que el modelo es el apropiado cumpliendo con el principio de parsimonia, Barahona [18] en su investigación determina el modelo multiplicativo $ARIMA(0,1,1) \times (1,0,0)_{12}$ reportando indicadores de $AIC = 364.89$ y un $BIC = 371$ valores que son inferiores frente a otros modelos, mientras que Quispe [4]determina su modelo únicamente a través del indicador $BIC = 2476.5$ siendo este el menor valor frente a otros 3 modelos, en ambas investigaciones determina su modelo seleccionando el menor valor obtenido en los indicadores AIC y BIC; de esta manera se prueba que el modelo proporciona un mejor ajuste a la serie de tiempo. Así mismo, se ejecutaron los pronósticos de consumo mensual de energía eléctrica a 36 meses (3 años), con una capacidad de predicción del 2.79% considerándose de alta precisión, al respecto Barreto [6]realizan una comparación de capacidad predictiva entre el modelo ARIMA contra el modelo Multiplicativo Winter, seleccionando el modelo ARIMA por presentar un valor MAPE del 2.96%, siendo inferior frente al otro modelo, resultado que coincide con nuestra investigación. Por otra parte (Arévalo Cabrera & Santisteban Morales, 2020) seleccionaron el modelo multiplicativo estacional $ARIMA(2,1,2) \times (1,0,0)_{12}$ con una capacidad de predicción del valor $MAPE = 19.32\%$ resultado que difiere con nuestra investigación, pues los pronósticos del modelo serán buenos, pero no precisos.

5 CONCLUCIONES

En el análisis a la serie consumo de energía eléctrica se muestra que factores atípicos como la pandemia en el año 2020 que distorsionan la serie de consumo de energía eléctrica, este fenómeno afecta directamente en la identificación del modelo ARIMA, pues estos modelos capturan los últimos datos de la serie para su implementación en el modelo, ello desfavorece al pronóstico con el menor margen de error.

Se ha estimado los parámetros con el método de máxima verosimilitud del modelo multiplicativo estacional $ARIMA(0,1,3) \times (2,0,0)_{12}$, obteniendo los siguientes coeficientes a un nivel de confianza del 5 %, indicando que son óptimos: $MA(1) = -0.2258$, $MA(2) = 0.0312$, $MA(3) = -0.5168$, $SAR(1) = 0.2631$, $SAR(2) =$, y está representada por la siguiente ecuación general (3)

Los pronósticos generados por el modelo multiplicativo estacional $ARIMA(0,1,3) \times (2,0,0)_{12}$ tienen un grado de alta precisión, de acuerdo a las pruebas realizadas al modelo en cuanto se refiere a la evaluación de la capacidad predictiva, el modelo ha obtenido valores inferiores en los indicadores $MAPE = 2.79\%$ y un $MAE = 625.80$ frente a otros modelos candidato, demostrando de esta forma que el modelo es apropiado en todos los aspectos.

REFERENCIAS

[1] Banco Mundial, «https://www.worldbank.org,» 09 09 2023. [En línea]. Available: https://www.worldbank.org.

[2] B. Mundial, «https://www.worldbank.org,» 12 09 2022. [En línea]. Available: https://www.worldbank.org/en/news/feature/2018/04/18/access-energy-sustainable-development-goal-7.

[3] M. Phillips, «https://marketrealist.com,» Market Realist, 10 10 2022. [En línea]. Available: https://marketrealist.com/2014/09/must-know-factors-impact-electricity-demand/.

[4] Y. L. R. Quispe Pacco, «Modelo univariante para el consumo de energía eléctrica Doméstica en el Distrito de Ayaviri–Electro Puno, periodo 2004-2013,» UNA , Puno, 2015.

[5] K. Y. Vásquez Díaz y M. E. Gamonal Sánchez , «Modelo para el pronósito del consumo mensual de energía electrica, de la provincia de Bagua Grande, mediante la metología Box Jenkis par eñ año 2016,» Universidad de Pedro Ruíz Gallo, Lambayeque, 2019.

[6] L. G. Villarreal Escate y E. A. Marcelo Barreto, «El modelo estocástico univariante ARIMA como herramienta predictiva de la demanda de energía eléctrica residencial del sistema eléctrico Cusco,» Universidad Nacional de Ingeniería, Lima, 2021.

[7] F. Villareal, «Introducción a los Modelos de Pronósticos,» Universidad Nacional del Sur, Bahia Blanca, 2016.

[8] F. Parra, *Estadística y Machine Learning con R*, Madrid: https://bookdown.org/, 2019.

[9] E. Court Monteverde y E. Williams Rengifo, Estadísticas y econometría financiera, Argentica: Cengage Learning Argentina, 2005.

[10] D. Peña Sanchez de Rivera, Análisis de series temporales, Madrid: Alianza, 2005.

[11] Instituto Nacioanal de Estadística e Informática, Censos Nacionales 2017: XII de poblacion, VII de vivienda y III de Comunidades Indígenas, Lima: INEI, 2022.

[12] Electro Puno S.A.A, Plan estratégico Institucional 2017-2021, Electro Puno S.A.A, 2021, p. 74.

[13] E. E. Gallardo Echenique, Metodología de la Onvestigación. Manual autoinformativo interactiva I, vol. 1, Huancayo: Universidad Continental, 2017, p. 22.

[14] R. Hernadez Sampieri, C. Fernández Collado y P. Batista Lucio, Metodología de la Investigación, Cuarta ed., Mexico: MACGRAW-HILL, 2010.

[15] Electro Puno - S.A.A, Memoria Anual 2020, Empresa Regional de Serviciós Públicos de Electricidad , 2020.

[16] N. Huertas Gnecco y S. Donado Torres, «Evaluación crítica de la función Auto. Arima de R en la identificación de procesos de series de tiempo,» Universidad de los Andes, Bogotá , 2018.

[17] R. J. Hyndman y G. Athanasopoulos, «orecasting: principles and practice,» OTexts, 2018.

[18] J. I. Barahona Bonifaz, «Pronósticos del consumo y demanda de potencia máxima de energía eléctrica en la ciudad de Riobamba para el periodo 2017-2020,» Escuela Superior Politécnica de Chimborazo, Chimborazo, 2018.

[19] K. Y. Vásquez Díaz y M. Estefany Gam, «Modelo para el pronóstico del consumo mensual de energía eléctrica, de la provincia Bagua Grande, mediante la metodología de Box y Jenkins, para el año 2016,» Universidad Pedro Ruíz Gallo, Lambayeque, 2019.

Biografías:

Dennis Uriel Añasco-Chata, Universidad Nacional del Altiplano Puno – Perú. Ingeniero Estadístico e Informático, desarrollador de aplicaciones estadísticas e informáticas.

Percy Huata-Panca, Universidad Nacional del Altiplano Puno – Perú, Doctoris Scientiae en Economía y Gestión, Doctorado en Estadística Aplicada, Magister Scientiae en Informática, Ingeniero Estadístico, Docente nombrado a tiempo completo en la Universidad Nacional del Altiplano Puno – Perú.

Ecler Mamani-Vilca, Universidad Nacional Micaela Bastidas de Apurímac - Perú, Dr. en Ciencias de la Computación, desarrollador de aplicaciones multimedia y Software Educativo Intercultural, docente nombrado a tiempo completo en la Uni-versidad Nacional Micaela Bastidas de Apurímac.

Adolfo Carlos Jimenez-Chura, Universidad Nacional del Altiplano Puno – Perú, Doctoris Scientiae en Ciencias de la Computación, Magister Scientiae en Informática, Ingeniero de Sistemas, Docente nombrado a tiempo completo en la Univer-sidad Nacional del Altiplano Puno – Perú.

Pablo Cesar Tapia-Catacora, Universidad Nacional del Altiplano Puno – Perú, Doctoris Scientiae en Ciencias de la Computación, Magister en Contabilidad y Administración, Ingeniero de Sistemas, Docente nombrado a tiempo completo en la Universidad Nacional del Altiplano Puno – Perú.

C&T Riqchary
Revista de Investigación en Ciencia y Tecnología
ISSN: 2810-8124 (en línea) / ISSN: 2706-543x

Vol. 5 Núm. 1 (2023)
Publicado: 15/10/23
DOI: https://doi.org/10.57166/riqchary/v5.n1.2023.113

Análisis del comportamiento geomecánica del macizo rocoso en galería nivel 4040, mina Victoria

Analysis of the geomechanical behavior of the rock mass in gallery level 4040, Victoria mine

David F. Catalán-Mendoza [A], José A. Cárdenas-Catalán [B]. Wilson J. Mollocondo-Flores [C]

0000-0001-7077-8254[A],0000-0002-1155-0943 [B] y 0000-0003-3445-7941[C]

(Recepción: 15/04/2023 y aceptación 10/10/2023)

Resumen— La tendencia en los países con expriencia en explotación minera, es tener practicas seguras para conservar su recurso más valioso, su personal; con tal motivo nos planteamos el objetivo de determinar el comportamiento geomecanico del macizo rocoso analizados mediante clasificaciones geomecánicas galería nivel 4040 en la mina victoria, Andahuaylas 2021; el estudio es de tipo aplicada, nivel; descriptivo – explicativo. Los resultados se obtuvieron mediante la aplicación de programas computacionales como el Roclab, Dips y Phases; el valor del RMR (corregido) es de 48 lo que califica como regular; lo que conlleva que en excavación como Corte piloto y banqueo con 1.5 a 3 m de avance en el corte piloto, los pernos de roca deben pernos sistematicos de 4m, espaciado 1.5-2 m en techo y paredes con malla soldada; el concretro lanzado de 50 a 100 mm en techo y 30 mm en los lados. La investigación concluye que en zonas de taludes en dominio estructural (D.E) 1 no existe riesgo pues no hay polos, en D.E 4 existe concentración de polos por ende posibilidad de deslizamiento. En estructuras de túneles en el D.E 2 no existe ni volteo ni vuelco, en D.E 3 y 5 existen cuñas puesto el primer caso se debe aplicar shotcrete, y para el segundo aplicar de una fina capa de shotcrete.

Palabra clave: clasificación geomecánica, macizo rocoso, resistencia, teoría de Meyerhof

Abstract— The trend in countries with experience in mining exploitation is to have safe practices to conserve their most valuable resource, their personnel; For this reason, we set ourselves the objective of determining the geomechanical behavior of the rock mass analyzed through geomechanical classifications gallery level 4040 in the Victoria mine, Andahuaylas 2021; The study is of applied type, level; descriptive – explanatory. The results were obtained by applying computer programs such as Roclab, Dips and Phases; the RMR value (corrected) is 48, which qualifies as regular; which means that in excavation such as pilot cut and benching with 1.5 to 3 m of advance in the pilot cut, the rock bolts must be systematic bolts of 4 m, 1.5-2 m spacing on the ceiling and walls with welded mesh; the concrete thrown from 50 to 100 mm on the roof and 30 mm on the sides. The investigation concludes that in slope areas in structural domain (D.E) 1 there is no risk since there are no poles, in D.E 4 there is a concentration of poles therefore the possibility of landslide. In tunnel structures in D.E 2 there is no overturning or overturning, in D.E 3 and 5 there are wedges since in the first case shotcrete must be applied, and for the second case a thin layer of shotcrete must be applied.

Keyword: geomechanical classification, rock mass, strength, Meyerhof theory

A. *David F. Catalán_Mendoza[A] está en Universidad Nacional Micaela Bastidas de Apurimac. Correo electrónico: davidcatalan@unamba.edu.pe*
B. *José A. Cárdenas Catalán [B] está con Departamenteo Ingenierias de la Universidad Nacional Micaela Bastidas de Apurimac, Correo electrónico: jcardenas@unamba.edu.pe.*
C. *Wilson J. Mollocondo Flores, Departamenteo Ciencias Básicas de la Universidad Nacional Micaela Bastidas de Apurimac, Correo electrónico: wmollocondo@unamba.edu.pe.*

1 INTRODUCCIÓN

La explotación minera en los países internacionales han optado a tener practicas seguras para salvaguardar el recurso más valioso, su personal, **[1]** razón a ello la tecnificación y aplicar el conocimiento de las herramientas y teorías para determinar el comportamiento de los macizos rocoso en su debido espacio y tiempo **[2]**.

La unidad minera victoria se vio obliga a realizar estudios al macizo rocoso en la galería en ejecución con fines múltiples sobre todo en evitar el desprendimiento de rocas y poder afectar a la vida humana que es recurso fundamental en la operación minera, en segundo lugar, los equipos y herramientas, el proceso y entre otros afectando el proceso y de responsabilidad administrativa y judicial por lo tanto la investigación es aplicar la teoría y contribuir ese conocimiento en esta mina

Analizando la realidad problemática estimar cual será el comportamiento geomecánico del macizo rocoso en la ejecución de la galería nivelo 4040 de la mina victoria. Es preciso la formulación del problema, considerando como problema general ¿ Cuál es el comportamiento del macizo rocoso analizados mediante clasificaciones geomecánicas galería nivel 4040, mina victoria, Andahuaylas 2021?

El estudio se basa en una justificación práctica, debido a que al concluirla podremos conocer el comportamiento del macizo rocoso analizados mediante clasificaciones geomecánicas, galería nivel 4040, mina victoria, Andahuaylas 2021, de la misma manera tiene justificación teórica.

Según Ochoa **[3]** en la tesis de posgrado cuyo objetivo es propuesta de clasificación geomecánica que permita evaluar el comportamiento del macizo rocoso en el frente de arranque del túnel Levisa-Mayari Tramo IV, mediante una investigación aplicada, obtuvo el macizo rocoso esta fracturado con tres familias de grietas predominantes y según las clasificaciones geomecánicas el mismo se clasifica de calidad muy mala a mala. A partir de ajustes de parámetros del índice de calidad Q, se obtiene una ecuación de cálculo con resultados similares a los medidos, Q=2*10-07 GSI4.244 ecuación de ajuste de tipo exponencial R^2=0.999, lo que representa el comportamiento en tiempo real del macizo rocoso en el túnel según la clasificación de GSI ajustada finalmente, concluye que del análisis de las diferentes clasificaciones geomecánicas usadas en túneles, de las características geomecánicas del macizo rocoso y del ajuste a sus parámetros, que la clasificación GSI de Hoek y Marinos con los ajustes pertinentes es la que más se adecuada a las particularidades de macizo rocoso estudiado, la clasificación Q de Barton ofrece mejores resultados para proyectar soluciones de sostenimiento

A nivel nacional, según Valdivia **[4]** en su estudio "Estimación del comportamiento geomecánico, mediante método de Kriging para el macizo rocoso del cerro Callacpuma - sector Huayrapongo, 2020", una investigación aplicada, con alcance descriptivo y diseño no experimental, encontró en sus resultados: fueron RCU (resistente a muy resistente), JRC (rugosa), RQD (media a buena), espaciado (medianamente junto). Llegando a concluir que la estimación representa una fuente confiable para identificar las zonas críticas del macizo; y deben ser tomadas en cuenta por parte de los resposables para establecer las medidas preventivas necesarias.

En función a lo descrito párrafos anteriores nos planteamos el problema de investigación ¿Cuál será comportamiento geomecanico del macizo rocoso analizados mediante clasificaciones geomecánicas galería nivel 4040, mina victoria, Andahuaylas 2021?

El planteamiento se justifica por ser aporte de la teoría y con la finalidad de determinar el comportamiento geomecanico de las rocas mediante las clasificaciones geomecánicas para su proyección de la ejecución de la galería nivel 4040 en producir el mineral de manera segura y eficiente previniendo los accidentes sobre todo cuando exista los desprendimientos de las rocas.

Ademas se establecico el objetivo determinar el comportamiento geomecanico del macizo rocoso analizados mediante clasificaciones geomecánicas galería nivel 4040, mina victoria, Andahuaylas 2021; mientras que la hipótesis a contrastar es el comportamiento geomecanico del macizo rocoso analizados mediante las clasificaciones geomecánicas son estables.

2 MARCO TEÓRICO

2.1 Clasificación Geomecánicas

Las clasificaciones usadas fueron:

– *Clasificación RMR de Bieniawski* : Esta clasificación geomecánica RMR o Rock Mass Rating fue desarrollado por Bieniawski, en base a más de 300 casos reales de túneles, cavernas, taludes y cimentaciones [5]. Actualmente se usa la edición 1989, que coincide sustancialmente con la de 1979. [6] Permite a su vez relacionar índices de calidad con parámetros de diseño y de sostenimiento de túneles, se obtiene estimando los parametros: resistencia a la compresión simple de la matriz rocosa, R.Q.D. Grado de fracturamiento del macizo rocoso, Espaciado de las discontinuidades, Condiciones de las discontinuidades, Presencia del agua, Orientación de las discontinuidades [7].

– *Clasificación de Barton*: Se basa mayoritariamente en casos tipo de estabilidad en excavaciones subterráneas, el Norgerian Geotechnical Institute (N.G.I.) propone un índice para determinar la calidad del macizo rocoso en túneles y taludes

C&T Riqchary
Revista de Investigación en Ciencia y Tecnología
ISSN: 2810-8124 (en línea) / ISSN: 2706-543x

Vol. 5 Núm. 1 (2023)
Publicado: 15/10/23
DOI: https://doi.org/10.57166/riqchary/v5.n1.2023.113

[8]. Considera 6 parametros: R.Q.D, Índice de fracturamiento (Jn), Índice de la rugosidad, relleno y continuidad de las discontinuidades (Jr), [9]Índice de la alteración de las discontinuidades (Ja), Coeficiente reductor por la presencia de agua (Jw), el coeficiente que tiene en cuenta la influencia del estado tensional sobre el macizo rocoso (SRF) [10].

– *Clasificación según índice GSI:* Es un sistema de caracterización de macizos rocosos desarrollado en la ingeniería de mecánica de rocas para satisfacer la necesidad de datos de entrada confiables [11]; relacionados con las propiedades del macizo rocoso requeridos para diseño de túneles, taludes o cimentaciones en rocas [12]. El carácter geológico del material rocoso, junto con la valoración visual del mismo, son usados directamente como datos de entrada para la selección de los parámetros para predecir la resistencia y deformabilidad del macizo rocoso [13]. Esto también provee un método en campo para caracterizar macizos rocosos difíciles de describir [14] .

2.2 Diseño del sostenimiento

A través del Q de Barton se pueden determinar parámetros adicionales relacionados con el sostenimiento:

a) Máximo vano sin sostener

Es la longitud máxima de perforación y voladura que se puede realizar en un solo ciclo sin fortificación [15]. Estas limitaciones en la longitud del avance están definidas por las características geotécnicas de la masa rocosa que se esté excavando [16].

b) Carga de roca sobre la clave (Pr)

Se refiere a la presión permanente que soporta el techo de la excavación [17], puede determinarse mediante la siguiente expresión según el número de familias de discontinuidades:
Para macizos con menos de tres familias de discontinuidades

$$Pr(kp/cm2) = (2\sqrt{}Jn)/(3Jr\ 3\sqrt{}Q)$$

Para macizos rocosos con más de tres familias de discontinuidades

$$Pr(kp/cm2) = 2/(3Jr\ 3\sqrt{}Q)$$

3 METODOLOGÍA

El tipo de investigación es aplicada; según Egg [18] consiste en utilizar los conocimientos y aplicarlos para resolver problemas; mientras que para le diseño experimento según Behar [19] considera que el diseño no experimental es observar los sucesos sin participar y dejar que ocurran naturalmente.
Respecto a la población según, Lerma [20]considera la población como un conjunto de elementos del mismo tipo con cualidades fijas, por lo tanto nuestra población está compuesta por galería principal nivel 4040, labores de exploración, chimenea y un tajeo en preparación de la mina Victoria. Mientras que para la muestra se usa el maestro noprobaiblistico mediante la tecnica de criterio del investigador lo que nos permite establecer la muestra como la Galería del Nivel 4040 de la Unidad Mina Victoria

4 RESULTADOS Y DISCUSIÓN

Parámetros de Resistencia para cada Dominio Estructural en la Túneles de la mina victoria

Los parámetros de los dominios estructurales:

- DE-1 y DE-4 en zonas de taludes
- DE-2, DE-3, y DE-5 en túneles

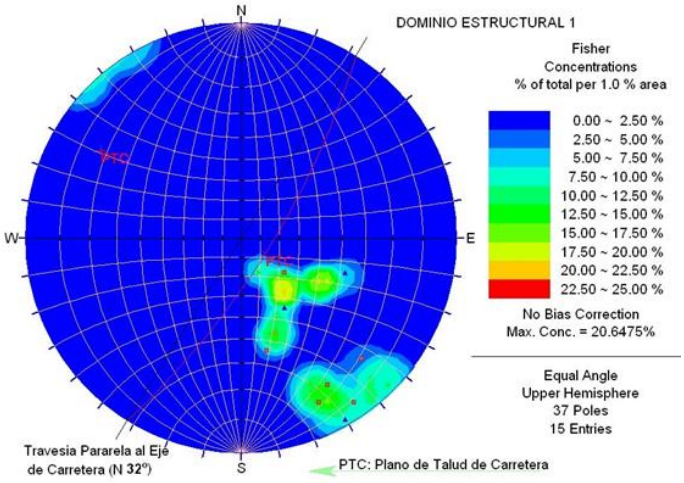


Fig. 1. Concentración de familias. D: E-1

De la figura 1, existe concentraciones importantes de diaclasas en el sentido del eje de vía, hacia el talud, en dirección N32° E. En este dominio se encuentran 2 zonas de alteración (margas) con azimut de 140° y un buzamiento 50°, que fueron atravesados perpendicularmente, por la vía, no se tiene inestabilidad en el talud. No existe ningún tipo de riesgo debido a la ausencia de discontinuidades. No hay polos

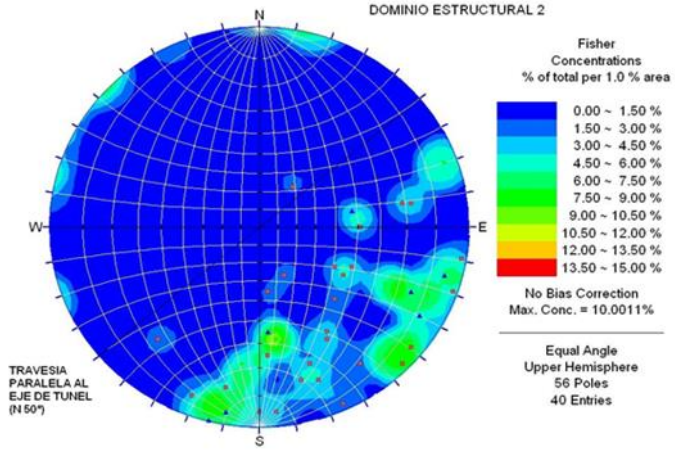


Fig. 2. Diagrama estereográfico de contornos de densidades de polos, concentración de familias D.E-2

Revista de investigación en ciencia y tecnología
Vol. 5 Núm. 1 (2023) - publicado: : 15/10/23
DOI https://doi.org/10.57166/riqchary/v5.n1.2023.113

Mientras que en la figura 2, existe concentraciones de diaclasas, una falla y discontinuidades dispersas que convergen al hastial derecho del túnel (eje N50E), que pueden ocasionar inestabilidad en el entorno de la superficie de excavación por lo que se recomienda colocar una fina capa de concreto lanzado (shotcrete). En el sets de concentraciones se genaran planos de corte, se visualiza la formación de una cuña pequeña por el lado izquierdo de la direccion de la excavación, por lo que se evaluó la inestabilidad y se recomienda tomar medidas de refuerzo para asegurar el equilibrio aplicando shotcrete

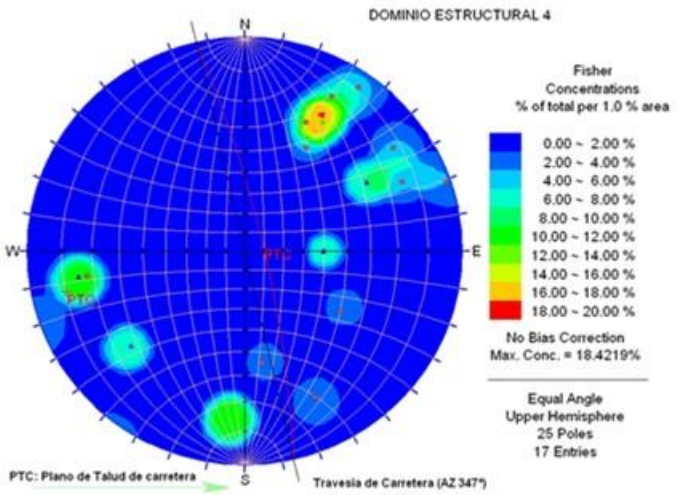


Fig. 3. Diagrama de densidad de polos de la dispersión de polos Del D.E-3

Asi mismo en la figura 3, existe concentraciones importantes de diaclasas hacia el hastial derecho del túnel (eje N52E); teniendo en consideración el efecto de la meteorización en las fisuras, se deben tomar las precauciones de estabilidad con una fina capa de concreto lanzado. El sets de concentraciones muestran interseciones forman pequeñas cuñas por el lado izquierdo de la direccion de excavación. Evaluado el lugar, para asegurar el equilibrio se debe aplicar shotcrete

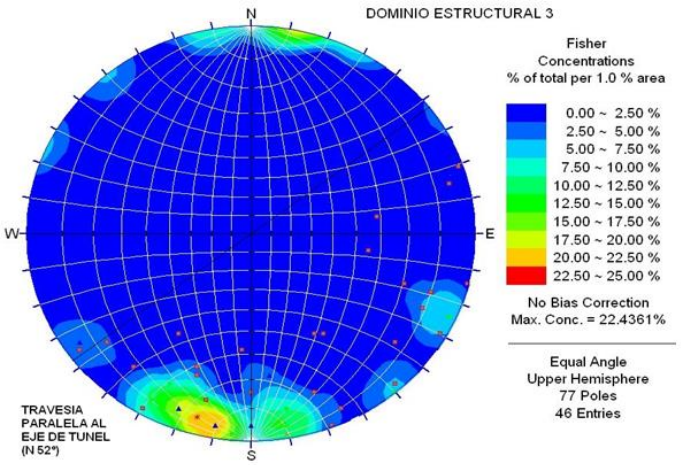


Fig. 4. Diagrama Diagrama de densidad de polos de la dispersión de polos mostrado en Del D.E-4

De la figura 4, existen concentraciones puntuales e importantes de diaclasas con distribución en ambos hastiales, tomando como referencia el eje de la vía (Az. 347º) y encontrándose 2 fallas (Buz. 65º), continuas y rellenas con marga, pueden condicionan inestabilidad en el talud. Posible vuelco por la cresta del talud debido a la mayor concentración de polos; Existe una concentración de polos aproximadamente el 50 % del set, que podría ocasionar un deslizamiento

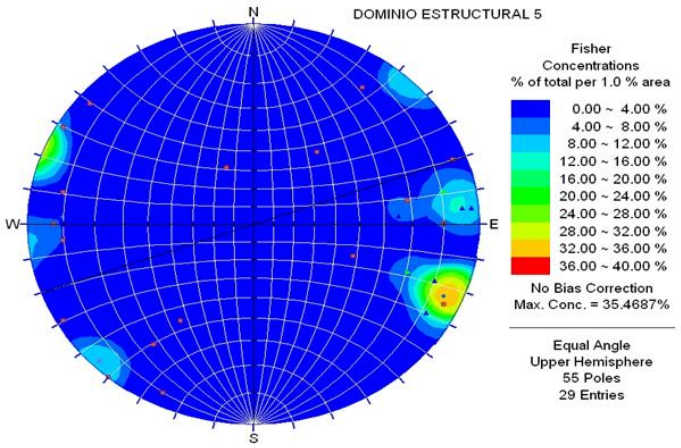


Fig. 5. Diagrama de densidad de polos de la dispersión de polos Del D.E-5

Presencia de concentraciones importantes de diaclasas en dirección hacia el hastial derecho del eje del túnel (N71E). No hay inestabilidad. En el sets de concentraciones, cuyas interseciones formaran pequeñas cuñas. Se visualiza una cuña pequeña por el lado izquierdo de la direccion de la excavación. Evaluado el equilibrio del macizo, requiere la aplicación de una fina capa de shotcrete.

Por último el valor de RMR (básico) es de 53, el RMR (corregido) es 48 y RMR (condiciones secas) es de 58; y segun la clasificación es considera como regular (rango de 60-41) y lo que esta considerado en excavación como Corte piloto y banqueo con 1.5 a 3 m de avance en el corte piloto. El sostenimiento se instala después de cada voladura. Completo sostenimiento de los 10 m hasta el frente. En pernos de roca se considera Pernos sistemáticos de 4 m de longitud, espaciados 1.5-2 m en el techo y las paredes con malla soldada en el techo. Para concreto de lanzado debe ser 50-100 mm en el techo y 30 mm en los lados.

C&T Riqchary
Revista de Investigación en Ciencia y Tecnología
ISSN: 2810-8124 (en línea) / ISSN: 2706-543x

Vol. 5 Núm. 1 (2023)
Publicado: 15/10/23
DOI: https://doi.org/10.57166/riqchary/v5.n1.2023.113

5 Discusión

Para el dominio estructural uno y cuatro sobre zonas de taludesse puede afirmar que no existe ningún tipo de riesgo debido a la ausencia de discontinuidades; No hay polos en el primero. Mientras para el dominio estrucutral cuatro es posible vuelco por la cresta del talud debido a la mayor concentración de polos; Existe una concentración de polos aproximadamente el 50 % del set, que podría ocasionar un deslizamiento.

De otro lado para los dominios estructurales 2, 3, 5 y 6 que se refiere a túneles: para DE-2 en el sets de concentraciones se genaran planos de corte, se visualiza la formación de una cuña pequeña por el lado izquierdo de la direccion de la excavación, por lo que se evaluó la inestabilidad y se recomienda tomar medidas de refuerzo para asegurar el equilibrio aplicando shotcrete. Se concluye que no existe ni volte ni vuelco. Para DE-3 el sets de concentraciones muestran interseciones forman pequeñas cuñas por el lado izquierdo de la direccion de excavación. Evaluado el lugar, para asegurar el equilibrio se debe aplicar shotcrete. Para el DE-5 el sets de concentraciones, cuyas interseciones formaran pequeñas cuñas. Se visualiza una cuña pequeña por el lado izquierdo de la direccion de la excavación. Evaluado el equilibrio del macizo, requiere la aplicación de una fina capa de shotcrete

6 Conclusiones

Para las zonas de taludes en Dominio estructural (D.E) 1 no existe riesgo por ausencia de discontinuidades; para D.E 4 se encontró un vuelco de cresta de talud lo que generaría un deslizamiento.

En estructuras de túneles en D.E 2 no existe ni volteo ni vuelco. Para D.E 3 muestra pequeñas cuñas y para equilibrar se debe aplicar shotcrete. Para D.E 5 se encontró una cuña pequeña en el lado izquierdo de excavación por lo que se requiere aplicar una fina capa de shotcrete.

Se tiene un RMR (corregido) de 48 y de acuerdo a la clasificación geomecánica este valor nos indica que es una roca de regular a buena

Referencias

[1] O. U. Gonzales, R. G. Molina y D. F. Patarroyo, «Condiciones de seguridad y salud en el trabajo, una revision teórica desde la mineria Colombiana,» *Revista Venezolana de Gerencia*, vol. 24, nº 85, 2019.

[2] M. R. Hernandez, «Mineria ilicita: incidencia en la convivencia y seguridad ciudadana en diez departamentos de Colombia,» *Revista criminalidad*, vol. 58, nº 1, pp. 47-65, 2016.

[3] A. Ochoa Quezada, Propuesta de una clasificacion geomecanica para la evaluacion del comportamiento del macizo rocoso en el frente de arranque del tunel Levisa-Mayari Tramo IV, Moa, cuba, 2018.

[4] A. F. Valdivia Donayre, Estimacion del comportamiento geomecanico, mediante metodo de kriging para el macizo rocoso del cerro callacpuma-sector huayrapongo,2020, cajamarca, 2020.

[5] M. Romana, J. B. Seron y E. Montalar, «La clasificación geomecánica SMR: aplicación, experiencias y validación,» V Simposio nacioal sobre taludes y laderas inestbles, Valencia, 2001.

[6] Z. Bieniawski, Rock mechanics design in mining and tunneling. A.A. Balkema, Rotterdam, 272pp, 1989.

[7] B. Celada y e. al, «Actualización y mejora del RMR,» *Ingeopres*, vol. 234, pp. 18-22, 2014.

[8] L. A. Torres, A. A. Ruiz, J. D. Huaricancha y J. G. Vizcarra, «Geomecánica aplicada al control de las labores mineras para la minimización de caída de rocas en la mina Madrugada Cía. minera Huinac SAC-Ancash-2009,» *Aporte santiaguino*, vol. 3, nº 1, pp. 35-40, 2010.

[9] M. Perez, E. Garcia, C. Vega, J. Montoya y P. Noriega, «Estudio comparativo entre sistemas de clasificación geomecánica en un depósito tipo Pórfido,» *Boletin de ciencias de la tierra*, pp. 34-44, 2018.

[10] D. Castro, D. Carrera, B. Muñoz y A. Garcia, «Correlación entre el Índice RMR de Bieniawski y el Índice Q de Barton en Suelos de Baja Calidad,» *Revista de la construcción*, vol. 9, nº 1, pp. 107-119, 2010.

[11] M. O. Morales, «Caracterización geotécnica y determinación de ángulos detalud en yacimiento franke,» Universidad de Chile, Santiago de Chile, 2009.

[12] P. Ramirez y L. Alejano, «Mecanica de rocas: Fundamentos e ingenieria de taludes,» Universidad polit{ecnica de Madrid, Madrid, 2004.

[13] P. Marinos y E. Hok, GSI: a geologically friendly tool for rock mass strength estimation. In ISRM international symposium. International Society for Rock Mechanics, 2000.

[14] A. Ochoa, M. Cartaya y J. L. Blanco, «Clasificación geomecánica óptima para evaluar el macizo rocoso en el frente de arranque del Tramo IV del túnel Levisa-Mayarí,» *Mineria y geología*, vol. 36, nº 1, pp. 50-64, 2020.

[15] H. Gaspar y E. Illanes, «Comparativo entre la clasificación geomecánica de indice de resistencia geólogica y rango del maxizo roccoso en el sostenimiento de la masa rocosa,» Universidad nacional de Huancavelica, Huancavelica, 2015.

[16] A. A. Nguema, M. C. Pier, G. Romero y A. Recio, «Comportamiento estructural del macizo rocoso de la

cantera Baney, Bioko Sur, Guinea Ecuatorial,» *Mineria y geología*, vol. 38, nº 3, pp. 266-276, 2022.

[17] L. V. Carreño, C. A. Ochoa, H. A. Fuentes y J. M. Carvajal, «Estimación de propiedades mecánicas de roca utilizando inteligencia artificial,» *Ingenieria y ciencias*, vol. 7, nº 14, pp. 83-103, 2011.

[18] E. A. Egg, Aprender a investigar: Nociones básicas para la investigación social, 1a ed., Córdoba, Argentina: Brujas, 2011, p. 190.

[19] D. S. Behar, Metodología de la investigación, Bogotá, Colombia: Shalom, 2008, p. 94.

[20] H. D. Lerma, Metodología de la investigación propuesta, anteproyecto y proyecto, 4a ed., Bogotá, Colombia: Ecoe Ediciones, 2009, p. 190.

Revista de Investigación en Ciencia y Tecnología
ISSN: 2810-8124 (en línea) / ISSN: 2706-543x

Vol. 5 Núm. 1 (2023)
Publicado: 15/10/23
DOI: https://doi.org/10.57166/riqchary/v5.n1.2023.115

Análisis del rendimiento en Clústeres basados en PelicanHPC

Analysis of performance in PelicanHPC-based Clusters

John A. Aguirre-Carrasco[A], Luis M. Alfaro-Chirinos [B] y Alejandrina Huaylla-Quispe [C]

ORCID:0000-0002-1522-2963[A], ORCID:0000-0002-0475-6994[B] y ORCID: 0000-0002-0258-8632[C]

(Recepción: 15/04/2023 y aceptación 05/10/2023)

Resumen— La implementación de clústeres de alto rendimiento tienen aplicación en entornos de agrupamiento de nodos físicos independientes y de manera virtualizada mediante un middleware especializado o un Sistema Operativo dedicado, con el objetivo de obtener mayor recurso computacional a partir de equipos básicos. Ante ambas opciones, surge la pregunta: ¿Qué implementación arroja mejores resultados en cuanto al rendimiento?, por lo que el presente estudio tiene por finalidad realizar un análisis al rendimiento de clústeres en dos escenarios: diez equipos de cómputo homogéneos interconectados a una red local por cable trenzado CAT 6a y un switch; y de otra parte, diez máquinas virtualizadas en el software VirtualBox, con recursos repartidos equitativamente como son: memoria RAM, procesadores y disco. La metodología adoptada fue experimental en cinco ciclos repetitivos y trabajando en función a las medias del tiempo de ejecución El algoritmo de prueba es un contador de números primos codificado en lenguaje C++, generando su archivo de salida y posterior ejecución a nivel de consola sobre la terminal en el Sistema Operativo PelicanHPC (variante de CD live del SO Debian). Los resultados son uniformes inicialmente, trabajando de 1 a 10 nodos, donde el resultado demuestra más de 30% de mejora en el clúster tradicional frente a un progreso irregular del clúster virtualizado, puntos que serán objeto del análisis descriptivo en base a resultados obtenidos.

Palabras clave: clúster de alto rendimiento, paralelismo, virtualización, PelicanHPC.

Abstract— The implementation of high-performance clusters has application in environments of grouping of independent physical nodes and in a virtualized manner through specialized middleware or a dedicated Operating System, with the aim of obtaining greater computational resources from basic equipment. Given both options, the question arises: Which implementation yields better results in terms of performance? Therefore, the purpose of this study is to carry out an analysis of the performance of clusters in two scenarios: ten homogeneous computing devices interconnected to a local network. by braided CAT 6a cable and a switch; and on the other hand, ten virtualized machines in the VirtualBox software, with resources distributed equally such as: RAM memory, processors and disk. The methodology adopted was experimental in five repetitive cycles and working based on the average execution time. The test algorithm is a prime number counter coded in C++ language, generating its output file and subsequent execution at the console level on the terminal. on the PelicanHPC Operating System (live CD variant of the Debian OS). The results are initially uniform, working from 1 to 10 nodes, where the result demonstrates more than 30% improvement in the traditional cluster compared to irregular progress in the virtualized cluster, points that will be the subject of descriptive analysis based on the results obtained.

Keywords: high performance cluster, parallelism, virtualization, PelicanHPC.

A. John A. Aguirre Carrasco· Universidad Nacional Micaela Bastidas de Apurímac, jaguirre@unamba.edu.pe
B. Luis Miguel Alfaro Chirinos, Universidad Nacional Micaela Bastidas de Apurímac, lalfaroch@unamba.edu.pe
C. Alejandrina Huaylla Quispe, Universidad Nacional Micaela Bastidas de Apurímac, ahuaylla@unamba.edu.pe

1 INTRODUCCIÓN

El uso de los clústeres nace a partir de la década de los 60, donde se proyecta la conexión de computadores en una primeriza ARPANET (Advanced Research Projects Agency Network) o en español Red de la Agencia de Proyectos de Investigación Avanzada, toman como solución el agrupamiento de ordenadores funcionales independientes para componer una supercomputadora que disponga de altos recursos computacionales (tales como procesadores, memorias principales, disco duro, etc.). **[1]- [2]**

Gracias al avance de la tecnología en todo el campo de la informática, mejora también el tema de redes y comunicaciones, que permite reforzar el vínculo computador – clúster **[2]**, **[3]**, **[4]**, por la naturaleza de conectividad hacia una Red de Área Local (LAN) y la implementación de protocolos de Internet tales como el TCP/IP **[1]**.

La necesidad y las nuevas tecnologías permiten asimilar también el tema de la programación concurrente y distribuida, más específicamente el paralelismo, que consiste en enfocar el trabajo de más de un procesador en una sola tarea específica, contando con los recursos del ordenador **[3]**, **[5]**. Entonces las bases del clúster computacional toman una nueva perspectiva operacional, clasificada más adelante en las definiciones de Clústeres de Alto Rendimiento, Clústeres

de alta disponibilidad y clústeres de alta eficiencia **[6]**.

Los clústeres de alta eficiencia o HPCC (High Performance Computing Clúster) son el enfoque de este trabajo de investigación **[7]**, **[5]**. Mas específicamente en los resultados que arrojan dos implementaciones desarrolladas hoy en día: la tradicional con ordenadores independientes en unidades funcionales y maquinas virtualizadas, que es el modo tendencia de trabajo hoy por hoy en los data center de las diferentes organizaciones **[8]**.

La motivación que nos lleva a realizar este trabajo es alcanzar un análisis del rendimiento en dos escenarios de implementación: una tradicional y otra mediante un software que virtualiza equipos. Con este análisis se pone a disposición de los investigadores dos panoramas claros de cómo el rendimiento puede variar dependiendo de los recursos con los que cuenta o podría contar para implementar un clúster, utilizando para ello el proceso deductivo - inductivo del método científico **[9]**. A su vez que según el enfoque de la solución nos lleva a nuevas preguntas como: ¿es mejor o peor implementar un clúster virtualizado?, ¿cuál es el máximo rendimiento obtenido en un clúster tradicional?, ¿existe la implementación mixta (tradicional y virtualizada) en clústeres y cómo es el análisis de su rendimiento?, etc., preguntas que nacen a partir del análisis de rendimiento de clústeres.

Los aportes de esta investigación se encuentran en el análisis experimental singular en cada implementación propuesta, para después contrastar ambos resultados estadísticos; de tal manera que se pueden observar patrones con una tendencia a la baja o patrones irregulares anticipando a los resultados. Para lo cual en el punto dos, vemos los métodos y procedimientos utilizados a lo largo de la experimentación, en el punto tres tenemos los resultados obtenidos de manera experimental, en el punto cuatro las discusiones a partir de investigaciones anteriores, en el punto cinco las conclusiones del grupo de investigación y finalmente las referencias que son una parte esencial de todo trabajo de investigación científico.

2 MÉTODOS Y PROCEDIMIENTOS

2.1 Clúster de alto rendimiento implementado mediante virtualización

La virtualización de Sistemas Operativos trabaja sobre un software que simulación bajo arquitecturas x86/amd64 llamado Oracle VM VirtualBox en su versión 7.0.4, donde es posible implementar la instalación de sistemas virtualizados llamados "sistemas invitados", que se apoyan física y lógicamente sobre un computador, denominado "anfitrión" **[5]**, **[10]**, **[11]**.

Se debe señalar que los data center actuales trabajan de esta manera para lograr un trabajo distribuido y compartido debido al alto coste de los equipos, con la indicación que el trabajo sobre licencia de operación resulta conveniente por las funciones completas del virtualizador **[10]**, **[12]**.

Como componente middleware y Sistema Operativo que se utilizó para la virtualización se tiene a PelicanHPC en su versión GNOME 4.1; que resulta una distribución de otro SO libre: Debian **[13]**. Este middleware toma las siguientes características:

TABLA 1
Características de maquina virtualizada sobre PelicanHPC

Componente	Características
Procesador	*1 procesador 2.9GHz*
Memoria principal	*2GB*
Disco	*10 GB*

Evaluación experimental para diez máquinas virtuales

El equipo anfitrión es un equipo portátil Dell Alienware R4 17, con un procesador I7 de séptima generación de 8 núcleos a 2.9GHz, memoria principal de 32GB DDR4 y un disco de 1TB SSD M.2 NvME PCIe. El Sistema Operativo del computador anfitrión es Windows 10 Home con una arquitectura de 64 bits **[14]**.

Características que son suficientes para soportar la virtualización de diez máquinas virtuales, las cuales trabajaron en una red interna, trabajando sobre el adaptador PCnet -FAST III (Am79C973).

2.2 Clúster de alto rendimiento implementado de forma tradicional

Es el modo de trabajo primigenio de implementación, desde la década de los 60 y de los cuales se tiene resultados positivos **[12]**. El trabajo empírico trabaja en base a una red en un entorno punto a punto sobre el protocolo TCP/IP, por lo que se tomará el uso de un switch y el trabajo de envío y recepción de paquetes mediante la dirección MAC de la tarjeta de red del equipo. Las características del equipo son las siguientes:

TABLA 2
Características de los equipos de cómputo

Componente	Características
Procesador	*I5 4GEN 3.2GHz 4 NÚCLEOS*
Memoria Principal	*8GB DDR3 1333MHz*
Almacenamiento	*512GB SSD SATA*
Sistema Operativo	*Windows 10 Pro 64 bits*

Evaluación experimental para diez equipos de cómputo de escritorio

Sobre el Switch para la conectividad local (red LAN) se tiene un equipo en la marca D-Link DES-1024D de 10/100Mbits de transmisión y 24 puertos y diez patch cord armados con cable UTP Cat6A, RJ45 Cat6A SATRA crimpeados de acuerdo al estándar de cableado T-568B **[15]**, con dimensiones desde 1 metro a 4 metros. Se indica que los patch cord fueron

C&T Riqchary
Revista de Investigación en Ciencia y Tecnología
ISSN: 2810-8124 (en línea) / ISSN: 2706-543x

Vol. 5 Núm. 1 (2023)
Publicado: 15/10/23
DOI: https://doi.org/10.57166/riqchary/v5.n1.2023.115

testeados y su funcionamiento estuvo correcto hasta la obtención de resultados.

2.3 Sobre la aplicación

Se tomó por casos prácticos el algoritmo de búsqueda de números primos debido al orden de su complejidad O(n^2), que nos permite obtener los resultados en tiempos prudentes **[16]**.

TABLA 3
Búsqueda de números primos con límite superior en base a exponentes del número dos desde 2^0 hasta 2^17

Lim	Números primos	Conteo
2^0	---	*0*
2^1	*2*	*1*
2^2	*2,3*	*2*
2^3	*2,3,5,7*	*4*
2^4	*2,3,5,7,11,13*	*6*
2^5	*2,3,5,7,11,13,17,19,23,29,31*	*11*
2^6	*2,3,5,7,11,13,17,19,23,29,31,37,41,43,47,53,59,61*	*18*

Evaluación práctica en papel

Donde su programación plasmada en un algoritmo en el lenguaje C++ (IDE ZinjaI) resulta ser sencillo de la siguiente manera:

```
#include <iostream>
#include<iomanip>
#include<time.h>
#include<unistd.h>
#include<math.h>
using namespace std;
bool primo(int n){
	bool rpta=false;
	int cont=0;
	for(int i=1;i<=n;i++){
		if(n%i==0){
			cont++;
		}
	}
	if(cont==2){
		rpta=true;
	}return rpta;
}
int Numprimos(int m){
	int contador=0;
	for(int j=2;j<=m;j++){
		if(primo(j)){
			contador++;
		}
	}
	return contador;
}
int main(int argc, char *argv[]) {
	int k=pow(2,0);
	cout<<setw(20)<<"N";
	cout<<setw(20)<<"S";
	cout<<setw(20)<<"TIEMPO"<<endl;
	for(long p=k;p<=131072;p=p*2){
		cout<<setw(20)<<p;
		double tiempoinvertido=0.0;
		clock_t inicio=clock();
		cout<<setw(20)<<Numprimos(p);
		clock_t fin=clock();
		tiempoinvertido+=(double)(fin-
		inicio)/CLOCKS_PER_SEC;
		cout<<setw(20)<<tiempoinvertido<<endl;
	}
	return 0;
}
```

2.4 Booteo de PelicanHPC

Se contó con dos contextos: el virtualizado y el tradicional, en caso del contexto virtualizado, nos limitamos a vincular el origen al archivo ISO del SO PelicanHPC: "*pelicanhpc-v4.1-gnome.iso*".

Para el contexto tradicional se booteo el archivo ISO de PelicanHPC "*pelicanhpc-v4.1-gnome.iso*", mediante el ejecutable Rufus4-1.exe, a un esquema de partición MBR, sobre diez memorias pendrive de 4GB de capacidad.Se debe indicar que el riesgo operacional de los equipos no se ve comprometido debido a la naturaleza LIVE del SO, sabiendo que las modificaciones dentro del SO LIVE no se mantienen al reiniciar el equipo [17].

2.5 Ejecución sobre consola

Para la ejecución se toma como referencia al archivo C++ del algoritmo de búsqueda de números primos, ubicada en la carpeta *Documents*. Para lo cual procede la generación del archivo de salida, con el cual se ejecutará el algoritmo. Procedemos a primeramente trabajar en modo super usuario, mediante el código: "***user@pel1:~/Documents$*** *sudo su*", a continuación se procedió a la creación del archivo de salida (el que necesitamos para la ejecución) mediante el comando: "***root@pel1:/home/user/Documents#*** *g++ *nombre_archivo.cpp* -o *nombre_salida.out**". Concluido este paso, procedemos a la ejecución del archivo de salida, mediante el comando: "***root@pel1:/home/user/Documents#*** *mpirun -np 1 ./*nombre_salida.out**",para a continuación tener el siguiente resultado:

Revista de investigación en ciencia y tecnología
Vol. 5 Núm. 1 (2023) - publicado:15/10/2023
DOI https://doi.org/10.57166/riqchary/v5.n1.2.23.115

N	S	TIEMPO
1	0	1e-06
2	1	1e-06
4	2	1e-06
8	4	1e-06
16	6	2e-06
32	11	4e-06
64	18	1.2e-05
128	31	3.9e-05
256	54	0.000149
512	97	0.000493
1024	172	0.001895
2048	309	0.007481
4096	564	0.029209
8192	1028	0.115536
16384	1900	0.462548
32768	3512	1.83687
65536	6542	7.3272
131072	12251	29.3093

Fig. 1. Resultados de ejecución del archivo de salida, correspondiente al algoritmo de búsqueda de números primos, donde la columna N corresponde al límite en base a exponentes en base dos, que va desde 2^0 hasta 2^17. La columna S corresponde a los números primos que se encontraron en dicho intervalo y la columna TIEMPO cuantifica el tiempo en unidad de segundos.

Se hace la indicación que el procedimiento explicado se replicó trabajando desde un nodo (maestro), incrementándose secuencialmente hasta los diez nodos (un nodo maestro + nueve nodos esclavos); en los dos contextos inicialmente presentados: forma virtualizada y forma tradicional.

Además de obtener los resultados, se obtuvo la media de cinco pasadas en cada caso de evaluación para obtener resultados significativos, mostrados a continuación en la sección resultados.

3 RESULTADOS

Los resultados se muestran de manera independiente, en los dos contextos, primeramente, en el entorno virtualizado; donde se tomaron tres casos interpolados, mostrando los resultados en un diagrama lineal. Pasamos a mostrar los resultados experimentales.

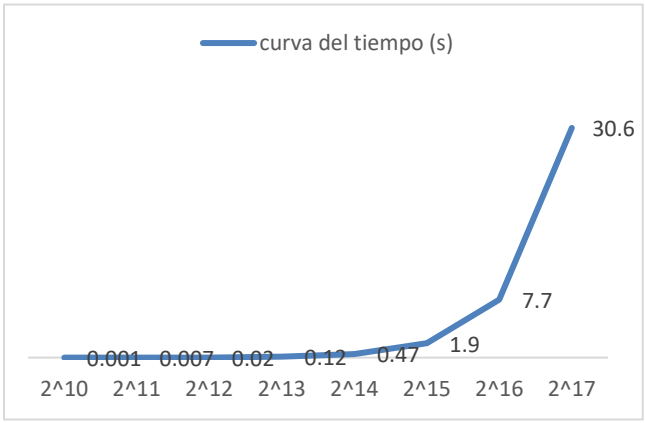


Fig. 2. Análisis de rendimiento en máquinas virtualizadas a un nodo, tomados desde el valor 2^10 al 2^17, debido a la fácil lectura.

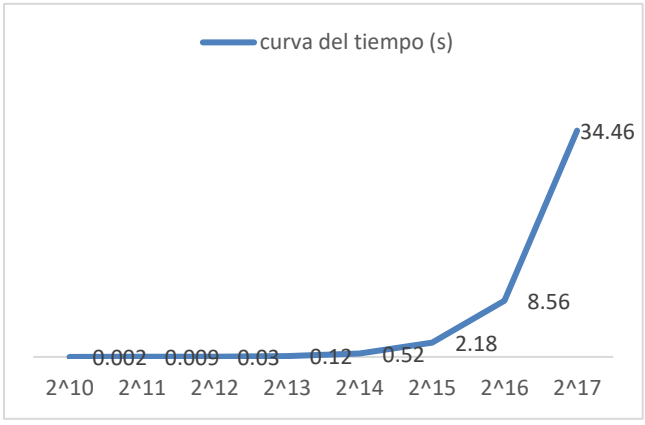


Fig. 3. Análisis de rendimiento en máquinas virtualizadas a cinco nodos, tomados desde el valor 2^10 al 2^17, donde se aprecia una ligera mejora frente a un nodo (Fig. 1).

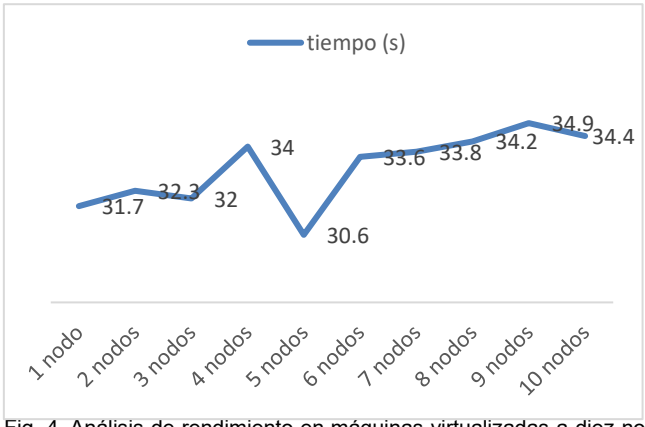


Fig. 4. Análisis de rendimiento en máquinas virtualizadas a diez nodos, tomados desde el valor 2^10 al 2^17, donde se aprecia una caída en el rendimiento comparado a la Fig. 2 y Fig. 3

Una evaluación en la búsqueda de números primos en el límite superior (2^17=131072) da un resultado notable, a la par que representan valores en la media de cinco evaluaciones, que pueden visualizarse en el siguiente gráfico.



Fig. 5. Análisis de rendimiento en máquinas virtualizadas desde un nodo hasta los diez nodos, tomando el mayor de los valores (2^17), donde se aprecia un desarrollo irregular

El contexto que se verá ahora pertenece a la forma tradicional, donde analógicamente veremos datos interpolados a un nodo, cinco nodos y diez nodos, como se muestran a continuación.

C&T Riqchary
Revista de Investigación en Ciencia y Tecnología
ISSN: 2810-8124 (en línea) / ISSN: 2706-543x

Vol. 5 Núm. 1 (2023)
Publicado: 15/10/23
DOI: https://doi.org/10.57166/riqchary/v5.n1.2023.115

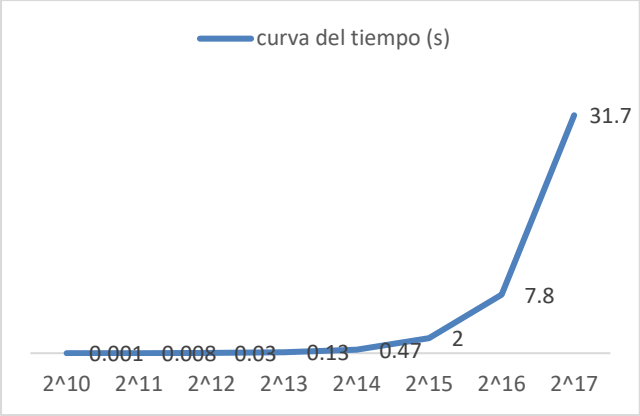


Fig. 6. Análisis de rendimiento en el clúster tradicional a un nodo, tomados de similar manera, desde el valor 2^10 al 2^17,

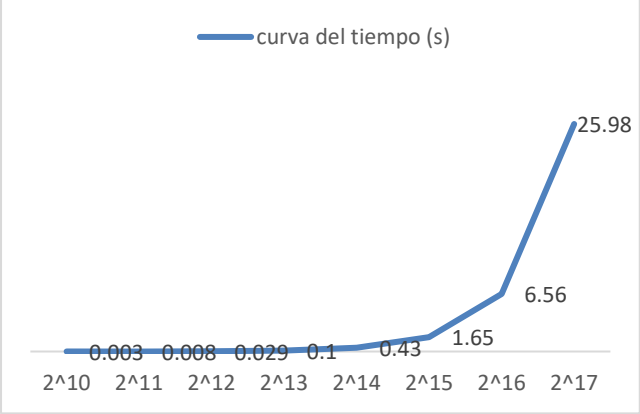


Fig. 7. Análisis de rendimiento en el clúster tradicional a cinco nodos, tomados del intervalo desde 2^10 al 2^17, Notando una mejora frente a los resultados mostrados en la Fig. 6.

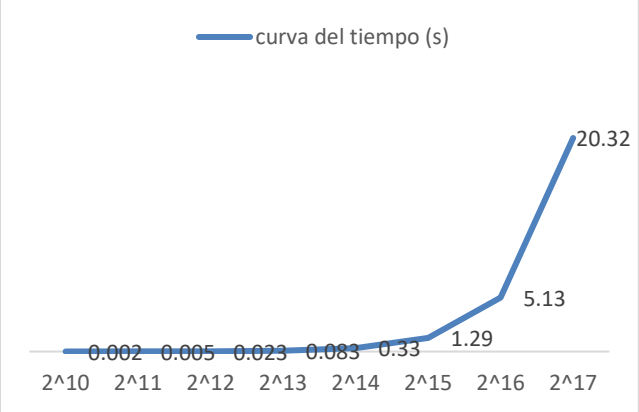


Fig. 8. Análisis de rendimiento en el clúster tradicional a diez nodos, tomados del intervalo desde 2^10 al 2^17, Notando una mejora frente a los resultados mostrados en la Fig. 6 y Fig. 7

La comparativa a lo largo de los diez nodos, trabajándose de uno a uno frente al mayor de los valores de búsqueda, se plasma en la siguiente figura:

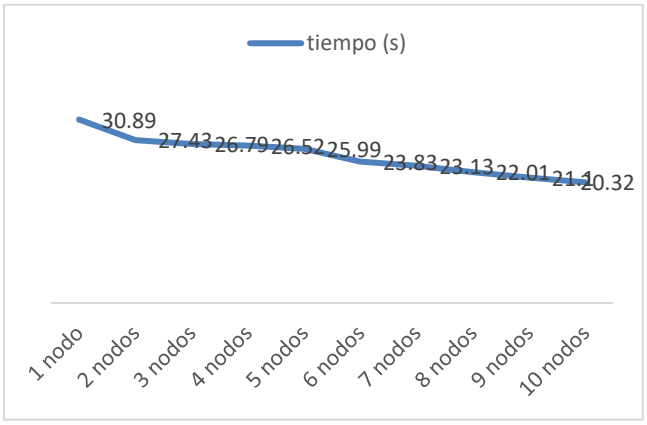


Fig. 9. Análisis de rendimiento en el clúster tradicional desde un nodo hasta los diez nodos con respecto al mayor de los valores de búsqueda, denotando un decremento lógico correcto durante el incremento de nodos computacionales.

4 Discusiones

Se discute la diferencia de equipos, debido al carácter heterogéneo entre ambos contextos (clúster virtualizado y clúster tradicional) por lo que una comparación objetiva y empírica se considerará al comparar en un entorno homogéneo donde las capacidades tanto de hardware y software se equiparan de manera imparcial.

No obstante, se debe rescatar los resultados macro, donde podemos ver una atrofia computacional cuando se sobrecarga el componente de hardware al virtualizar maquinas sobre una única maquina anfitrión, aunque los recursos sean relativamente "potentes", tarde o temprano será evidente la caída operacional-computacional en bloque, resumen que se puede sustentar en la Fig. 5 donde no existe una conclusión clara, más bien una proyección no lógica donde el tiempo se incrementa mientras se incrementa los nodos virtuales (cuando la proporción entre ambas partes debería ser inversa, es decir a más nodos, menor tiempo de ejecución). De otra parte, en el contexto tradicional, se logra apreciar una mejora del 34% de un nodo a diez nodos (el tiempo pasa de 30.89s a 20.32s), siendo el camino más recomendable para implementar clústeres de alto rendimiento.

Los resultados de la investigación convergen con en trabajo de Miguel G Xavier et. al. [5], [18]; puntualizando que el rendimiento nativo de los clústeres (implantación tradicional) tiene mejores resultados que el entono virtualizado.Una desventaja clara en la implementación de clústeres de computadoras es la parte energética, puesto que al tener varios nodos funcionales; también el consumo energético y la facturación por la misma serán directamente proporcionales, por lo que implementar sistemas energéticos como por ejemplo energías fotovoltaicas o eólicas pueden compensar su costo-operatividad, aunque el coste inicial en la implementación de estos últimos no resulta barato pero que a la larga se obtiene buenos resultados coste – beneficio.

5 CONCLUSIONES

Se concluye que la manera de obtener rendimiento computacional al implementar clústeres computacionales de alto rendimiento (HPCC); es realizando una conexión de área local o una red LAN donde se puedan conectar físicamente los nodos, apoyado sobre una red de alta velocidad que soporte un buen ancho de banda (en Gbps), aportando mayor performance al cometido. Además, que el manejo actual de servidores en data center, viendo la caída que manifiestan al sobrecargarse de máquinas virtuales [19]; pueda hacerse un balanceo equitativo (de similar designación de recursos), de tal manera que no se llegue a saturar el performance y el rendimiento. Por otra parte, buscar financiamiento para la adquisición de estos equipos no es nada fácil en una institución gubernamental del estado peruano.

REFERENCIAS

[1] W. Chuliang , L. Minglu , Zhigang Wang and L. Xinda, "Automatic Performance Tuning for the Virtualized Cluster System," 2009.

[2] A. Y. J, J. P. Walters and S. P. Crago, "Supporting High Performance Molecular Dynamics in vittualized clusters using IOMMU, SR-IOV and GPUDirect," 2015.

[3] J. Xunfei, B. Tuguldur and N. Lam, "High performance cluster monitoring system," *Conference: 2018 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, 2018.

[4] Z. Wu, C. Zhangxin, C. C. Douglas and T. Weigin , "High Performance Computing and Applications," *Second international conference HPCA 2009 Shangai, China,* 2010. https://doi.org/10.1007/978-3-642-11842-5

[5] M. G. Xavier, M. V. Neves, F. D. Rossi, T. C. Ferreto, T. Lange and C. A. F. De Rose, "Performance Evaluation of Container-based Virtualization for high performance computing enviroments," 2013. https://doi.org/10.1109/PDP.2013.41

[6] T. Thandar, C. Sung-Do and S. P. Jong , "Availability Modeling and Analysis on Virtualized Clustering with rejuvenation," *international journal of computer science and network security,* vol. 8, no. 9, 2009.

[7] R. M. Yoo, C. J. Hughes, L. Konrad and R. Rajwar, "Performance Analysis of High Performance Computing Applications on the Amazon Web Services Cloud," 2013.

[8] C. Paidamoyo, L. Chunyuan, M. Nikhil, C. Lawrence and R. Henao, "Survival Cluster Analysis," 2020.

[9] R. Hernández, C. Fernández and P. Baptista, Metodología de la Investigación, Mexico: McGraw-Hill, 2014.

[10] U. Deshpande, B. Wang, S. Haque, M. Hines and K. Golapan, "MemX: Virtualization of Cluster-wide Memory," 2010. https://doi.org/10.1109/ICPP.2010.74

[11] D. Hasenkamp, A. Sim, M. Wehner and K. Wu, "Finding Tropical Cyclones on a Cloud Computing Cluster: using parallel virtualization for large-scale climate simulation analysis," 2010. https://doi.org/10.1109/CloudCom.2010.68

[12] B. Tierney, W. Johnston, B. Crowley, G. Hoo, C. Brooks and D. Gunter, "The NetLogger Methodology for High Performance Distributed Systems Performance Analysis," 1998. https://doi.org/10.2172/764331

[13] DistroWatch, "https://distrowatch.com/," [Online]. Available: https://distrowatch.com/table.php?distribution=pelicanhpc. [Accessed 20 septiembre 2023].

[14] DELL Technologies, "https://www.dell.com," [Online]. Available: https://www.dell.com/support/manuals/es-pe/alienware-17-laptop/alienware-17-r4-setupandspecifications/especificaciones?guid=guid-1eae1626-6999-4d20-a522-a574d6b4761c&lang=es-mx. [Accessed 18 septiembre 2023].

[15] FS community, "https://community.fs.com," [Online]. Available: https://community.fs.com/es/blog/t568a-vs-t568b-difference-between-straight-through-and-crossover-cable.html. [Accessed 20 septiembre 2023].

[16] M. C. Cowgill, R. J. Harvey and I. T. Watson, "A Genetic Algorithm Approach to cluster Analysis," 1999. https://doi.org/10.1016/S0898-1221(99)00090-5

[17] Keepcoding TechSchool, "https://keepcoding.io," [Online]. Available: https://keepcoding.io/blog/que-es-live-cd/. [Accessed 20 septiembre 2023].

[18] M. G. Xavier, M. V. Neves and C. A. F. De Rose , "A Performance Comparison of Container-based Virtualization Systems for Mapreduce Clusters," 2014. https://doi.org/10.1109/PDP.2014.78

[19] A. J. Younge, R. Herschel, J. T. Brown, G. V. Laszewski, J. Qiu and G. C. Fox, "Analysis of Virtualization Technologies for High Performance Computing Enviroments," 2011. https://doi.org/10.1109/CLOUD.2011.29

BIOGRAFÍAS

John Abraham Aguirre Carrasco, Ingeniero Informático y Sistemas egresado de la Universidad Nacional Micaela Bastidas de Apurímac.

Luis Miguel Alfaro Chirinos, Ingeniero Informático y Sistemas, maestro en Gestión Pública de la Universidad Cesar Vallejo, docente adscrito al departamento de Ingeniería, Escuela Académica Profesional de Ingeniera informática y Sistemas de la Universidad Nacional Micaela Bastidas de Apurímac.

Alejandrina Huaylla Quispe, Ingeniero de Sistemas e Informática, maestra en Gestión Pública de la Universidad Cesar Vallejo, Egresada de la maestría en Proyectos de Inversión de la Universidad Tecnológica de los Andes, docente contratado adscrito al departamento de Ingeniería, Escuela Académica Profesional de Ingeniera informática y Sistemas de la Universidad Nacional Micaela Bastidas de Apurímac.

C&T Riqchary
Revista de Investigación en Ciencia y Tecnología
ISSN: 2810-8124 (en línea) / ISSN: 2706-543x

Vol. 5 Núm. 1 (2023)
Publicado: 15/10/23
DOI: https://doi.org/10.57166/riqchary/v5.n1.2023.117

La mejor opción para predecir el precio máximo de las acciones de Intel Corporation: ¿Árbol de regresión o Regresión Lineal Múltiple?

The best choice for predicting Intel Corporation's peak stock price: Regression Tree or Multiple Linear Regression?

Edgar Eloy Carpio Vargas [A], Alicia Roxana Mayda Huanca [B]. D. German Rafael Espinoza Rivas [C], Ecler Mamani Vilca[D] y [E]Betsabe M. Ccolqque Ruiz

ORCID:0000-0001-6457-4597[A], ORCID:0009-0006-4618-8872[B], ORCID: 0000-0002-6547-7761[C], ORCID: 0000-0002-5205-3660[D] ORCID 000-0002-3089-8750[E]

(Recepción: 15/04/2023 y aceptación 08/10/2023)

Resumen— El objetivo del estudio fue comparar el rendimiento del Árbol de regresión frente al Modelo de regresión lineal múltiple en relación al precio de apertura y volumen de ventas diarias de las acciones de Intel Corporation. Se llevó a cabo una investigación descriptiva correlacional de tipo no experimental con diseño transversal, utilizando una muestra por conveniencia. La muestra consistió en 410 registros recopilados desde mayo de 2018 hasta octubre de 2019, obtenidos a través de revisión documental. Los resultados obtenidos mostraron que el Árbol de regresión estableció que la variable más significativa para explicar el precio máximo de las acciones fue el precio de apertura, descartando la variable de volumen. El Error Medio Cuadrático obtenido fue de 1.4480 dólares. Por otro lado, el Modelo de regresión lineal múltiple, utilizando la técnica de eliminación de datos atípicos, presentó un Error Estándar Residual de 0.2257 dólares. En conclusión, se determinó que el modelo más adecuado para predecir el precio máximo de las acciones de Intel Corporation es el Modelo de Regresión Lineal Múltiple con eliminación de puntos atípicos.

Palabras clave: Árboles de regresión, regresión lineal multiple

Abstract— The objective of the study was to compare the performance of the Regression Tree versus the Multiple Linear Regression Model in relation to the opening price and daily sales volume of Intel Corporation shares. A descriptive correlational non-experimental correlational research with cross-sectional design was conducted using a convenience sample. The sample consisted of 410 records collected from May 2018 to October 2019, obtained through documentary review. The results obtained showed that the Regression Tree established that the most significant variable to explain the maximum stock price was the opening price, discarding the volume variable. The Mean Squared Error obtained was $1.4480. On the other hand, the multiple linear regression model, using the outlier elimination technique, presented a Residual Standard Error of 0.2257 dollars. In conclusion, it was determined that the most adequate model to predict the maximum price of Intel Corporation shares is the Multiple Linear Regression Model with the elimination of outlier points.

Keywords: Multiple Linear Regression, Regression Trees

A. *Edgar Eloy Carpio Vargas Departamento de Estadística e Informática – UNA Puno. ecarpio@unap.edu.pe*
B. *Alicia Roxana Mayda Huanca, Escuela de Ingeniería Estadística e Informática – UNA Puno: amaydana@epg.unap.edu.pe*
C. *German Rafael Espinoza Rivas, trabaja en el Departamento de Ingeniería – UNAMBA Perú: gespinoza@unamba.edu.pe*
D. *Ecler Mamani Vilca Departamento de Ingeniería informática y Sistemas UNAMBA, Perú : eclervirtual@unamba.edu.pe*
E. *Betsabe Milagros Ccolqque Ruiz, Departamento de Ingeniería informática y Sistemas UNAMBA bccolque@unamba.edu.pe*

1 INTRODUCCIÓN

La toma de decisiones juega un papel fundamental en cualquier organización o empresa. Por tanto, resulta imperativo buscar técnicas y metodologías estadísticas precisas que nos ayuden a resolver problemas de incertidumbre que surgen al tomar decisiones. Hoy en día, se generan grandes volúmenes de datos a diario, los cuales necesitan ser tratados a través de metodologías que sean capaces de generar información útil para investigar, predecir o tomar decisiones. Además, gracias al avance tecnológico, no solo podemos almacenar la información, sino también procesarla y generar conocimiento. [1].

Intel (Integrated Electronics Corporation) es el mayor fabricante de circuitos integrados del mundo según su cifra de negocio anual. La compañía estadounidense es la creadora de la serie de procesadores x86, los procesadores más comúnmente encontrados en la mayoría de las computadoras personales [2]. Sin embargo, para los accionistas e inversionistas de dicha empresa es muy riesgoso vender acciones sin saber si esta le generará ganancias o pérdidas de dinero al vender su bien, en este sentido, se plantea predecir el Precio máximo que una acción puede tomar durante un determinado tiempo, de esta forma los accionistas e inversionistas tendrán una mejor aproximación, sobre cuanto sería el Precio máximo que alcanzaría una acción. Si bien hoy en día, existe una gran diversidad de técnicas y metodologías estadísticas para predecir el costo de las acciones de las empresas en el mercado, se decidió realizar este estudio comparativo y nos planteamos la siguiente interrogante: ¿Cuál es la mejor opción para predecir el precio máximo de las acciones de Intel Corporation: Árboles de regresión o Regresión Lineal Múltiple

Los trabajos revisados fueron:

Gimenéz [3] indicaron que hacer pronósticos precisos de índices bursátiles es prácticamente imposible. Incluso teniendo modelos que parecieran dar buenos resultados, como los que elegimos finalmente, éstos pueden presentar problemas en la realidad debido al sobreajuste o “overfitting”.

Alvarez [4] indicó que la metodología Ítem a Ítem planteada por Amazon resulto útil para predecir el orden de preferencia de los ligandos que se unen a un determinado blanco sin embargo los resultados predictivos no resultaron suficientemente buenos. Si bien la medida aprendió la tendencia para cada proteína esta no pudo generalizar la información.

En cusco Candia [5] comprobó que, según los resultados encontrados, el algoritmo de árboles de decisión “Random Forest”, fue el algoritmo que tuvo el mejor performance para la predicción del rendimiento académico de los ingresantes en los primeros semestres a la UNSAAC con un 69% de predicción, el segundo algoritmo con mejor performance fue algoritmo de Regresión Logística con un 68% para el presente caso de estudio.

1.1 Árboles de Regresión

Los árboles de clasificación se emplean para asignar sujetos a las clases de una variable dependiente a partir de sus mediciones en uno o más predictores. Modernamente, los árboles de clasificación (AC) constituyen uno de los recursos instrumentales básicos de la llamada "minería de datos" [6]. Heredia [7] define los árboles de clasificación y regresión, conocidos como algoritmos CART (del inglés classification and regression trees), constituyen una aproximación multivariada no paramétrica que permite identificar y dimensionar las variables X de mayor impacto en una variable Y. Los modelos CART particionan los datos en forma recursiva de modo tal de conformar subconjuntos cada vez más homogéneos en base a criterios de partición de las variables explicativas. Cada árbol se obtiene a partir de la clasificación de un nodo parental o raíz que contiene la totalidad de los datos, mediante un algoritmo de partición especificado en función de un criterio de partición referido al tamaño del nodo formado o a la variabilidad contenida en los datos del nodo, también el algoritmo CART. Un árbol de regresión o de clasificación consiste en un conjunto de reglas determinadas por un procedimiento de ajuste mediante particiones binarias recursivas.

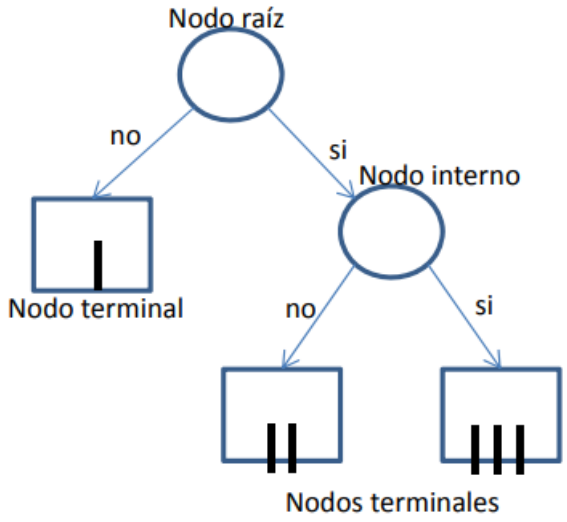


Fig. 1. Elementos del Árbol de regresión

En la fig. 1, el árbol tiene tres niveles de nodos, el primer nivel tiene un único nodo en la cima llamado nodo raíz. Un nodo interno en el segundo nivel, y tres nodos terminales que están respectivamente en el segundo y tercer nivel. El nodo raíz y el nodo interno son particionados cada uno en dos nodos en el siguiente nivel los cuales son llamados nodos hijos (o ramas) izquierdo y derecho.

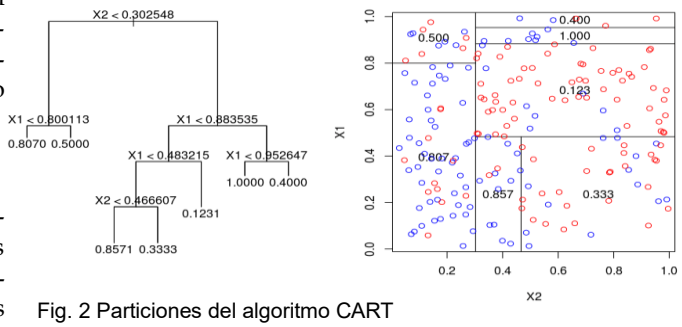


Fig. 2 Particiones del algoritmo CART

los árboles de regresión permiten predecir variables respuestas continuas, las observaciones según umbrales de las variables regresoras, considerando la suma de cuadrados de la respuesta como medida de heterogeneidad dentro de cada partición. Como en el árbol de clasificación, la medida de heterogeneidad entre las observaciones que quedan dentro de un nodo debe ser menor que la calculada entre las observaciones de distintos nodos.

C&T Riqchary
Revista de Investigación en Ciencia y Tecnología
ISSN: 2810-8124 (en línea) / ISSN: 2706-543x

Vol. 5 Núm. 1 (2023)
Publicado: 15/10/23
DOI: https://doi.org/10.57166/riqchary/v5.n1.2023.117

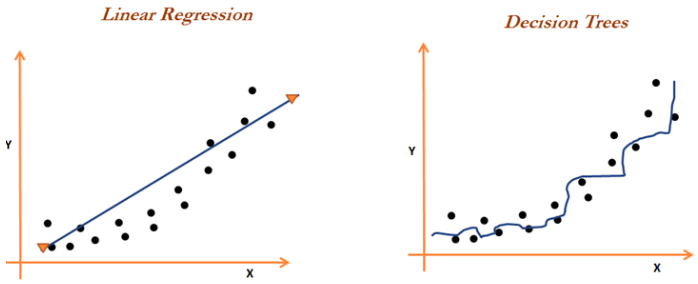


Fig. 3. Regresión lineal y el Árbol de decisión [7]

Podado del árbol, con la finalidad de reducir la varianza del modelo y así disminuir el test error, se somete al árbol a un proceso de poda, intenta encontrar el árbol más sencillo (menor tamaño) que consigue explicar las observaciones, y así para identificar el valor óptimo de penalización α. Por defecto, esta función emplea la desviación para guiar el proceso de poda y manteniendo la estructura robusta que consigue un test error bajo. La selección del sub-árbol optimo puede hacerse mediante la validación cruzada, sin embargo, dado que los árboles se crecen lo máximo posible (tienen muchos nodos terminales) no suele ser viable estimar el test error de todas las posibles sub-estructura que se pueden generan [8].

1.3 Algoritmo para crear un árbol de regresión con pruning

Paso 1:

Se emplea recursive binary splitting para crear un árbol grande y complejo (To) empleando los datos de training y reduciendo al máximo posible las condiciones de parada. Normalmente se emplea como única condición de parada el número mínimo de observaciones por nodo terminal.

Recursive binary splitting

El objetivo del método recursive binary splitting es encontrar en cada iteración el predictor Xj y el punto de corte (umbral) s tal que, si se distribuyen las observaciones en las regiones $\{X|Xj<s\}$ y $\{X|Xj\geq s\}$, se consigue la mayor reducción posible en el RSS. El algoritmo seguido es [9]:

a. El proceso se inicia en lo más alto del árbol, donde todas las observaciones pertenecen a la misma región.

b. Se identifican todos los posibles puntos de corte (umbrales) s para cada uno de los predictores ($X1$,$X2$,…,Xp). En el caso de predictores cualitativos, los posibles puntos de corte son cada uno de sus niveles. Para predictores continuos, se ordenan de menor a mayor sus valores, el punto intermedio entre cada par de valores se emplea como punto de corte.

c. Se calcula el RSS total que se consigue con cada posible división identificada en el paso b en ecuación (1).

$$\sum_{i:x_i \epsilon R_1(j,s)} \left(y_i - \hat{y}_{R_1}\right)^2 + \sum_{i:x_i \epsilon R_2(j,s)} \left(y_i - \hat{y}_{R_2}\right)^2 \quad (1)$$

donde el primer término es el RSS de la región 1 y el segundo término es el RSS de la región 2, siendo cada una de las regiones el resultado de separar las observaciones acordes al predictor j y valor s.

d. Se selecciona el predictor Xj y el punto de corte S que resulta en el menor RSS total, es decir, que da lugar a las divisiones más homogéneas posibles. Si existen dos o más divisiones que consiguen la misma mejora, la elección entre ellas es aleatoria.

e. Se repiten de forma iterativa los pasos 1 a 4 para cada una de las regiones que se han creado en la iteración anterior hasta que se alcanza alguna norma de stop. Algunas de las más empleadas son: que ninguna región contenga un mínimo de n observaciones, que el árbol tenga un máximo de nodos terminales o que la incorporación del nodo reduzca el error en al menos un % mínimo.

Paso 2:

Se aplica el cost complexity pruning al árbol To para obtener el mejor sub-árbol en función de α. Es decir, se obtiene el mejor sub-árbol para un rango de valores de α.

Cost complexity pruning es un método de penalización de tipo Loss + Penalty, similar al empleado en ridge regression o lasso. En este caso, se busca el sub-árbol T que minimiza la ecuación:

$$\sum_{j=1}^{|T|} \sum_{i \in R_j} \left(y_i - \hat{y}_{R_j}\right)^2 + \alpha |T| \quad (2)$$

donde |T| es el número de nodos terminales del árbol.

El primer término de la ecuación se corresponde con el sumatorio total de los residuos cuadrados RSS. Por definición, cuantos más nodos terminales tenga el modelo menor será esta parte de la ecuación. El segundo término es la restricción, que penaliza al modelo en función del número de nodos terminales (a mayor número, mayor penalización). El grado de penalización se determina mediante el tuning parameter α. Cuando α=0, la penalización es nula y el árbol resultante es equivalente al árbol original. A medida que se incrementa α la penalización es mayor y, como consecuencia, los árboles resultantes son de menor tamaño. El valor óptimo de α puede identificarse mediante cross validation.

Paso 3:

Identificación del valor óptimo de α mediante k-cross-validation. Se divide el training data set en K grupos. Para k=1, …,

k=K:

- Repetir pasos 1 y 2 empleando todas las observaciones excepto las del grupo ki.
- Evaluar el mean squared error para el rango de valores de α empleando el grupo ki.
- Obtener el promedio de los K mean squared error calculados para cada valor α.

Paso 4:
Seleccionar el sub-árbol del paso 2 que se corresponde con el valor α que ha conseguido el menor cross-validation mean squared error en el paso 3.

1.2 La regresión lineal múltiple

Un modelo de regresión lineal múltiple es un modelo estadístico versátil para evaluar las relaciones entre un destino continuo y los predictores **[10]**. La permite generar un modelo lineal en el que el valor de la variable dependiente o respuesta Y se determina a partir de un conjunto de variables independientes llamadas predictores X ´s. Es una extensión de la regresión lineal simple, por lo que es fundamental comprender esta última. Los modelos de regresión múltiple pueden emplearse para predecir el valor de la variable dependiente o para evaluar la influencia que tienen los predictores sobre ella (esto último se debe que analizar con cautela para no malinterpretar causa-efecto) **[11]**.

Los modelos lineales múltiples siguen la siguiente ecuación:

$$Y_i = (\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_{ni}) + e_i$$

$$i = 1,2,3,\ldots,n \qquad (3)$$

$\boldsymbol{Y_i}$**:** Es la variable dependiente o respuesta.
$\boldsymbol{\beta_0}$**:** Es la ordenada en el origen, el valor de la variable dependiente Y cuanto todos predictores son cero.
$\boldsymbol{\beta_i}$**:** Es el efecto promedio que tiene el incremento en una unidad de la variable predictora
$\boldsymbol{Xi}$: Es la variable independiente o explicativas
$\boldsymbol{ei}$**:** Es el residuo o error, la diferencia entre el valor observado y el estimado por el modelo.

Condiciones para la Regresión Lineal Múltiple:

Los modelos de correlación lineal múltiple requieren de las mismas condiciones que los modelos lineales simples más otras adicionales, estas son [11]:

a. **No colinialidad o multicolinialidad:** En los modelos lineales múltiples los predictores deben ser independientes, no debe haber colinialidad entre ellos. La colinialidad ocurre cuando un predictor está linealmente relacionado con uno o varios de los otros predictores del modelo o cuando es la combinación lineal de otros predictores. Como consecuencia de la colinialidad no se puede identificar de forma precisa el efecto individual que tiene cada una de las variables colineales sobre la variable respuesta, lo que se traduce en un incremento de la varianza de los coeficientes de regresión estimados hasta el punto que resulta prácticamente imposible establecer su significancia estadística. Además, pequeños cambios en los datos provocan grandes cambios en las estimaciones de los coeficientes [12].

b. **Normalidad:** Los residuos deben seguir una distribución normal.

c. **Homocedasticidad:** La varianza del error debe ser constante para todos los valores del predictor.
d. **Independencia:** Los errores deben ser independientes entre sí.

Es importante tener en cuenta que estas condiciones son necesarias pero no suficientes para garantizar un buen ajuste del modelo.

Teniendo la variable dependiente (Y), es el precio máximo de venta que la acción ha logrado durante la negociación regular por un determinado periodo de tiempo; durante el día, para esta variable la unidad de medida es el dólar ($).

Variables independientes (X's)
X_1: Precio de apertura en el cual la acción es abierta al mercado (esta varia a lo largo del día). Al igual que la variable anterior, la unidad de medida es el dólar ($).
X_2: Volumen de ventas de un determinado periodo de tiempo de la empresa Intel. Para esta variable la unidad de medida es la cantidad de ventas que se registró, durante el día.

Torres [13] indica que, las fuentes de información, son todos aquellos medios de los cuales procede la información, que satisfacen las necesidades de conocimiento de una situación o problema presentado, que posteriormente será utilizado para lograr los objetivos esperados, para el estudio la obtención de datos fue mediante el internet, los cuales han sido obtenidos de la página https://www.nasdaq.com/es/market-activity/stocks/intc , "Nasdaq" , la técnica utilizada fue documental y el instrumento registro de la Base de Datos.

2 MATERIALES Y MÉTODOS

El diseño de investigación es cuantitativo de tipo descriptivo correlacional con diseño no experimental [14], debido al modo de obtencion de las variables. Por otro lado, se puede indicar que los datos son de corte transversal y fueron obtenidos en un corte de tiempo determinado.

La población para el estudio estuvo conformada por las acciones de la empresa Intel a través del tiempo y la muestra se

Revista de Investigación en Ciencia y Tecnología
ISSN: 2810-8124 (en línea) / ISSN: 2706-543x

Vol. 5 Núm. 1 (2023)
Publicado: 15/10/23
DOI: https://doi.org/10.57166/riqchary/v5.n1.2023.117

define como no probabilística obtenida por conveniencia durante el periodo de estudio, se consideró a todas las acciones de la empresa Intel Corporation las mismas que se vendieron desde mayo del 2018 hasta octubre del 2019, se obtuvieron , 410 registros.
Las técnicas estadísticas aplicadas para el procesamiento y obtención de modelos luego de la recolección de datos fueron: la regresión lineal múltiple y los árboles de regresión, utilizando la metodología machine learnign.

3 Resultados

Para obtener el árbol de regresión se particiono la data en datos de entrenamiento y validación, quedando 80% para entrenamiento y 20% para validación, siendo el total de la muestra 410 registros, de ellos, 328 para entrenamiento y 82 para validación.

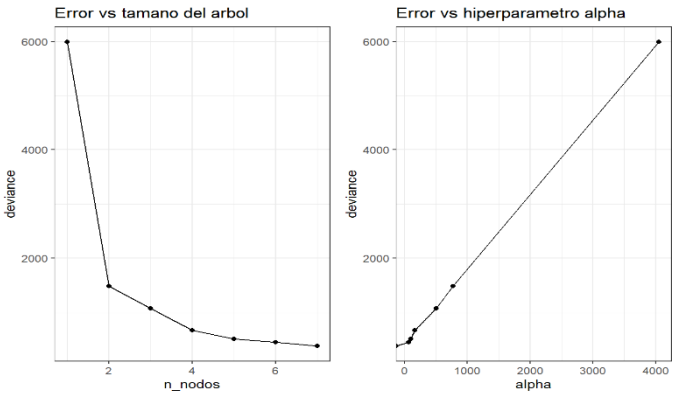


Figura 4. Número de nodos para el entrenamiento

El número de nodos para establecer el Árbol de regresión, en 7, que nos garantiza la no presencia de mucha variación en cuanto al error, observando la desviación deben ser muy pequeños. También la finura nos muestra que podríamos haber tomado 4 nodos.

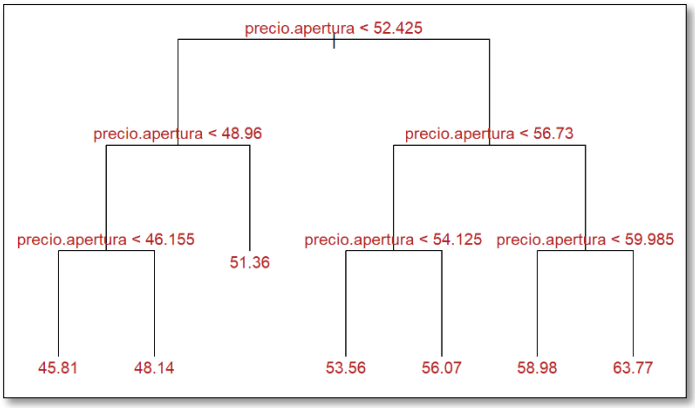


Fig. 5. Árbol de Regresión con 7 nodos terminales.

Se utilizó el software R, que presenta una solución para el Precio de apertura de las acciones de Intel.

Interpretación del modelo.
Intel alcanza el siguiente el precio máximo de acciones:

- $63.77 cuando el precio de apertura es mayor a $59.985, mayor a $56.73 y mayor a $52.425.
- $58.98 cuando el precio de apertura es menor a $59.985, mayor a $56.73 y mayor a $52.425
- $56.07 cuando el precio de apertura es mayor a $54.125, menor a $56.73 y mayor a $52.425
- $53.56 cuando el precio de apertura es menor a $54.125, menor a $56.73 y mayor a $52.425
- $51.36 cuando el precio de apertura es mayor a $48.96 y menor a $52.425
- $48.14 cuando el precio de apertura es mayor a $46.155, menor a $48.96 y menor a $52.425
- $45.81 cuando el precio de apertura es menor a $46.155, menor a $48.96 y menor a $52.425

La fig. 6. muestra a continuación corresponde al algoritmo podado.

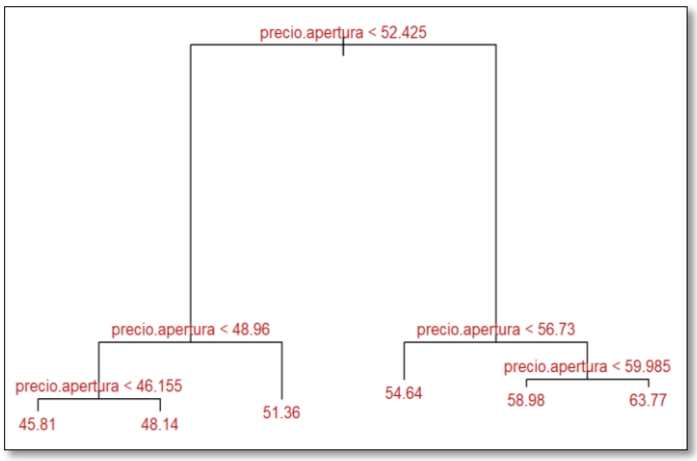


Fig. 6. Representación del Árbol de regresión podado con el algoritmo pruning.

Tabla 1. Comparación de errores para el Árbol de regresión inicial y podado.

Comparación	Árbol de Regresión	
	Sin podar	Podado
MSE	1.448015 $	1.61 $

Se observa en la tabla 1, que el valor de MSE sin podar (1.44) es menor que el árbol podado(1.61), por lo tanto, elegimos la solución sin podar.

Para el modelo de regresión lineal múltiple. A partir de los 410 de registros recopilados a partir del Nasdaq, R nos muestra los coeficientes de β_0, β_1 y β_2 para el primer modelo y las estimaciones queda así:

$$Y = 0.1454 + 1.00100X_1 + 0.00000001492X_2 \quad (4)$$

Las dos variables independientes precio de apertura y volumen resultan ser significativos p<α (0.05) excepto el intercepto. El modelo en su conjunto observando la prueba F es significativa y la bondad de ajuste es bastante alta $R^2 = 0.9934$. lo que indica que, es capaz de explicar el 99,34% de la variabilidad Observada en el precio máximo. El p-valué del modelo es significativo 2.2e-16 por lo que, se puede aceptar que el modelo no es por azar, al menos uno de los coeficientes parciales es distinto de 0.

Comprobación de supuestos:

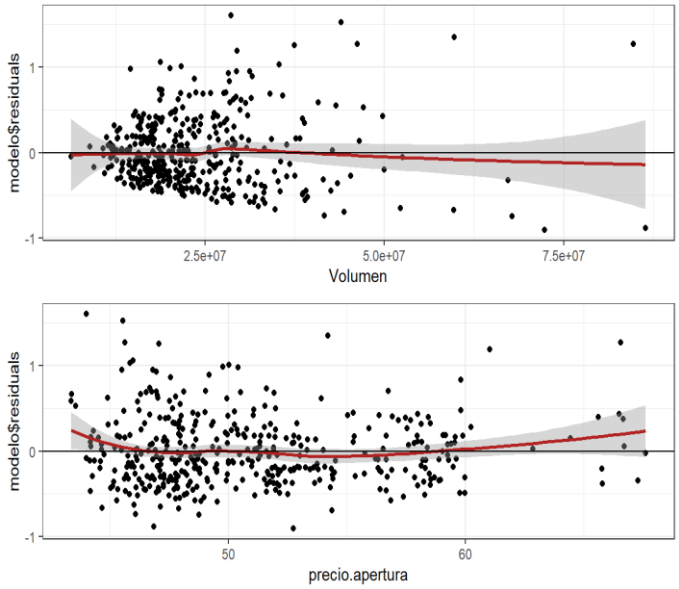


Fig. 7. Linealidad de las variables precio apertura y volumen de las acciones Intel Corporation, se observa el supuesto de linealidad de las variables independientes X's, que evidentemente, al estar la línea dentro de los intervalos se observa linealidad para el precio de apertura y el volumen de ventas de las acciones del año 2019.

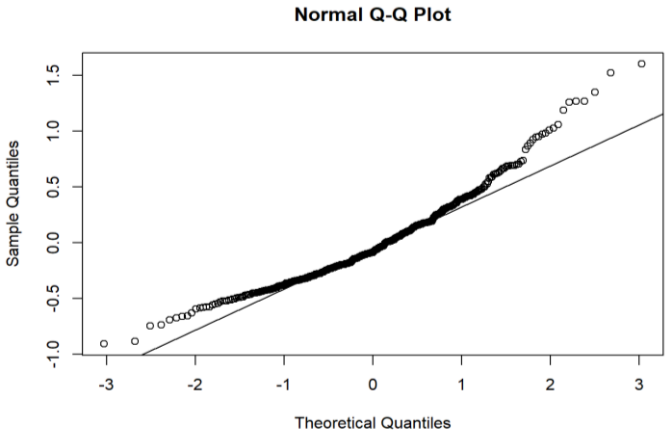


Fig. 8, se observa que los datos presentan normalidad de errores, en el apartado siguiente se realiza el supuesto de normalidad.

Mediante la prueba de Shapiro Wilks, se concluye por p(0.0000000002096)<α(0.05), en este caso se acepta la hipótesis Ha, es decir, no existe normalidad en los errores, y se rechaza la hipótesis H_0.

La prueba de homocedasticidad (test de Breusch Pagan), se concluye, r p(0.00000000000000022)<α(0.05), se acepta la hipótesis Ha, es decir, no existe homocedasticidad en las variables.

Prueba de autocorrelación (Durwin Watson), se concluye p(0.176)>α(0.059), entonces, existe autocorrelación entre las variables, es decir se acepta H0, y se rechaza Ha.
En resumen, se tiene los siguientes resultados: no existe cumplimiento de linealidad, normalidad y Homocedasticidad al contrario de autocorrelación que si cumple.
A pesar del incumplimiento de supuestos, el modelo arroja un R^2 alto 0.9934 y el ANVA significativo.
A continuación, se realiza un análisis de puntos Outliers.

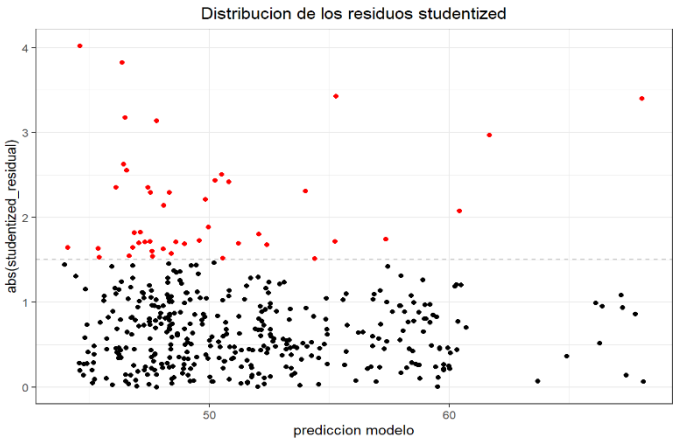


Fig. 9, Se aprecia muchos valores atípicos. Usaremos un corte de 1.5σ. los cuales afectan potencialmente a la estimación de los parámetros, por ende, se realizó la limpieza de datos, con el método de eliminación de puntos Outliers

Obtención del segundo modelo de RLM, para este segundo modelo, luego de haber utilizado el método de eliminación de puntos Outlier queda así:

$$Y = 0.01750 + 1.00300X_1 + 0.00000001199X_2 \quad (5)$$

Las dos variables independientes precio de apertura y volumen resultan ser significativas p<α (0.05) excepto el intercepto. El modelo en su conjunto observando la prueba F es significativa y la bondad de ajuste es bastante alta R^2=0.9968. El error estándar residual este disminuye en comparación al modelo anterior (0.4072 frente a 0.2849 del modelo 2).

Comprobación del supuesto de normalidad:
Mediante la prueba de Kolmogorov Smirnov se determinó, p=0.02752<α=0.05, en este caso se acepta la hipótesis Ha, es decir, no existe normalidad en los errores, y se rechaza la hipótesis H_0.

C&T Riqchary
Revista de Investigación en Ciencia y Tecnología
ISSN: 2810-8124 (en línea) / ISSN: 2706-543x

Vol. 5 Núm. 1 (2023)
Publicado: 15/10/23
DOI: https://doi.org/10.57166/riqchary/v5.n1.2023.117

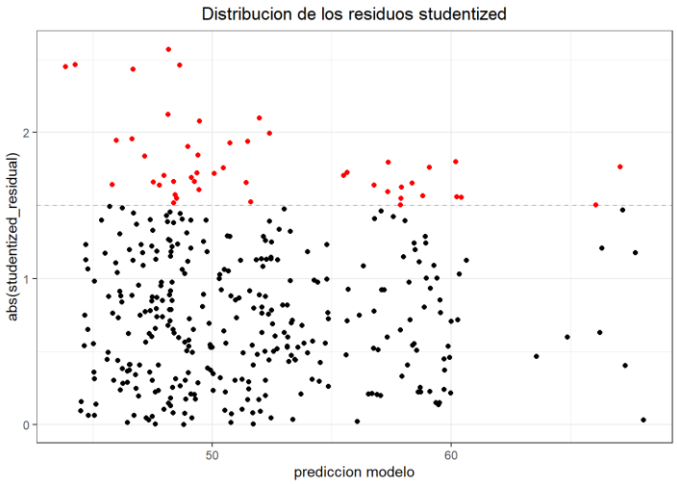


Fig. 10. Gráfico de datos Estandarizados mayores a 1.5 sigma, los valores son excluidos para obtener mejores resultados.

A partir de la limpieza de datos, el modelo de regresión lineal múltiple, queda definido como:

$$Y = 0.02856 + 1.003X_1 + 0.000000009405X_2 \quad (6)$$

Las dos variables independientes precio de apertura y volumen resultan ser significativas p<α (0.05) excepto el intercepto. El modelo en su conjunto observando la prueba F es significativa y la bondad de ajuste es bastante alta R^2=0.9979. El error estándar residual este disminuye en comparación al modelo anterior 0.2849 frente a 0.2257 del modelo 3.

Interpretación del modelo:

$$Y = 0.02856 + 1.003PrecioApert. + 0.000000009405Volum \quad (7)$$

β_1: Por cada unidad monetaria del precio de apertura en el cual la acción es abierta al mercado, el precio máximo de venta en un día, se ve incrementada en 1.003.
β_2: Por cada venta que se realiza el precio máximo de venta de las acciones de Intel, incrementa en 0.000000009405.

Comprobación de los supuestos del tercer modelo de regresión (7).

Linealidad, normalidad, homocedasticidad, autocorrelación análisis de inflación de varianza, cumple todos.

Tabla 2. Resumen de errores de la Regresión lineal múltiple.

Resumen	Regresión Lineal Múltiple		
	Primer modelo	Segundo modelo	Tercer modelo
R^2	0.9934	0.9968	0.9979
Estándar error residual	0.4072	0.2849	0.2257

Determinación de la mejor opción entre Árbol de regresión o Regresión Lineal Múltiple:

Tabla 3, muestra el resumen de los errores MSE y R^2 de Regresión lineal múltiple con la eliminación de puntos atípicos Outlier, obteniendo de esta forma un R^2 muy bueno 0.9979 lo cual indica que el 0.99% de los datos se ajustan al modelo y el Error estándar residual con $ 0.2257, seguidamente se muestra el algoritmo Árbol de regresión, para la validación sin podado se encontró un MSE de $1.4480; y finalmente se concluye que la Regresión lineal múltiple tiene menor error frente al Árbol de regresión.

Tabla 3. Resumen de errores para la Regresión lineal y Árbol de regresión.

Algoritmo	Regresión lineal múltiple	Árbol de regresión
	Tercer Modelo	Validación sin podado
R^2	0.9934	_____
Error Estándar/MSE	$ 0.2257	$ 1.448015

4 Discusiones

Lizares [1] comparo la Regresión logística binaria y Arboles de clasificación (CHAID) para evaluar el rendimiento académico. Concluye que, Según la evaluación de la clasificación de los modelos optamos por la Técnica de Árboles de clasificación, siendo la más óptima por tener mayor Sensibilidad=77,6% AUC=90,1%, Gini =80,2% y Kappa=0,589. A diferencia de esta tesis, en la que se empleó el Árbol de regresión por la presencia de variables cuantitativas, se encuentro mayor MSE $1.448015 comparado al Error Estándar Residual de la RLM.

También Espinosa [15] quien concluye que realizando un análisis técnico de los datos históricos del valor de las acciones durante un plazo determinado, es posible predecir el cambio de precio de una acción, con la finalidad de ayudar al inversor al momento de tomar la decisión de compraventa trading de sus acciones, utilizando para esto regresión múltiple, lo que concuerdan con los resultados del presente trabajo de investigación al tener la regresión lineal múltiple un Error Estándar de 0.2257, se confirma que el modelo de Regresión múltiple es buen candidato al momento de enfrentarse a un problema de predicción de variables continuas.

Confrontando con Sepúlveda [16] que, en su estudio comparo árbol regresión CART y Regresión Lineal, encontró que, el error predictivo de la regresión lineal siempre es menor que el del CART. Aconteció algo similar en esta tesis, en la que adicionalmente se emplearon técnicas estadísticas como la eliminación de puntos Outlier para la Regresión lineal múltiple

y la poda en el árbol de regresión, sin embargo, se concluyó que el modelo de regresión lineal múltiple tiene menor error frente el árbol de regresión para predecir el precio máximo en función al precio de apertura y el volumen de ventas de Intel.

5 CONCLUSIONES

Después de varias etapas desarrolladas finalmente se concluye que el mejor modelo para predecir el precio máximo de acciones de Intel es el modelo de regresión lineal Múltiple con eliminación de Outliers, debido a que este presenta menor error frente a un árbol de regresión. Este modelo ayudara a reducir la incertidumbre en los accionista o Inversionistas de la empresa Intel Corporation.

Referencias

[1] M. Lizares, Comparación de los modelos de clasificación: regresión logística y árboles de clasificación para evaluar el rendimiento académico., 2017.

[2] A. García De Mendoza Ortega, "Intel," Universidad Jesuista de Guadalajara, Guadalajara, 2022.

[3] R. Giménez Fernández and . P. Zamorano, "Modelos predictivos de índices bursátiles relevantes para la economía chilena," Universidad de Chile, Santiago, 2014.

[4] M. L. Alvarez, "Predicción de afinidad de uni´on de ligandos en protéinas," Universidad de Buenos Aires, Buenos Aires, 2016.

[5] D. I. Candia Oviedo, "Predicción del rendimiento académico de los estudiantes de la UNSAAC a partir de sus datos de ingreso utilizando algoritmos de aprendizaje automático," Universidad Nacional San Antonio Abad del Cusco, Cusci, 2019.

[6] J. Bacallao Gallestey and J. M. Parap, "Árboles de regresión y otras opciones metodológicas aplicadas a la predicción del rendimiento académico," *Educación Médica Superior*, 2004.

[7] S. Rosales Heredia, C. Bruno and . M. Balzarini, "Identificación de relaciones entre rendimientos y variables ambientales vía árboles de clasificación y regresión (CART)," *Interciencia*, vol. 35, no. 12, pp. 876-882, 2010.

[8] S. C. S. Pineda, Comparación de árboles de regresión y clasificación regresion logística, 2009.

[9] J. Felipe Díaz and J. Carlos Correa, "Comparación entre árboles de regresión CART," *Comunicaciones en Estadística*, vol. 6, no. 2, pp. 175-195, 2013.https://doi.org/10.15332/s2027-3355.2013.0002.05

[10] IBM, "IBM Cognos Analytics," 2022. [Online]. Available: https://www.ibm.com/docs/es/cognos-analytics/11.1.0?topic=tests-multiple-linear-regression. [Accessed 12 10 2022].

[11] J. Amat Rodrigo, "cienciadedatos.net," https://cienciadedatos.net/documentos/25_regresion_lineal_multiple, 2016. [Online]. Available: https://cienciadedatos.net/documentos/25_regresion_lineal_multiple. [Accessed 2 10 2022].

[12] D. R. Tobergte and S. Curtis, "Introducción a la econometría: Un enfoque moderno," *Journal of Chemical Information and Modeling*, vol. 53, no. 9, 2013.

[13] M. Torres, K. Paz and F. Salazar, "Métodos de recolección de datos para una investigación," no. 2, pp. 2-21, 2019.

[14] R. Hernadez Sampieri, C. Fernández Collado and P. Batista Lucio, Metodología de la Investigación, Cuarta ed., Mexico: MACGRAW-HILL, 2010.

[15] J. I. Espinosa Muñoz, "na aproximación a la predicción del valor de acciones en la bolsa de valores aplicando técnicas de Data Mining," Universidad Politécnica de Madrid, Madrid, 2015.

[16] J. . F. D. Sepúlveda Díaz and J. C. Correa Morales, "Comparación entre árboles de regresión CART y regresión lineal," *Comunicaciones en Estadística 6.2*, pp. 175-195, 2013. https://doi.org/10.15332/s2027-3355.2013.0002.05

Biografías

Edgar Eloy Carpio Vargas, Dr. en estadística e informática, especialista en ciancia de datos, docente de la facultad de Ingenieria Estadistica e Informática, docente RENACYT.

Alicia Roxana Mayda Huanca, Ing Estadistico e informatico de la Universidad Nacional del Altiplano.

Espinoza Rivas German Rafael, especialista en geología, geotecnia y medio ambiente, con estudios de Ingeniería Geológica, Ingeniería Civil, y Topografía en universidades del Perú, maestría en Ingeniería Civil y Medio Ambiente en la Universidad de Utah, USA y segunda maestría en Ingeniería Ambiental en la Universidad Nacional del Altiplano. Doctor en Ciencia, Tecnología y Medio Ambiente.

Ecler Mamani Vilca, Universidad Nacional Micaela Bastidas de Apurímac - Perú, Dr. en Ciencias de la Computación, desarrolla-dor de aplicaciones multimedia y Software Educativo Intercultural.

Betsabe M. Ccolqque Ruiz, docente en la Universidad Nacional Micaela Bastidas de Apurímac, Ingeniero Informático y Sistemas con Magister en SEGURIDAD DE LA INFORMACIÓN Y TECNOLOGÍA y estudiante de doctorado.

C&T Riqchary
Revista de Investigación en Ciencia y Tecnología
ISSN: 2810-8124 (en línea) / ISSN: 2706-543x

Vol. 5 Núm. 1 (2023)
Publicado: 15/10/23
DOI: https://doi.org/10.57166/riqchary/v5.n1.2023.118

Determinación del mejor algoritmo de reconocimiento de imágenes de camélidos sudamericanos mediante machine learning, Apurímac, 2023

Determination of the best image recognition algorithm for south american camelids using machine learning, Apurímac, 2023

Filio Carrasco-Sauñe [A], Manuel J. Ibarra-Cabrera [B]

ORCID:0009-0006-5867-275X [A], ORCID:0000-0001-6711-4916 [B]

(Recepción: 15/04/2023 y aceptación 08/10/2023)

Resumen— Las alpacas y las vicuñas pertenecen a la familia de los camélidos sudamericanos, la alpaca es un animal doméstico, mientras que la vicuña es un animal silvestre que generalmente vive en alturas superiores a los 3000 m.s.n.m. El problema es que en el caso de las vicuñas es difícil realizar el reconocimiento y conteo y porque por su naturaleza de ser animal silvestre huyen de las personas y de cualquier otro objeto extraño que no esa de su entorno; el otro problema es que no se tiene una base de datos de imágenes o fotos de vicuñas para hacer un estudio de reconocimiento de imágenes. Este trabajo consistió en recolectar 146 imágenes de vicuña, de los cuales 95, fueron consideradas para el entrenamiento 41 para la validación y 10 para las pruebas; las imágenes fueron recolectadas en un centro poblado del distrito de Cotaruse, provincia de Aymaraes en la región Apurímac de Perú. Para el caso de las alpacas, las imágenes se obtuvieron del repositorio de Kaggle y se trabajó con 142 imágenes de las cuales 95 fueron consideradas para el entrenamiento, 41 para la validación y 10 para las pruebas. Los resultados muestran que el algoritmo o modelo Mask R-CNN obtiene valor para accuracy de 1,0 para las vicuñas y de 0,95 para las alpacas; estos valores son los más eficientes en relación con los encontrados por Yolo V8 y SSMD.

Palabras clave: camélidos, reconocimiento, red neuronal, vicuña

Abstract— Alpacas and vicuñas belong to the South American camelid family; the alpaca is a domestic animal, while the vicuña is a wild animal that generally lives at altitudes above 3,000 meters above sea level. The problem is that in the case of vicuñas it is difficult to recognize and count them because they are wild animals and because they flee from people and any other foreign object that is not part of their environment; the other problem is that there is no database of images or photos of vicuñas to carry out an image recognition study. This work consisted of collecting 146 vicuña images, of which 95 were considered for training, 41 for validation and 10 for testing; the images were collected in a village center in the district of Cotaruse, province of Aymaraes in the Apurimac region of Peru. In the case of alpacas, the images were obtained from the Kaggle repository and 142 images were used, of which 95 were considered for training, 41 for validation and 10 for testing. The results show that the Mask R-CNN algorithm or model obtains a value for accuracy of 1.0 for vicuñas and 0.95 for alpacas; these values are the most efficient in relation to those found by Yolo V8 and SSMD.

Keywords: camelids, neural network, recognition, vicuña

A. Filio Carrasco-Sauñe, Escuela Profesional de Ingenieria Informática y Sistemas de la Universidad Nacional Micaela Bastidas de Apurímac-Perú, 112134@unamba.edu.pe.

B. Manuel J. Ibarra-Cabrera, Departamento Académico de Informática y Sistemas de la Universidad Nacional Micaela Bastidas de Apurímac-Perú, mibarra@unamba.edu.pe.

1 Introducción

Actualmente, existen cuatro especies de camélidos sudamericanos bastante populares: llamas y alpacas que son conocidos como animales domésticos; y vicuñas y guanacos que son conocidos como animales silvestres.

La vicuña es uno de los animales que vive en los andes del Perú y resulta complicado determinar con precisión la cantidad exacta de estos seres vivos. Esto se debe a que llevar a cabo el monitoreo de la población de animales silvestres resulta difícil y todavía se basa en el conteo manual o en el uso de transectos, que es un método de muestreo para calcular la densidad poblacional.

Las cuatro especies de los camélidos sudamericanos son: llamas y alpacas que son conocidos como animales domésticos; y vicuñas y guanacos que son conocidos como animales silvestres [1]. La vicuña es una especie animal silvestre que está en peligro de extinción, debido a que su lana es de alta calidad y tiene un costo elevado en el mercado [2], por tanto, parece ser atractivo en término de costos para las personas que los cazan furtivamente [3]. Mayormente viven sobre los 3000 m.s.n.m. en la zona sierra o los andes de los países de Argentina, Bolivia, Chile, Ecuador y Perú [4].

El problema es que realizar el conteo de animales silvestres es un proceso complejo, específicamente el conteo de vicuñas [5] se realiza una vez al año o cada dos años en una actividad que se llama "vicuña chaco", y se usa el método conteo por aproximación, o transecto [6]. Para realizar el conteo, previamente se requiere el reconocimiento, por esta razón, esta investigación se enfoca en el reconocimiento de imagen de vicuña.

Un segundo problema asociado al conteo de las vicuñas, es que no existen imágenes de vicuñas disponibles en las bases de datos de libre uso y acceso para realizar reconocimiento y conteo utilizando algoritmos de machine learning. Una alternativa para obtener las imágenes de animales silvestres, es uso de un dron o cámaras fotográficas escondidas, y así tener una opción de realizar el reconocimiento, conteo y monitoreo de animales [7].

Por otro lado, se tiene disponible una base de datos de imágenes para la clasificación de alpacas en *Kaggle* [8] (repositorio de imágenes para realizar entrenamientos y pruebas con algoritmos de machine learning); sin embargo, no se ha podido encontrar base de datos para llama, guanaco ni vicuña, que también son especies que pertenecen a los camélidos sudamericanos, esto dificulta realizar entrenamiento y pruebas con algoritmos de machine learning.

Los algoritmos de machine learning cada vez son más importantes y se pueden aplicar a diversas áreas, consiste en entrenar un conjunto de imágenes en base ciertas reglas o datos, después se tiene un modelo entrenado y finalmente se realizan las pruebas para determinar la precisión [9]. Existen algoritmos para el reconocimiento de objetos, y están basados en redes neuronales, redes bayesianas y basados en características [10], [11].

Existen algunos autores que han estudiado el reconocimiento de animales, así por ejemplo Xiao et al. [12] realizaron una investigación cuyo objetivo fue proponer un método para la identificación o reconocimiento de vacas basado en visión por computadora. La metodología que siguieron fue obtener las imágenes en un ambiente libre en un granero, debido a los cambios en la posición y el ángulo de la cámara, los patrones de las vacas grabados a menudo se deformaron, lo que dificultó la identificación individual de las vacas. Primero, se obtiene una imagen de vista superior de una vaca y se utiliza una Mask R-CNN mejorada para segmentar esta imagen y extraer las características de la forma del lomo de la vaca. Luego, se utilizó un enfoque de Fisher para seleccionar el mejor subconjunto de características y se aplicó un clasificador de máquina de vectores de soporte (SVM) para identificar vacas individuales. Para verificar la efectividad del algoritmo de detección de objetivos, el método propuesto se compara con el modelo Mask R-CNN tradicional, y se comparan la precisión, la recuperación, la puntuación F1, el tiempo de ejecución promedio por imagen y la precisión promedio del modelo Mask R-CNN mejorado. 98,21%, 96,48%, 97,34%, 1,02s y 97,39%, respectivamente. Para la identificación individual de las vacas se utiliza un clasificador SVM entrenado en función de las características de forma obtenidas. El método propuesto logra una precisión de identificación de vacas del 98,67% basándose en un conjunto de datos que contiene imágenes de vista superior de 48 vacas. Los resultados demuestran la efectividad del método de identificación de vacas propuesto y su importante potencial para su uso en el manejo de precisión de las vacas lecheras.

Por otro lado, Li et al. [13] realizaron una investigación que usa el aprendizaje profundo para reconocer las características de vacas individuales. Alexnet se utiliza como red esqueleto para una red neuronal convolucional liviana que puede reconocer vacas individuales en imágenes con fondos complejos. Se agregó un módulo de inicio mejorado y un mecanismo de atención para extraer características en múltiples escalas para mejorar la detección de puntos característicos. En los experimentos, se recogieron imágenes de vista lateral de 13 vacas. El método propuesto logró una precisión del 97, 95% en la identificación de vacas con un único tiempo de entrenamiento de solo 6s, que es una sexta parte del Alexnet original. Para verificar la validez del modelo, el conjunto de datos y los parámetros experimentales se mantuvieron constantes y se compararon con los resultados de Vgg16, Resnet101, Mobilenet V2 y GoogLenet. El modelo propuesto garantiza una alta precisión y tiene el tamaño de parámetro más pequeño de 6,51 MB, que es 1,3 veces menor que el de la red Mobilenet V2, famosa por su peso ligero.

El objetivo de este proyecto de investigación fue determinar el mejor algoritmo de Machine Learning para el reconocimiento de dos especies de los camélidos sudamericanos; específicamente en el reconocimiento de imágenes de vicuñas, para lo cual se recolectó imágenes de Apurímac mediante cámara fotográfica; y por otro lado, en el reconocimiento de imágenes de alpacas de acuerdo al dataset de la base de datos en Kaggle. El entrenamiento se realizó con tres modelos o algoritmos de redes neuronales convolucionales: Yolo V8 [14], Mask R-CNN Resnet101 [15] y SSMD (Single Shot Multibox Detector (SSMD) mobileNet 2.0 [16].

2 MÉTODO

2.1 Ámbito de estudio

Este estudio se realizó en la Universidad Nacional Micaela Bastidas de Apurimac (UNAMBA) en la ciudad de Abancay; y los datos fueron recolectados en el distrito de Cotaruse y provincia de Aymaraes en el año 2023.

2.2 Tipo y nivel de investigación

Este trabajo es una investigación de tipo aplicada y de nivel descriptivo.

2.3 Procedimiento. fue el siguiente:

a) Se recolectó las imágenes de vicuñas mediante cámara fotográfica de un celular *Iphone 11 pro max*, con una resolución máxima de 2268x4032 pixeles, y la calidad de

Revista de Investigación en Ciencia y Tecnología
ISSN: 2810-8124 (en línea) / ISSN: 2706-543x

Vol. 5 Núm. 1 (2023)
Publicado: 15/10/23
DOI: https://doi.org/10.57166/riqchary/v5.n1.2023.118

video de 30 fotogramas por segundo.

b) Se eliminaron las imágenes que estuvieran desenfocadas, con mala calidad o que tengan muchas imágenes superpuestas.
c) Se separó aproximadamente 70,0% imágenes para entrenamiento, 20,0% para validación y 10,0% de las imágenes para pruebas.
d) Luego, sólo a las imágenes separadas para el entrenamiento se procedió a etiquetarlas utilizando el software de etiquetado CVAT [17], que es una plataforma especializada para realizar anotaciones sobre imágenes.
e) Se descargó de *Kaggle* las imágenes de alpaca y se procedió a entrenar con las diferentes técnicas de machine learning.
f) Luego, se entrenó las imágenes de vicuña y alpaca con las diferentes técnicas de machine learning, y se utilizó un *cuaderno de Google Collab* con la cuenta de correo electrónico de *filiocs27@gmail.com*. El entrenamiento demoraba alrededor de 40 minutos con una cuenta estándar de Google, lo cual no permitía completar el proceso de enternamiento; por lo que se tuvo que adquirir y utilizar una cuenta *premium de Google* para que el entrenamiento no se corte, esto se consigue pagando 9,90 dólares adicionales. El entrenamiento se realizó utilizando herramientas de programación y librerías con sus distintas versiones, por ejemplo: cudatoolkit==11.0.3, cudnn==8.0.5.39, numpy==1.19.2, scikit-learn==1.2.0, scipy==1.10.1, seaborn ==0.10.0, pandas==2.0.3, pillow==10.0.0, matplotlib==3.6.0, tensorflow-gpu==2.4.0, opencv-python==4.5.4.60, ipython ==8.10.0.
g) Luego se realizaron las pruebas con cada una de las técnicas con las imágenes separadas para el testing (imágenes sin etiquetar).

2.4 Población

Las imágenes o fotos de vicuñas se obtuvieron en el distrito de Cotaruse provincia de Aymaraes en la región Apurímac, cuyas coordenadas geográficas son 14°31'43.4"S y 73°27'15.6"W. Se obtuvieron un total de 146 imágenes de buena calidad. La Fig. 1 muestra las imágenes tomadas en los andes de Apurímac a más de 3000 m.s.n.m.



Fig. 1. Imágenes tomadas en el campo

Para el proceso de entrenamiento se obtuvo una muestra de 96 imágenes para entrenamiento, 41 imágenes para la validación y 10 para el testing. Este muestreo fue de forma aletoria.

Para el caso de las alpacas la población fue extraída del repositorio de imágenes *Kaggle*, con un total de 142 imágenes de alpacas; de las cuales 102 se utilizaron para el entrenamiento, 30 para validación, y 10 para el testing.

2.5 Materiales e Instrumentos

Dentro de los materiales utilizados están: a) Cámara fotográfica b); Google Collab; y c) Recurso repositorio digital Kaggle.
Dentro de los instrumentos utilizados están: a) Registro de imágenes de vicuñas; b) Cuaderno notebook cuyo archivo tiene extensión *.ipynb; y c) Registro de imágenes de Alpacas

3 Resultados y Discusión

3.1 Pruebas y resultados para las vicuñas

La tabla 1 muestra la Matriz de Confusión para las vicuñas, en la cual se puede observar los resultados de los tres modelos entrenados, Yolo V8, Mask R-CNN Resnet101 y SSMD mobileNet 2.0, "TP" representa a los verdaderos positivos, "FP" los falsos positivos "FN" falsos negativos y "TN" verdaderos negativos.

TABLA 1
Matriz de confusión para vicuña

Algoritmo o modelo de Machine Learning	TP	FP	FN	TN
Yolo V8	11	0	2	0
Mask R-CNN Resnet101	11	0	0	0
SSMD mobileNet 2.0	10	1	2	0

La tabla 2 muestra los resultados estadísticos de las pruebas realizadas con las imágenes de vicuñas en los modelos o algoritmos de Yolo V8, Mask R-CNN Resnet101 y SSMD mobileNet 2.0, asimismo se tienen tres variables de medición "Accuracy", "Precision" y "Recall". Los resultados muestran que Mask R-CNN Resnet101 tiene el mejor resultado con un valor de 100%, luego esta Tolo V8 y finalmente SSMD. Estos resultados fueron realizados considerando 200 épocas de entrenamiento.

TABLA 2
Estadísticas para vicuña

Algoritmo o modelo de Machine Learning	Accuracy	Precision	Recall
Yolo V8	1,00	1,00	0,85
Mask R-CNN Resnet101	1,00	1,00	1,00
SSMD mobileNet 2.0	0,92	0,90	0,83

La tabla 3 muestra las predicciones realizadas a un conjunto de 10 fotos de imágenes de vicuñas, donde "N" significa un seudónimo asignado a la figura, "cantidad" significa la cantidad de vicuñas que hay en la foto, y seguidamente están los modelos: Yolo V8, Mask R-CNN y SSMD. Así por ejemplo, la figura con pseudónimo "V10", contiene 2 imágenes de vicuñas, y Yolo V8 obtiene un valor 0,803 y 0,94; Mask R-CNN 0,997 y 0,997; y SSMD 0,65 y 0,94. Por lo que se puede observar que Mask R-CNN obtiene mejores resultados.

TABLA 3
Predicciones para Vicuña

N	Cantidad	Yolo V8	Mask R-CNN Resnet101	SSMD mobileNet 2.0
V01	1	0,92	0,997	0,89
V02	1	0,92	0,996	0,90
V03	1	0,92	0,999	0,89
V04	1	0,96	0,998	0,88
V05	1	0,93	0,998	0,86
V06	1	0,90	0,998	0,87
V07	1	0,84	0,999	0,44
V08	1	0,91	1,000	9,83
V09	1	0,89	0,999	0,79
V10	2	0,83-0,94	0,997-0,997	0,65-0,94

La Fig. 2 muestra el resultado de las predicciones en las vicuñas, en la que se observa que Mask R-CNN obtiene los mejores valores con 0,997 y 0,997 para cada vicuña.

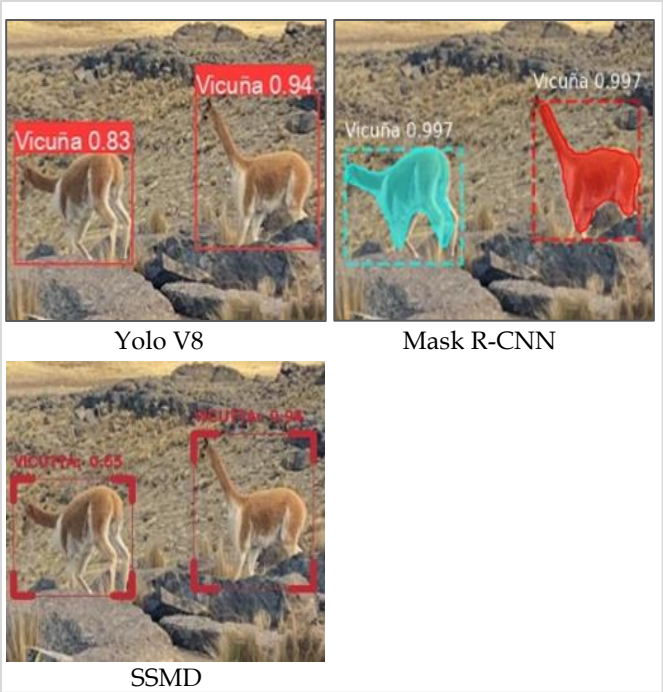


Fig. 2. Predicciones para imágenes de vicuña.

Para el caso de Mask R-CNN, la Fig. 3 muestra las métricas de precisión y recall versus las épocas, y lo cual nos indica que a partir de la época 60 aproximadamente ya se muestra un valor estabilizado en el proceso de entrenamiento; de forma similar, la Fig. 4 muestra la pérdida versus las épocas, y lo cual nos indica a partir de la época 80 ya se muestra un valor estabilizado en el proceso de entrenamiento.

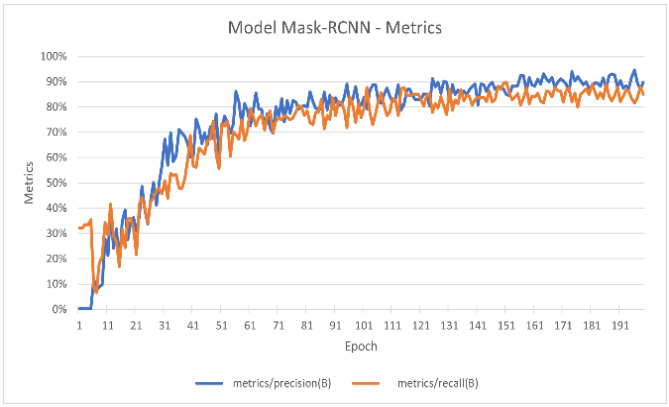


Fig. 3. Métricas vs época en Mask R-CNN

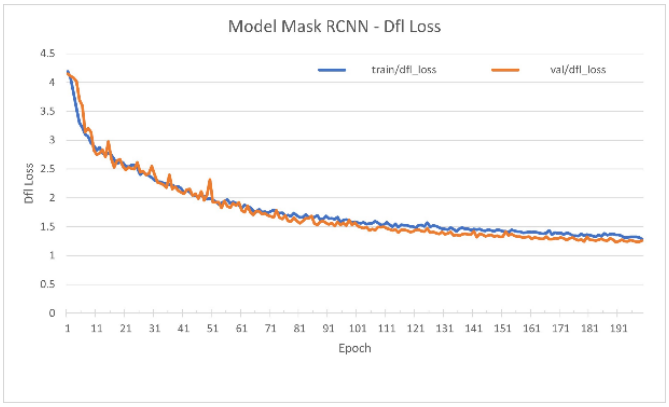
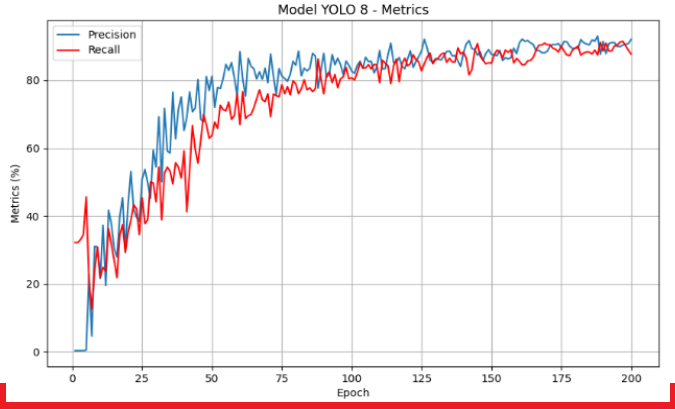


Fig. 4. Pérdida vs época en Mask R-CNN

Para el caso de Yolo V8, la Fig. 5 muestra las métricas de precisión y recall versus las épocas, y lo cual nos indica que a partir de la época 75 aproximadamente ya se muestra un valor estabilizado en el proceso de entrenamiento; de forma similar, la Fig. 6 muestra la pédida versus las épocas, y lo cual nos indica que en la época 125 aproximadamente ya se muestra un valor estabilizado en el proceso de entrenamiento.



Revista de Investigación en Ciencia y Tecnología
ISSN: 2810-8124 (en línea) / ISSN: 2706-543x

Vol. 5 Núm. 1 (2023)
Publicado: 15/10/23
DOI: https://doi.org/10.57166/riqchary/v5.n1.2023.118

Fig. 5. Métricas vs época en Yolo V8

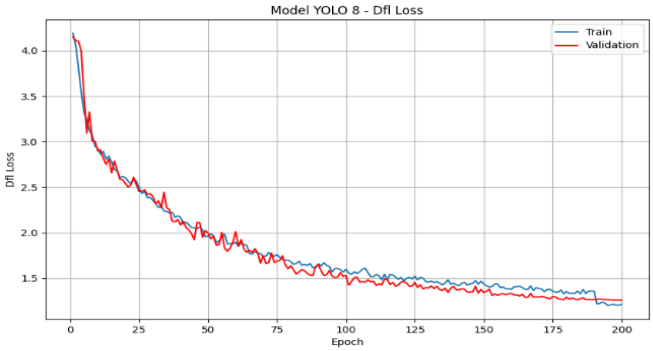


Fig. 6. Pérdida vs época en Yolo V8

3.2 Pruebas y resultados para las alpacas

La tabla 4 muestra la Matriz de Confusión para las alpacas, en la cual se puede observar los resultados de los tres modelos entrenados, Yolo V8, Mask R-CNN Resnet101 y SSMD mobileNet 2.0, "TP" representa a los verdaderos positivos, "FP" los falsos positivos "FN" falsos negativos y "TN" verdaderos negativos.

TABLA 4
Matriz de confusión para alpaca

Algoritmo o modelo de Machine Learning	TP	FP	FN	TN
Yolo V8	*34*	*0*	*4*	*0*
Mask R-CNN Resnet101	*21*	*0*	*1*	*0*
SSMD mobileNet 2.0	*26*	*0*	*10*	*0*

La tabla 5 muestra la Matriz de Confusión para las Alpacas, en la cual se puede observar los resultados de los tres modelos entrenados, Yolo V8, Mask R-CNN Resnet101 y SSMD mobileNet 2.0, "TP" representa a los verdaderos positivos, "FP" los falsos positivos "FN" falsos negativos y "TN" verdaderos negativos.

TABLA 5
Estadísticas para alpaca

Algoritmo o modelo de Machine Learning	Accuracy	Precision	Recall
Yolo V8	0,89	1,00	0,89
Mask R-CNN Resnet101	0,95	1,00	0,95
SSMD mobileNet 2.0	0,72	1,00	0,72

La tabla 6 muestra las predicciones realizadas a un conjunto de 10 fotos de imágenes de alpacas, conde "N" significa un seudónimo asignado a la figura, "cantidad" significa la cantidad de alpacas que hay en la foto, y seguidamente están los modelos: Yolo V8 Mask R-CNN y SSMD. Así, por ejemplo, la figura "A10", contiene 3 imágenes de alpacas, y Yolo V8 obtiene un valor 0,57, 0,71 y 0,27; Mask R-CNN 0,998, 0,996 y 0,953; y SSMD 0,68, 0,85 y 0,56. Por lo que se puede observar que Mask R-CNN tiene mejores resultados

La Fig. 7 muestra el resultado de las predicciones en las alpacas, en la que se observa que Mask R-CNN obtiene el mejor valor de 0,999; esto implica que existe un 99.9% de certeza en que la imagen sea alpaca; en segundo lugar, se encuentra la predicción con Yolo V8 que tiene un valor de 0,84 y en tercer lugar se encuentra la predicción realizada con SSMD con un valor de 0,70.

TABLA 6
Predicciones para Alpaca

N	Cantidad	Yolo V8	Mask R-CNN Resnet101	SSMD mobileNet 2.0
A01	1	0,84	0,999	0,70
A02	1	0,84	0,999	0,53
A03	2	0,59-0,83	0,996-0,997	0,70-0,87
A04	2	0,97-0.98	0,999-0,988	0,81-0,88
A05	2	0,91-0.98	1,000-0,998	0,70-
A06	2	0,93-0,94	0,999-0,998	0,64 0,54
A07	1	0,91	1,000	0,68
A08	2	0,84-0,87	1,000-0,999	0,86-
A09	3	0,82-0,82	1,000-0,999-	0.81-0,75-0,75
A10	3	0,57-0,71-0,27	0,998-0,996-0,953	0,68-0,85-0,56

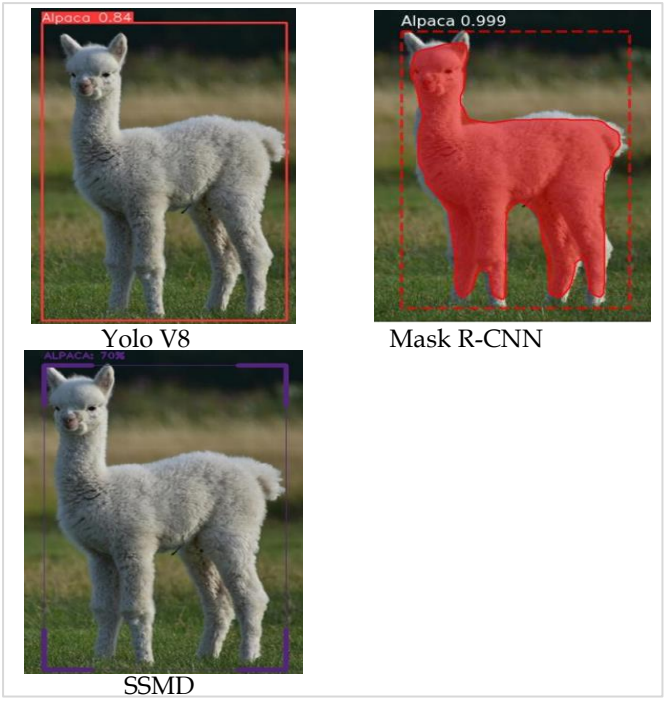


Fig. 7. Predicciones para imágenes de alpaca

3.3 Discusión

Según Xiao et al. [12], el reconocimiento de imágenes de 48 vacas lecheras mediante el modelo Mask R-CNN tradicional obtuvieron resultados satisfactorios con una precisión de 0,98, y la recuperación de 0,96 en una escala de 0 a 1. Para la identificación individual de las vacas se utiliza un clasificador SVM entrenado en función de las características

de forma obtenidas. Estos resultados están alineados con los resultados obtenidos por Xiao et al. [12], puesto que para el caso de las vicuñas se obtuvieron el modelo de Mask R-CNN valores de precisión 1,00 y recuperación de 0,95, y pare el caso de las vicuñas; y para el caso de las alpacas precisión de 1,00 y recuperación de 0,95. Es preciso mencionar, que no existe pruebas realizadas con camélidos sudamericanos, pero si existen investigación con vacas.Por otro lado, lo investigado por Li et al. [13] en el cual usan el aprendizaje profundo para reconocer las características de vacas individuales. En los experimentos muestran que, el método propuesto tiene una precisión del 97,95% en la identificación de vacas lecheras y con un único tiempo de entrenamiento de solo 6 s, que es una sexta parte del Alexnet original y que es 1,3 veces menor que el de la red Mobilenet V2, famosa por su peso ligero. La investigación de Li et al. [13] se enfoca prioritariamente en el tiempo de respuesta en la identificación, mientras que nuestra investigación es orientada para camélidos sudamericanos enfocado en el reconocimiento de imágenes y no en el tiempo de respuesta.

4 Conclusiones

Este trabajo de investigación muestra que, para el algoritmo más eficiente (en base a la variable precisión) para el reconocimiento de camélidos sudamericanos, es Mask R-CNN Resnet101, seguido de Yolo V8 y finalmente SSMD. Los algoritmos fueron probados con 146 imágenes de vicuñas tomadas en Apurímac-Perú y 142 imágenes de alpacas extraídas de Kaggle.

Agradecimientos

Queremos agradecer a CONCYTEC-FONDECYT por el apoyo en el desarrollo de esta investigación, mediante el fondo concursable que tiene contrato número 376-2019.

Referencias

[1] H. D. Yacobaccio, "The domestication of South American camelids: a review," *Anim. Front.*, vol. 11, no. 3, pp. 43–51, 2021, https://doi.org/10.1093/af/vfaa065

[2] J. E. Quispe Coaquira, T. Herrera Mamani, E. Apaza Zúñiga, L. Clavetea Quisca, and Z. Maquera Marón, "Características tecnológicas de la fibra de vicuñas en semicautiverio de la Multicomunal Picotani-Región Puno," *Rev. Investig. Vet. del Perú*, vol. 29, no. 2, pp. 522–532, 2018, https://doi.org/10.15381/rivep.v29i2.14475

[3] G. Rodríguez García, "Animales...? En peligro de extinción o en peligro de que los extingamos?," *Ius Prax.*, vol. 14, no. 1, pp. 301–315, 2008, doi: http://dx.doi.org/10.4067/S0718-00122008000100011.

[4] J. Laker, J. Baldo, Y. Arzamendia, and H. D. Yacobaccio, "La vicuña en los Andes," *Investig. Conserv. y manejo vicuñas*, pp. 37–50, 2006, [Online]. Available: https://www.vicam.org.ar/publi/VicuniaAndes.pdf

[5] R. Hoffmann, K. Otte, and C. Ponce, *El manejo de la vicuña silvestre Tomo I*, vol. I. Alemania: Sociedad Alemana de Cooperación Técnica- GTZ, 1983.

[6] G. Gallardo, A. Nuñez, and L. F. Pacheco, "Transectos lineales como opción para estimar abundancia de vicuñas (Vicugna vicugna): Estudio de caso en el Parque Nacional Sajama, Bolivia," *Ecol. en Boliv.*, vol. 45, pp. 64–72, 2010.

[7] K. S. Christie, S. L. Gilbert, C. L. Brown, M. Hatfield, and L. Hanson, "Unmanned aircraft systems in wildlife research: Current and future applications of a transformative technology," *Front. Ecol. Environ.*, vol. 14, no. 5, pp. 241–251, 2016, doi: 10.1002/fee.1281, https://doi.org/10.1002/fee.1281

[8] Kaggle, "Image Dataset," 2023. https://www.kaggle.com/datasets/sid4sal/alpaca-dataset-small (accessed Jan. 02, 2023).

[9] J. Hu, X. Wu, and M. Dai, "Estimating the population size of migrating Tibetan antelopes Pantholops hodgsonii with unmanned aerial vehicles," *Oryx*, pp. 1–9, 2018, https://doi.org/10.1017/S0030605317001673

[10] H. Nguyen *et al.*, "Animal recognition and identification with deep convolutional neural networks for automated wildlife monitoring," in *2017 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*, 2017, pp. 40–49. https://doi.org/10.1109/DSAA.2017.31.

[11] A. Gomez, G. Diez, A. Salazar, and A. Diaz, "Animal identification in low quality camera-trap images using very deep convolutional neural networks and confidence thresholds," in *International Symposium on Visual Computing*, 2016, pp. 747–756. https://doi.org/10.1007/978-3-319-50835-1_67

[12] J. Xiao, G. Liu, K. Wang, and Y. Si, "Cow identification in free-stall barns based on an improved Mask R-CNN and an SVM," *Comput. Electron. Agric.*, vol. 194, 2022, https://doi.org/10.1016/j.compag.2022.106738

[13] S. Li *et al.*, "Individual dairy cow identification based on lightweight convolutional neural network," *PLoS One*, vol. 16, no. 11 November, 2021, https://doi.org/10.1371/journal.pone.0260510

[14] M. Hussain, "YOLO-v1 to YOLO-v8, the Rise of YOLO and Its Complementary Nature toward Digital Manufacturing and Industrial Defect Detection," *Machines*, vol. 11, no. 7, p. 677, 2023, https://doi.org/10.3390/machines11070677

[15] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask r-cnn," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2961–2969, https://doi.org/10.1109/ICCV.2017.322

[16] S. A. Magalhães *et al.*, "Evaluating the single-shot multibox detector and YOLO deep learning models for the detection of tomatoes in a greenhouse," *Sensors*, vol. 21, no. 10, p. 3569, 2021, https://doi.org/10.3390/s21103569

[17] CVAT, "CVAT-Open Data Annotation Platform," 2023. https://www.cvat.ai/

Biografías

Filio Carrasco Sauñe, bachiller en Ingeniería Informática y Sistemas de la Universidad Nacional Micaela Bastidas de Apurímac, actualmente es analista y programador de software de la empresa Avalon Soluciones S.A.C.

Manuel Jesús Ibarra Cabrera, doctor en ciencias de la computación e investigador en las áreas de ingeniería de software, serious game, informática educativa, computación móvil, IoT e industria y sociedad. Docente universitario de pre y pos grado en la Universidad Nacional Micaela Bastidas de Apurímac, Universidad Tecnológica de los Andes, Universidad Nacional del Altiplano y la Universidad Nacional Mayor de San Marcos. Actualmente, es miembro activo de conferencias importantes en latinoamerica: LACLO, CLEI, CONTIE, HCI, SABTIC, CISTI, Decisioning y otros.

C&T Riqchary
Revista de Investigación en Ciencia y Tecnología
ISSN: 2810-8124 (en línea) / ISSN: 2706-543x

Vol. 5 Núm. 1 (2023)
Publicado: 15/10/23
DOI: https://doi.org/10.57166/riqchary/v5.n1.2023.119

Caracterización de los suelos de la comunidad de Huayllati, Apurímac

Characterization of the soils of the community of Huayllati, Apurimac

Norma Aguirre-Quispe[A], Niki Franklin Flores-Pacheco [B], Salvador Quispe-Chipana[C] y Maria Luisa Flores-Pacheco [D]

ORCID:0009-0004-5131-703X[A], ORCID:0000-0002-2772-0951[B], ORCID:0000-0002-5509-9437[C] y ORCID:0009-0004-7485-8875[D]

(Recepción: 25/09/2023 y aceptación 10/10/2023)

Resumen— La investigación se realizó con el propósito de conocer las caracteristicas quimicas y físicas de los suelos. La metodología de investigacion fue no experimental, de tipo descripitvo, de corte tranversal con un enfoque cuantitativo. La población en estudio fueron los suelos de los sectores de Accahuaray, Topopampa y Padroma de la Comunidad de Huayllati, provincia de Grau, region Apurimac. La muestra fue tomada de manera aleatoria y estratificada y posteriormente analizadas en laboratorio. El resultado mostró que las características químicas y físicas de los suelos, muestran diferencias que nos llevo a la conclusión que el sector Padroma, tiene suelos con mejores caracteristicas químicas y el sector Topopampa con mejores caracteristicas físicas, superior a los otros sectores.

Palabras clave: Caracteristicas, químicas, físicas, suelos, sectores

Abstract— The research was carried out with the purpose of knowing the chemical and physical characteristics of the soils. The research methodology was non-experimental, descriptive, cross-sectional with a quantitative approach. The study population was the soils of the Accahuaray, Topopampa and Padroma sectors of the Huayllati Community, Grau province, Apurimac region. The sample was taken in a random and stratified manner and subsequently analyzed in the laboratory. The result showed that the chemical and physical characteristics of the soils show differences that led us to the conclusion that the Padroma sector has soils with better chemical characteristics and the Topopampa sector with better physical characteristics, superior to the other sectors.

Keywords: Characteristics, chemical, physical, soils, sectors

A. *Norma Aguirre Quispe[A] Ingeniero Agroecoloco Rural - Universidad Nacional Micaela Bastidas de Apurimac: aguin815@gmail.com*
B. *Niki Franklin Flores Pacheco[B] Ingeniero Agronomo – Docente del Departamente Academico de Ingenieria - Universidad Nacional Micaela Bastidas de Apurimac: nflores@unamba.edu.pe.*
C. *Salvador Quispe Chipana [C] Ingeniero Agronomo – Docente de la Escuela Profesional de Ingenieria Agropecuaria - Universidad Nacional de San Antonio Abad del Cusco: salvador.quispe@unsaac.edu.pe*
D. *Maria Luisa Flores Pacheco [D] Ingeniero Agronomo - Universidad Tecnologica de los Andes:: marialuisaflorespacheco7@gmail.com*

1 INTRODUCCIÓN

El suelo es un recurso muy importante en la producción de alimentos, en la generación de riqueza y bienestar. Para lograr el uso sostenible es importante conocer sus caracteristicas químicas, físicas y entre otros factores que influyen en su conservación y uso.

El uso de lo suelos, sin un análisis físico-quimico y descuidando practicas agricolas de conservación genera erosion, sobrepastoreo y acelera la perdida de su capacidad productiva. Por lo tanto, el suelo es primordial para el desarrollo social y económico de la Comunidad de Huayllati, este recurso asegura en gran parte la producción de alimentos de origen pecuario y agrícola que son nescesarios para la alimentación y el bienestar.

El estudio tuvo el propósito conocer las características químicas y físicas de los suelos. Porque es fundamental conocer las características para realizar programas planificación para el uso adecuado, implementando programas de abonamiento y recuperación que asegure una producción de alimentos adecuada para la comunidad.

En el estudio se observo que gran parte de los agricultores han perdido o no practican las técnicas de manejo apropiado de los suelos y que es nescesario poner en valor. Por tal motivo para iniciar un proceso de recuperación de los suelos es importante conocer las caracteristicas en los tres sectores Padroma, Topopampa y Accahuaray.Nuestro estudio se fundamenta en las

investigaciones realizadas por varios investigadores, que detallamos a continuación:

Maldonado **[1]**, realizo el estudio que tuvo como objetivo determinar la aptitud de uso del suelo en Karhuiza, La Paz. La metodología que utilizó en la investigacion fue de tipo descriptivo, no experimental. Los resultados en las características físicas que obtuvo a través del análisis de los perfiles, donde observó suelos superficiales, con textura franco-limoso, franco, arcillosos y arenosos, con estructura granular y permeabilidad adecuada. Las caracteristicas químicas observadas fueron suelos con un pH con intervalos de 4,07 a 6,85 es decir suelos ácidos muy cercanos a neutro; sales con intervalos de 0,85 a 5,37 (dS/m), suelos moderadamente salinos; CIC con intervalos de 7,72 a 61,78 cmol (+)/kg-1; grado de saturación de bases con intervalos de 5,63 a 96,7%. Concluyó que los suelos son superficiales, con bajos niveles de fertilidad natural, con problemas acidez y variabilidad textural.

Gomez **[2]**, realizo el estudio que tuvo como objetivo caracterizar morfológicamente, taxonómicamente, física y químicamente los suelos en la parroquia El Cambio, de El Oro. La metodología de la investigacion fue descriptiva no experimental. Obtuvo los resultados, en las caracteristicas químicas: Suelos con pH con intervalos de 7,01 a 8,95, capacidad de intercambio catiónico (CIC) con intervalos 21,6 a 52,2 meq/100g; Conductividad eléctrica (CE) con promedios de 0,14 dS m-1 y materia orgánica (MO) con intervalos de 0,02% - 2,55%. Las caracteriticas físicas de los suelos, con densidad aparente en promedio de 1,5 g cm-3 y texturas: franco-arenoso, franco-arcilloso, franco-limoso, franco-arenoso. Llegando a la conclusión que los suelos son cercanos a la neutralidad, suelos con texturas gruesas y contenido de arcillas, con problemas de salinidad, bajos niveles de materia organica, con cierto grado de compactación y variabilidad de texturas.

Vivas **[3]**, realizó el estudio que tuvo como objetivo: caracterizar y evaluar la fertilidad de los suelos en el valle de Chira-Piura. La metodología de investigacion utilizada fue descriptiva y no experimental. Los resultados en las caracteristicas químicas: Suelos con pH con intervalos de 7,02 a 8,04; conductividad eléctrica (CE) con intervalos de 0,37 a 0,67dS/m; materia orgánica (MO) con intervalos de 0,92 a 2,2%. Nitrógeno con intervalos bajos a medios; Fosforo con intervalos de 3,7 a 19,7ppm, Potasio con intervalos 122 a 250ppm, capacidad de intercambio catiónico (CIC) con intervalos de 12,3 a 25,8 meq/100 g; carbonatos con valores menores a 1,1 %. La clase textural, son suelos francos, francos-limosos, franco-arcillo-limosos y franco-arcillo-arenosos: Concluyó que la mayoría de suelos son cercanos a la neutralidad, con menores riesgos de daños por exceso de sales. macroelementos químicos bajos a medios, clases texturales francos con variantes a arcillosos y arenosos, nivel bajo de fertilidad; con aptitud para el cultivar bananos. El estudio realiza la conceptualización de las variables evaluadas.

Caracteriticas fisicoquímicas del suelo

Textura. – Es el contenido de las particulas de arcilla, limo y arena presente en el suelo. Estas fracciones se analizan en los laboratorios para determinar la clase textural a la que pertenece (Dorronsoro) **[4]**. Hay una variabilidad de los diametros de las particulas del suelo, que toman valores de <0, 002 mm para arcilla; 0,05 – 0, 002 mm para limo y 2,00 – 0, 05 mm para arena; y la combinación de estas determina la clase textural del suelo (Medina) **[5]**

Densidad real (Dr). – Entendida como la relación de la masa de particulas de suelo seco presente y el volumen real que ocupan las particulas. Los valores de la (Dr) estan entre los intervalos de 2,6 a 2,75 g/cm3, este parámetro presenta poca variación en los suelos, es útil para determinar el nivel de porosidad (Dorronsoro) **[4]**. Castillo **[6]** considera que la (Dr) de los suelos estan en relación con el peso seco (Pss) y el volumen real de las particulas (Vp), que están expresadas en unidades g/cm3. Determinar el peso específico del suelo, va a permitir realizar el calculo de la porosidad, y conocerlo es importante para la agricultura ya que nos dara una orientación sobre el nivel o grado de desarrollo que tiene el suelo; asimismo, conoceremos la relación que existe entre los componentes tanto mineral como orgánico (Castillo) **[6]**. Los valores o intervalos que permiten conocer la densidad real de los suelos y clasificarlos estan entre >2,80 (Alto); entre 2,60 – 2,80 (medio) y <2,4 (Muy bajo) (Castillo) **[6]**

La densidad aparente (DA). – Se considera a la masa o peso seco del suelo inalterado, es decir obtenido de su disposición natural y el volumen total que incluye los espacios porosos (Castillo) **[6]**. Por lo que al interpretarlas el resultado podemos concluir que una densidad aparente alta, es un indicador que estamos frente a un suelo compacto, es decir contiene un nivel alto de particulas granulares entre las que destaca la arena, y una densidad aparente baja indica lo contrario; pero no necesariamente es señal de un suelo optimo para el desarrollo de las plantas (Castillo) **[6]**. Asimismo, (Flores y Alcala) **[7]**. Mencionan que determinando la (DA) se calcula la porosidad total y se alcanza a conocer la densidad de las particulas que es útil para realizar

C&T Riqchary
Revista de Investigación en Ciencia y Tecnología
ISSN: 2810-8124 (en línea) / ISSN: 2706-543x

Vol. 5 Núm. 1 (2023)
Publicado: 15/10/23
DOI: https://doi.org/10.57166/riqchary/v5.n1.2023.119

estimaciones de la masa de capa arable del suelo. También (Cumba et al) **[8]** considera que la densidad aparente es el procedimiento que permite identificar la relación entre la masa de los suelos secos y su volúmen bajo condiciones naturales. Por lo tanto, la (DA) (g/cm3) presenta variaciones con respecto a la textura de los suelos (Dorronsoro) **[4]**

pH del suelo. – Considerado como el indicador para medir la acidez o alcalinidad, su medida esta basada en concentraciones de hidrogeniones a una escala que va de 0 a 14, el valor de neutralidad es la escala 7 (Gardí et al) **[9]** y (Porta et al) **[10]**

Conductividad eléctrica (CE). - Es el parámetro o indicador para estimar la capacidad de la solucion de un suelo para transportar corriente eléctrica en relacion del contenido de sales disueltos en la solución. A traves de este indicador se estima los niveles de sales que estan disueltos en el suelo (Bazán) [11]. La salinidad esta referido a los solutos inorgánicos que se encuentran disueltas en agua siendo los esenciales Na+, Mg+2, Ca+2, K+, Cl-, SO4-2, HCO3-, y CO32- que estan presentes en mayor concentracion (Bazán) [11]. Conocer la CE, nos permitirá tomar decisiones sobre la conducción de nuestro suelo. Facilitará la decisión para instalar los cultivos y/o variedades a sembrar, que estará en función al nivel de tolerancia a las sales presentes en el suelo (INTAGRI) [12] y (Gutiérrez) [13]

Capacidad de campo (CC). – Es el valor o escala que tienen los suelos de retener agua o humedad de un suelo, luego de haber sido humedecido como producto de la lluvia o el riego y su posterior drenaje libre, evitando perdidas por evapotranspiración, hasta que el potencial hídrico del suelo logre estabilizarce (Flores y Alcalá) [7].

Porosidad. – Es el porcentaje o proporción del volumen de suelo que esta libre es decir no esta ocupado por fracciones solidas. Por lo general el volumen de un suelo lo conforman el 50% materia solida (5% materia organica y 45% minerales) y el 50% conformado por espacio poroso (Dorronsoro) [4]

Justificación
Conocer las caracteristicas físicas y químicas de los suelos es importante porque nos permitió tener una informacion real de la calidad, el contenido nutricional y las condiciones de conservación o degradación de los suelos de la comunidad de Huayllati. Esta información permitirá a las autoridades que promueven el desarrollo realizar acciones de planificación para recuperar la calidad y la fertilidad de los suelos. Ademas, esta información ayudara a identificar que actividades productivas debemos realizar y si es pertinente sembrar un cultivo, desarrollar una crianza o instalar frutales en nuestros suelos.

Al conocer las caracteristicas físicas y químicas se podrán proponer el uso y manejo pertinente de los suelos, conservando y recuperando su capacidad productiva, que incidirá favorablemente en la producción agropecuaria y por ende en mejorar las condiciones sociales y económicas de los sectores Topopampa, Padroma y Accahuaray en Huayllati.

Objetivo
Conocer las caracteristicas físicas y químicas de los suelos en Topopampa, Padroma y Accahuaray de Huayllati, Grau - Apurímac.

2 Método

2.1 Tipo y nivel de investigación

El estudio adoptó un enfoque cuantitivo, de tipo transversal y nivel descriptivo.

2.2 Diseño de la investigación

El estudio adopto un diseño no experimental; no se manipularon las variables, los datos fueron observadas y recojidas tal como se desarrollar en el medio. Todo el proceso se desarrolló sin intervención directa del investigador (Mousalli-Kayat) [14]

2.3 Población y muestra

La población estuvo constituida por terrenos de las familias de los sectores de Topopampa, Padroma y Accahuaray. Se hoyaron las calicatas y se recogieron las muestras y fueron llevadas al laboratorio para su análisis. El muestro utilizado fue el probabilistico sistematico estratificado. Las muestras se tomarón de los tres sectores garantizando la representación de cada estrato (Ñaupas et, al.) **[15]**

2.4 Analisis estadistico

Para el analisis estadistico previamente se realizarón las prueba de normalidad y homogeneididad de varianzas atraves del estadistico Shapiro Wilk y de Levene. Posteriorme se utilizo la estadistica descriptiva para responder al objetivo del studio a traves de tablas, en SPSS v25.

3 Resultados

Se obtuvieron los resultados de acuerdo a lo siguiente:

CARACTERISTICAS QUIMICAS

3.1 Conductividad eléctrica (CE)

TABLA 1
Conductividad eléctrica

Sector	Media	Máx	Mín	Desv. tipica
Accahuaray	0,08	0,11	0,05	0,03
Topopampa	0,14	0,20	0,10	0,04
Padroma	0,19	0,24	0,16	0,04

Los resultados de (CE) muestran que los tres sectores son suelos no salinos Medina, **[5].** Teniendo en Padroma valores de 0,19 mmhos/cm2; Topopampa 0,14 mmhos/cm2 y Accahuaray 0,08 mmhos/cm2.

3.2 Potencial de hidrogeniones (pH)

TABLA 2
Potencial de hidrogeniones (pH)

Sector	Media	Máx	Mín	Desv. tipica
Accahuaray	7,33	7,50	7,20	0,13
Topopampa	8,23	8,30	8,10	0,10
Padroma	8,38	8,50	8,20	0,13

Los resultados de pH que se obtuvieron en Padroma con valores de 8,38; Topopampa 8,23 y finalmente Accahuaray con 7,33. Los suelos cercanos a la escala de 7 de pH, son optimos para la producción de cultivos. Los niveles de pH de Padroma y Topopampa son relativamente alcalinos con presencia de problemas en las características físicas y deficiencia de nutrientes. Los suelos de Accahuaray son adecuados para la producción de cultivos Medina, [5]

3.3 Carbonato de calcio ($CaCo_3$)

TABLA 3
Carbonato de calcio

Sector	Media	Máx	Mín	Desv. tipica
Accahuaray	1,55	1,86	1,25	0,30
Topopampa	4,17	4,58	3,67	0,40
Padroma	4,59	4,90	4,32	0,26

Los resultados de ($CaCo_3$) muestran en Padroma una media de 4,59%, Topopampa 4,17% y Accahuaray 1,55%. El (CaCO3), influye en la absorción de micronutrientes. Los niveles altos de carbonatos es caracteriticas de los calcáreos y estan en relación con valores altos de pH (García) **[16]**

3.4 Materia orgánica (MO)

TABLA 4
Materia orgánica

Sector	Media	Máx	Mín	Desv. tipica
Accahuaray	1,74	2,46	1,10	0,58
Topopampa	0,77	0,93	0,63	0,13
Padroma	2,00	2,58	1,47	0,48

Los resultados observados en el contenido de (MO), en Padroma otuvo una media de 2,00%, Accahuaray 1,74% y Topopampa con 0,77%. El sector de Padroma tiene un nivel medio de (MO); Topopampa y Accahuaray, se encuentran ubicados en el nivel pobre (Medina) **[5]**

3.5 Nitrogeno (N)

TABLA 5
Nitrógeno

Sector	Media	Máx	Mín	Desv. tipica
Accahuaray	0,08	0,12	0,05	0,03
Topopampa	0,15	0,50	0,03	0,23
Padroma	0,28	0,80	0,07	0,35

Los resultados de (N) en Padroma se obtuvó una media de 0,28%, Topopampa 0,15% y Accahuaray con 0,08%. En Padroma hay un nivel alto de (N), Topopampa un nivel medio de (N) y Accahuaray un nivel bajo de (N) (Medina) **[5]**

3.6 Fósforo (P_2O_3)

TABLA 6
Fosforo

Sector	Media	Máx	Mín	Desv. tipica
Accahuaray	27,75	38,60	16,10	9,60
Topopampa	19,05	25,50	13,60	5,05
Padroma	14,58	17,00	11,20	2,64

Los resultados de (P_2O_3) en Accahuaray muestra una media de 27,75ppm, Topopampa 19,05ppm y Padroma 14,58ppm. Según (Medina) **[5]** los valores de Accahuaray y Topopampa obtuvierón un nivel alto de (P_2O_3) y Padroma un nivel bueno de (P_2O_3). El potasio es un nutriente esencial para la formación de la estructura y los frutos de las plantas.

C&T Riqchary
Revista de Investigación en Ciencia y Tecnología
ISSN: 2810-8124 (en línea) / ISSN: 2706-543x
Vol. 5 Núm. 1 (2023)
Publicado: 15/10/23
DOI: https://doi.org/10.57166/riqchary/v5.n1.2023.119

3.7 Potasio (K_2O)

TABLA 7
Potasio

Sector	Media	Máx	Mín	Desv. tipica
Accahuaray	156,00	175,00	144,00	13,44
Topopampa	112,00	116,00	105,00	4,97
Padroma	96,25	98,00	95,00	1,50

Los resultados de (K_2O) en Accahuaray muestran valores de 156,00ppm, Topopampa 112,00ppm y Padroma 96,25ppm. El (K_2O) es un macronutriente primario que logra influir en la calidad de los frutos, por lo que disponer de una cantidad optima de (K_2O) en los suelos, tendrá como efecto el incremento de los niveles de azucares en los frutos. Accahuaray y Topopampa tiene un nivel medio de (K_2O) y Padroma un nivel bajo de (K_2O) (Medina) [5]

3.8 Capacidad de intercambio catiónico (CIC)

TABLA 8
Capacidad de intercambio catiónico

Sector	Media	Máx	Mín	Desv. tipica
Accahuaray	13,30	16,06	10,17	2,80
Topopampa	9,46	11,23	8,02	1,49
Padroma	19,14	20,41	18,06	0,99

Los resultados de (CIC) en Padroma con 19,30 meq/100g, Accahuaray 13,30 meq/100g, y Topopampa 9,46 meq/100g. Los valores según (Medina) **[5]** en Padroma y Accahuaray tienen un nivel medio de CIC y Topopampa un nivel bajo de CIC.

CARACTERÍSTICAS FÍSICAS

3.9 Clase textural

TABLA 9
Clase textural

Sector	% Arena		% Limo		% Arcilla	
	Media	Desv. tipica	Media	Desv. tipica	Media	Desv. tipica
Accahuaray	49,50	3,11	26,00	2,94	24,50	5,07
Topopampa	55,25	11,67	27,50	6,67	17,25	6,09
Padroma	29,40	2,55	40,88	3,58	29,69	2,65

Los resultados de contenido de particulas en Accahuaray, con una media de 49,50% arena; 26,00% limo y 24,50% arcilla; Topopampa 55,25% arena, 27,50% limo 27,50% y 17,25% arcilla y Padroma 29,40% arena 40,88 limo y 29,69% arcilla. (Agroecología Tornos) **[17]**. Analizando los resultados en función al Triángulo Textural USDA, hay una variabilidad de clases texturales en los sectores, Topopampa franco-arenoso, Accahuaray franco-arcillo-arenoso y Padroma franco- arcilloso. Que nos llevo a concluir la existencia de variabilidad en los tres sectores.

3.10 Capacidad de campo (CC)

TABLA 10
Capacidad de campo

Sector	Media	Máx	Mín	Desv. tipica
Accahuaray	21,82	25,02	20,01	2,21
Topopampa	20,80	25,14	18,49	2,98
Padroma	29,20	30,71	28,12	1,19

Los resultados de (CC) en Accahuaray con 21,82%, Topopampa 20,80% y Padroma 29,20%. Segun (Flores y Alcalá) [7], los porcentajes de (CC) representan la cantidad de agua que puede retener un suelo, por lo que los resultados en los sectores tienen una relación con las clases texturales, descritas: Topopampa franco-arenoso; Accahuaray franco-arcilloso-arenoso y Padroma franco-arcilloso.

3.11 Densidad aparente (DA)

TABLA 11
Densidad aparente

Sector	Media	Máx	Mín	Desv. Típica
Accahuaray	1,50	1,56	1,43	0,05
Topopampa	1,53	1,59	1,47	0,05
Padroma	1,34	1,39	1,30	0,04

Los resultados de (DA) en Accahuaray con 1,50g/m3, Topopampa 1,53g/m3 y Padroma 1,34g/m3. (Flores y Alcalá) [7], la (DA) es el volúmen total incluyendo el espacio poroso y la masa del suelo seco. Hay una relación con las clases texturales de Topopampa franco-arenoso, Accahuaray franco-arcilloso-arenoso y Padroma franco-arcilloso.

3.12 Densidad real (Dr)

TABLA 12
Densidad real de los suelos

Sector	Media	Máx	Mín	Desv. tipica
Accahuaray	2,53	2,58	2,50	0,04
Topopampa	2,53	2,55	2,51	0,02
Padroma	2,54	2,56	2,51	0,02

Los resultados de (Dr) en Accahuaray con 2,53g/m3, Topopampa 2,53g/m3 y Padroma 2,54g/m3. (Castillo) [6] considera que la (Dr) esta en relacion con el el peso seco (Pss) del suelo y el volumen de sus partículas (Vp), expresado en g/cm3. Hay una relación con las clases texturales en Topopampa franco-arenoso, Accahuaray franco-arcilloso-arenoso y Padroma franco-arcilloso, se concluyo que los suelos en su mayoria estan clasificados

en el nivel bajo.

3.13 Porosidad

TABLA 13
Porosidad

Sector	Media	Máx	Mín	Desv. tipica
Accahuaray	40,13	42,80	37,84	2,18
Topopampa	38,86	41,20	36,40	2,24
Padroma	46,66	48,41	44,62	1,63

Los resultados de la porosidad en Accahuaray con 40,13%, Topopampa 38,86% y Padroma 46,66%. (Castillo) [6]. Establece que la porosidad son los espacios que no estan ocupados por las particulas o fracciones solidas. Hay una relación entre la porosidad y las clases texturales, Topopampa franco-arenoso en un nivel muy bajo, Accahuaray franco-arcilloso-arenoso en un nivel bajo y Padroma franco-arcilloso en un nivel medio.

4 Discusiones

Discusión de las características químicas de los suelos

La Conductividad eléctrica (CE), en Padroma, fue de 0,19 mmhos/cm2, Topopampa 0,14 mmhos/cm2 y Accahuaray 0,08 mmhos/cm2. La CE es un indicador del contenido de sales (milimhos/cm2) en los suelos. Medina [5], considera que los suelos con valores inferiores a <2.0 mmhos/cm2, son suelos no salinos. Es decir, los suelos son no salinos.

Potencial de hidrogenones (pH) en Padroma con 8,38; Topopampa con 8,23; y Accahuaray con 7.33. Al ser el pH un indicador que influye en la disponibilidad de nutrientes para el desarrollo de las plantas. Suelos con pH ubicado en las escalas extremas no son fértiles al no admitir nutrientes y no se recomienda cultivar en aquellos suelos (Medina) **[5]** Lo suelos de Padroma y Topopampa son relativamente alcalinos con presencia de problemas en las propiedades físicas y relativa deficiencia de nutrientes, por lo contrario, Accahuaray tiene suelos optimos para los cultivos.

Carbonato de calcio (CaCO3), en Padroma 4,59%; Topopampa 4,17% y Accahuaray 1,55%. El (CaCO3) tiene una influencia determinante para la absorción de los micronutrientes del suelo, con un nivel alto de carbonatos, los suelos adoptan las características de ser calcáreos y estan en relación con escalas altas de pH (García) [16]. Los suelos de Padroma y Topopampa tienen relativos niveles altos de (CaCo3), frente a Accahuaray que se ubica en un nivel medio, que es relativamente óptimo que permitirá la absorción de micronutrientes por los cultivos.

Materia orgánica (MO), en Padroma 2,00%; Accahuaray 1,74% y Topopampa 0,77%. La (MO) influye de manera determinante en la composición, la actividad microbiana, la estabilidad y disponibilidad de nutrientes del suelo para las plantas, asimismo la (MO) tiene la cualidad de mantener el pH estable, (Medina) [5]. El sector de Padroma, considerado con un nivel medio de (MO) y los sectores de Topopampa y Accahuaray, se encuentran ubicados en el nivel pobre (Medina) [5]

Nitrógeno (N) en Padroma con 0,28%; Topopampa 0,15% y Accahuaray 0,08%. El (N) constituye ser un macronutriente primario fundamental para las plantas, sobre todo de la estructura foliar (Medina) [5]. En Padroma hay un nivel alto de (N), Topopampa un nivel medio de (N) y Accahuaray un nivel bajo de (N).

Fosforo (P_2O_3), en Accahuaray con 27,75ppm, Topopampa 19,05ppm y Padroma 14,58ppm. El (P_2O_3) en los suelos, es un macronutriente primario, que es fundamental para el crecimiento adecuado de las plantas (Medina) [5] por lo que Accahuaray y Topopampa tiene un nivel alto de (P_2O_3) y Padroma un nivel bueno de (P_2O_3).

Potasio (K_2O), en Accahuaray con 156.00ppm, Topopampa 112.00ppm y Padroma 96.25ppm. El (K_2O) en los suelos, es un macronutriente primario que tiene influencia en la calidad de los frutos, un nivel optimo de (K2O) incrementa los niveles de azucares de los frutos. Accahuaray y Topopampa tiene un nivel medio de (K2O) y Padroma un nivel bajo de (K2O) (Medina) [5]

Capacidad de intercambio catiónico (CIC), en Padroma con 19.30 meq/100g, Accahuaray 13.30 meq/100g y Topopampa 9.46 meq/100g. (Medina) **[5]**, la (CIC) es un indicador de la las cualidades de los suelos para retener cationes, y esta en relacion con la disponibilidad y cantidad de nutrientes para las plantas. Los suelos con bajos niveles de (CIC) nos indican que tienen baja posibilidad de retener nutrientes, son suelos arenosos o pobres en materia orgánica (Medina) [5] en Padroma y Accahuaray tienen un nivel medio y Topopampa un nivel bajo.

Los resultados obtenidos estan en relación con los estudios de (Vivas) [3], que encontro (MO) con valores 0,92 a 2,2%. (N) con niveles bajos a medios; (P_2O_3) con valores de 3,7 a 19,7ppm.; (K_2O) en valores 122 a 250 ppm. Asimismo, (Maldonado) [1], encontró valores de pH que oscilan entre 4,07 a 6,85 considerados suelos ácidos cercanos a neutros. Contenido de sales con valores 0,85 a 5,37 (dS/m) suelos con caracteriticas relativamente salinos a moderados; Las bases intercambiables con valores de 0,7 a 32 cmol +)/kg-1, suelos con caracteriticas de bajos niveles de fertilidad; la (CIC) con valores 7,72 a 61,78 cmol (+)/kg-1, considerados suelos con bajos niveles de fertilidad. Por último (Gómez) [2], encontró valores de pH con promedios de 8,45, (CIC) promedios de 40 meq/100g, sin problemas de salinidad.

Discusión de las características físicas de los suelos

La textura es un indicador del tamaño y la proporción de las particulas presentes en los suelos (Agroecología Tornos) [17]. A traves del análisis de los suelos con el Triángulo Textural USDA, se identificaron la clase textural. Las clases de texturas

C&T Riqchary
Revista de Investigación en Ciencia y Tecnología
ISSN: 2810-8124 (en línea) / ISSN: 2706-543x

Vol. 5 Núm. 1 (2023)
Publicado: 15/10/23
DOI: https://doi.org/10.57166/riqchary/v5.n1.2023.119

que se identificarón en Accahuaray fuerón Franco-arcillo-arenoso; Topopampa franco-arenoso y Padroma franco-arcilloso

Capacidad de campo (CC), en Accahuaray 21,82%, Topopampa 20,80% y Padroma 29,20%. La (CC) del suelo es la representacion de la cantidad de agua que puede retener en contra de la fuerza de la gravedad, posterior al riego o la lluvia que humedecio el suelo (Flores y Alcalá) [7]. Los valores de (CC) establecen una relación con las clases texturales, descritas en Accahuaray franco-arcilloso-arenoso; Topopampa franco-arenoso y en Padroma franco-arcilloso

Densidad aparente (DA) en Accahuaray 1,50g/m3; Topopampa 1,53 g/m3 y Padroma 1,34g/m3. La (DA) representa la relacion de la masa del suelo seco y el volúmen total incluido el espacio poroso. Cuando se realiza el calculo de la porosidad total de un suelo, se puede estimar la masa de la capa arable (Flores y Alcalá) [7]. La (DA) tiene relacion con la clase textural de Accahuaray franco-arcilloso-arenoso; Topopampa franco-arenoso y Padroma franco-arcilloso.

Densidad real (Dr) en Accahuaray 2.53g/m3, Topopampa 2.53 g/m3 y Padroma 2.54g/m3. La (Dr) es la relación entre el peso del suelo (Pss) seco y el volumen real las partículas (Vp), que se expresan en g/cm3 (Castillo) [6]. La (Dr) tiene relación con la clase textural identificadas en Accahuaray franco-arcilloso-arenoso; Topopampa franco-arenoso y Padroma franco-arcilloso, clasificandose en el nivel bajo.

Porosidad de los suelos en Accahuaray con 40,13%; Topopampa 38.86% y Padroma 46.66%. El espacio libre que se encuentra entre las particulas solidas minerales u organicas se denomina porosidad del suelo (Castillo) [6]. La porosidad esta en relación con la clase textural de los suloes, tal es que en Accahuaray se tiene franco-arcilloso-arenoso, con un nivel bajo; Topopampa franco-arenoso con un nivel muy bajo y Padroma franco-arcilloso con un nivel medio.

Los resultados de la investigacion estan en relación con los estudios de Maldonado [1] encontrando en el Grado de saturación de bases en valores 5,63 a 96,7 %, evidenciando suelos acidos o con problemas de acidez. Tambien (Gómez) [2], en su estudio sobre la (DA) obtuvo 1,5 g/cm3 con compactación moderado. La (MO) con una media 0,5%. Suelos con textura arena, francos, franco-arcilloso, franco-limoso y franco-arenoso.

5 Conclusiones

Concluimos que las características químicas de los suelos de Accahuaray, Topopampa y Padroma de la comunidad de Huayllati – Grau – Apurímac, mostraron diferencias entre las variables evaluadas; observando que Padroma tiene mejores características químicas frente a los otros sectores.

Concluimos que las características físicas de los suelos de Accahuaray, Topopampa y Padroma de la comunidad de Huayllati – Grau – Apurímac, mostraron diferencias entre las variables, observándose que Topopampa presentó mejores características físicas frente a los otros sectores.

Agradecimientos

Los autores agradecen a los Comuneros de Accahuaray, Topopampa y Padroma de la comunidad de Huayllati – Grau – Apurímac, que brindaron su apoyo en la investigacion

Referencias

[1] Maldonado, "MaldonadoDeterminación de la aptitud de uso del suelo en la comunidad de Karhuiza La Paz," La Paz, 2016.

[2] Gómez, "Levantamiento de información agrológica de suelos dedicados a la producción agrícola en la parroquia el Cambio provincia de El Oro," El Oro, 2015.

[3] Vivas, "Caracterización y evaluación de los niveles de fertilidad en predios bananeros del valle del Chira región Piura," Piura, 2015.

[4] Dorronsoro, "Introducción a la Edafología," 2009.

[5] Medina, "Charla Técnica Virtual: "Interpretación de resultados del análisis de suelos"," INIA, Cusco, 2020.

[6] Castillo, "Selección y calibración de indicadores locales y técnico para evaluar la degradación de los suelos laderas, en la microcuenca Cuscamá El Tuma-La Dalia, Matagalpa," UNA, Matagalpa, 2005.

[7] Flores y Alcalá, "Manual de procedimientos analíticos. Laboratorio de Física de Suelos," UNAM, Mexico, 2010.

[8] Cumba et, al. "Propiedades morfológicas, físicas, químicas y mineralógicas de suelos del sur de Corrientes," Revista de la Asociacion Geologica Argentina, Buenos Aires, 2005.

[9] Gardi et, al. "Atlas de suelos de América Latina y el Caribe," 2013.

[10] Porta et, al. "Edafología para la agricultura y el medio ambiente," Madrid, 2003.

[11] Bazán, "Manual de procedimientos de los análisis de suelos y agua con fines de riego" 2017.

[12] INTAGRI, "La conductividad electrica del suelo en el desarrollo de los cultivos," 2022.

[13] Gutiérrez, "Geomorfología," UNAM, Mexico, 2008.

[14] Mousalli-Kayat, "Métodos y diseños de investigación cuantitativa," 2015.

[15] Ñaupas et, al. "Metodología de la investigación cuantitativa-cualitativa y redacción de la tesis," Ediciones de la U, Bogotá Colombia, 2014.

[16] Garcia, "Guía técnica: Fertilización en el cultivo de maíz blanco," UNALM-Agrobanco, Cusco, 2013.

[17] Agroecología-Tornos, "Cómo realizar un buen análisis del suelo para el cultivo. Extraído el 02 de noviembre del 2021 de: https://www.agroecologiatornos.com/como-realizar-un-buen-analisis-del-suelo-para-el-cultivo/," 2018.

[18] Cusi y. Aguirre, "Clasificación de tierras por su capacidad de uso mayor del distrito de Kosñipata-Paucartambo-Cusco," Cusco, 2015.

Biografías

- Norma Aguirre Quispe [A], Ingeniero Agroecologo Rural, Escuela Academico Profesional de Ingenieria Agroecologica y Desarrollo Rural de la Universidad Nacional Mi-caela Bastidas de Apurimac.

- Niki Franklin Flores Pacheco[B], Ingeniero Agronomo, adscrito al Departamento Academico de Ingenieria, Escuela Academico Profesional de Ingenieria Agroecologica y Desarrollo Rural de la Universidad Nacional Micaela Bastidas de Apurimac.

- Salvador Quispe Chipana[C], Doctor en Estadistica aplicada y Magister Scientiae en la especialidad: Economía grícola, docente nombrado a tiempo completo en la Escuela Profesional de Ingenieria Agropecuaria de la Universidad Nacional de San Antonio Abad del Cusco. Desarrolla la catedra de los cursos de Estadistica Agropecuaria, Diseños Experimentales y Economia Agricola.

- Maria Luisa Flores Pacheco[D], Ingeniero Agrónomo, Universidad Tecnológica de los Andes, Egresada de la Escuela de Post Grado – Maestría en Economía Mención Proyectos de Inversión Universidad Nacional San Antonio Abad del Cusco.