

INTERFACES



Revista digital de la Carrera de Ingeniería de Sistemas de la Universidad de Lima

Edición n.º 15 // Enero-julio 2022

Objetivo de *Interfases*

Interfases es una revista digital, gestionada por la Universidad de Lima, para la publicación de investigaciones originales en áreas temáticas relacionadas con las ciencias de la computación, gestión de datos, aprendizaje automático, sistemas de información, gestión de conocimiento y áreas afines. Se publican artículos científicos y avances de investigación, siempre que cumplan con el proceso de revisión por pares. La revista *Interfases* está indexada en CrossRef, Dialnet, Latindex y DOAJ: Directory of Open Access Journals, y se publica una vez al año, en diciembre. Sin embargo, a partir de julio del 2021, los manuscritos individuales se publicarán tan pronto como estén listos, añadiéndolos progresivamente al contenido de la edición en curso en la modalidad de publicación continua. Los artículos en publicación continua siguen el proceso de revisión por pares, y ya se pueden citar utilizando el año de publicación y el DOI.

Público objetivo

- Profesionales e investigadores relacionados con la ingeniería de sistemas y afines.
- Estudiantes y docentes del pregrado y posgrado de ingeniería de sistemas.
- Público interesado.

ISSN: 1993-4912 (electrónico)
doi: <https://doi.org/10.26439/interfases2022.n015>
Hecho el depósito legal en la Biblioteca Nacional
del Perú n.º 2020-09967

Periodicidad: Semestral
Edición: n.º 15
Fecha publicación: Julio del 2022
Arbitraje editorial: Revisión por pares
de doble ciego
Directorios y catálogos: CrossRef, Dialnet,
Latindex y DOAJ

Las opiniones expresadas en los artículos
firmados son de exclusiva responsabilidad de los
autores. Los contenidos de la revista *Interfases* son
de acceso abierto y se encuentran bajo la licencia
Creative Commons.

Universidad de Lima
Av. Javier Prado Este 4600,
Urbanización Monterrico Chico,
Lima 33, Perú.
Apartado postal 852, Lima 100, Perú.
Teléfono: (511) 437-6767, anexo
30131
fondoeditorial@ulima.edu.pe
<http://www.ulima.edu.pe/>

Edición, diseño y diagramación:
Fondo Editorial
Imagen de portada:

Carrera de Ingeniería de Sistemas
[http://www.ulima.edu.pe/pregrado/
ingenieria-de-sistemas](http://www.ulima.edu.pe/pregrado/ingenieria-de-sistemas)

Contacto: interfases@ulima.edu.pe

DIRECTOR

Julio Alejandro Padilla Solís
jpadilla@ulima.edu.pe

EDITOR

Hernán Nina Hanco
hninaha@ulima.edu.pe

EDITOR ASOCIADO

Marco Antonio Sotelo Monge
msotelo@ulima.edu.pe

EDITORES ASOCIADOS

Dr. Guillermo Antonio Dávila Calle
Universidad de Lima, Perú
gdavila@ulima.edu.pe

Ruth María Reategui Rojas, Ph. D.
Universidad Técnica Particular de Loja, Ecuador
rmreategui@utpl.edu.ec

COMITÉ EDITORIAL

Álvaro Talavera-López
ag.talaveral@up.edu.pe
Universidad del Pacífico, Lima, Perú

Carlos Mugruza-Vassallo
cmugruza@yahoo.com
Universidad Nacional Tecnológica del Cono Sur de Lima, Lima, Perú

César Beltrán-Castañón
cbeltran@pucp.pe
Pontificia Universidad Católica del Perú, Lima, Perú

Effie Lai-Chong Law
lai-chong.law@durham.ac.uk
Durham University, Inglaterra

Hugo Alatrasta-Salas
halatristas@up.edu.pe
Universidad del Pacífico, Lima, Perú

Ian D. Sanders
sandeid@unisa.ac.za
University of South Africa,, Pretoria, Sudáfrica

Juan Gutiérrez-Cárdenas
jmgutier@ulima.edu.pe
Universidad de Lima, Lima, Perú

Michael Dorin
michael_andrew.dorin@stud-mail.uni-wuerzburg.de
University of St. Thomas, Estados Unidos

Víctor Ayma-Quirita
vayma@ulima.edu.pe
Universidad de Lima, Lima, Perú

REVISORES

Dr. Alejandro Apaza Tarqui (Universidad Nacional del Altiplano Puno, Perú)
Dr. Isaias Bianchi (Al-Farabi Kazakh National University, Kazakhstan)
Mg. José Alberto Caballero Ortiz (Universidad de Lima, Perú)
Dr. Dennis Iván Candia Oviedo (Universidad Nacional de San Antonio Abad del Cusco, Perú)
Mg. Daniel Enrique Cárdenas Salas (Universidad de Lima, Perú)
Mg. William Alberto Chávez Espinoza (Universidad Pública de Navarra, España)
Dra. María de León Sigg (Universidad Autónoma de Zacatecas, México)
Dr. Ignacio Diaz-Cano (Universidad de Cádiz, Andalucía, España)
Dra. Simena Dinas (Universidad de San Buenaventura, Cali, Colombia)
Dr. Dario Francisco Dueñas Bustinza (Universidad Nacional Federico Villarreal, Perú)
Dr. Edwin Jonathan Escobedo Cárdenas (Universidade Federal de Ouro Preto, Brasil)
Mg. Jisbaj Gamarra Salas (Universidad de Granada, España)
Dr. Juan Gutiérrez Cárdenas (Universidad de Lima, Perú)
Mg. Jorge Luis Irey Nuñez (Universidad de Lima, Perú)
Mg. Cesar Stuardo Lucho Romero (Pontificia Universidad Católica del Perú)
Mg. William-Rogelio Marchand-Niño (Universidad Nacional Agraria de la Selva, Perú)
Mg. Pilar Alexandra Moreno (Universidad Nacional Abierta y a Distancia, Colombia)
Mg. Braulio Oscar Murillo Veliz (Pontificia Universidad Católica del Perú)
Mg. Iván Darío Peñaranda Arenas (Universidad de Granada, España)
Mg. Ferdinand Edgardo Pineda Ancco (Pontificia Universidad Católica del Perú)
Dr. José Antonio Pow Sang Portillo (Pontificia Universidad Católica del Perú)
Mg. Hernán Alejandro Quintana Cruz (Universidad de Lima, Perú)
Mg. Lennin Paul Quiroz Villalobos (Universidad de Lima, Perú)

Mg. Lourdes Ramirez Cerna (Universidade Federal de Ouro Preto, Brasil)

Mg. Liseth Urpy Segundo Carpio (Universidade de São Paulo, Brasil)

Dr. Ángel Leonardo Valdivieso Caraguay (Escuela Politécnica Nacional, Quito, Ecuador)

Dr. Ronny Villafuerte Serna (Universidad Nacional de San Antonio Abad del Cusco, Perú)

Mg. Rannoverng Yanac Montesino (Universidad Nacional Agraria de la Selva, Perú)

POLÍTICA EDITORIAL

ENFOQUE Y ALCANCE

Interfases es una revista digital, gestionada por la Universidad de Lima, para la publicación de investigaciones originales en áreas temáticas relacionadas con las ciencias de la computación, gestión de datos, aprendizaje automático, sistemas de información, gestión de conocimiento y áreas afines. Se publican artículos científicos y avances de investigación, siempre que cumplan con el proceso de revisión por pares. La revista *Interfases* está indexada en CrossRef, Dialnet, Latindex y DOAJ: Directory of Open Access Journals, y se publica una vez al año, en diciembre. Sin embargo, a partir de julio del 2021, los manuscritos individuales se publicarán tan pronto como estén listos, añadiéndolos progresivamente al contenido de la edición en curso en la modalidad de publicación continua. Los artículos en publicación continua siguen el proceso de revisión por pares, y ya se pueden citar utilizando el año de publicación y el DOI.

PROCESO DE REVISIÓN POR PARES

Los manuscritos originales e inéditos enviados a la revista *Interfases* siguen un proceso de evaluación en dos etapas.

En la primera, el editor examina el contenido para determinar si el manuscrito está alineado con el alcance y ha seguido las directrices para autores. Si el manuscrito no es aceptado, se devuelve al autor correspondiente con las razones detalladas que motivan la decisión adoptada por el editor.

Si el manuscrito es aceptado por el editor, este se envía a revisores externos expertos en el tema de investigación. Esta segunda evaluación corresponde a una revisión por pares doble ciego, donde el autor y revisores son anónimos.

El revisor evalúa el contenido del manuscrito y, basándose en su experiencia y conocimiento, adopta una de las siguientes recomendaciones:

1. El manuscrito es aceptado sin cambios o con cambios mínimos.
2. El manuscrito se acepta, a condición de realizar cambios importantes, de acuerdo con las observaciones del revisor. La versión corregida del manuscrito debe ser aprobada en una segunda revisión.
3. El manuscrito no se acepta por las contribuciones limitadas del estudio u otras consideraciones informadas por el revisor.

Con base en los comentarios de los revisores, el editor informa la decisión al autor correspondiente, quien tiene hasta treinta días para realizar los cambios al manuscrito (recomendación 1 y 2) o argumentar el porqué no se acepta (recomendación 3).

Una vez que los revisores reciben el manuscrito corregido, tienen hasta veinte días para informar el resultado de la nueva evaluación; posteriormente, emiten su recomendación final. Una vez que el editor recibe la segunda ronda de revisiones, toma una decisión para publicar el manuscrito y luego se le notifica al autor correspondiente.

Cualquier objeción del autor con respecto a la decisión del editor o hacia los comentarios de los revisores será resuelta por el Comité Editorial como instancia final.

La revista se adhiere a los criterios establecidos por el Guidelines on Good Publication Practice del Committee on Publication Ethics (COPE), el cual establece las sanciones en caso de plagio.

DIRECTRICES PARA AUTORES/AS

ENVÍO DEL MANUSCRITO

Interfases publica tres tipos de artículos: trabajos de investigación (hasta 5000 palabras), avances en investigación (hasta 2800 palabras) y revisiones (hasta 1500 palabras).

Todos los artículos se envían del mismo modo. Una vez que el editor verifique que el contenido del manuscrito pertenece al ámbito de *Interfases* lo remitirá a un proceso de revisión por pares. Este proceso (compuesto de dos rondas) toma aproximadamente de dos a tres meses, pero dependiendo de la complejidad del manuscrito, podría extenderse.

Los manuscritos enviados a *Interfases* no deben haberse publicado previamente ni estar en consideración para su publicación en otra revista.

Página del título

La página del título debe incluir:

- Un título conciso e informativo (hasta 30 palabras).
- El nombre completo de cada autor, incluyendo la afiliación institucional, la dirección de correo electrónico y el código ORCID.
- Resumen de 200-250 palabras. El resumen debe indicar la naturaleza y contribución del estudio. Evite las abreviaturas no definidas, las ecuaciones matemáticas o las referencias bibliográficas en el texto del resumen.
- Palabras clave (de 3 a 5) separadas por comas. Las palabras clave deben tomarse de la taxonomía de la IEEE Computer Society: <https://www.computer.org/digital-library/journals/sc/tsc-taxonomy-list>

Texto

Los trabajos enviados deben haber sido redactados en un documento Word (.doc o .docx), y aquellos aceptados para ser publicados deben usar la plantilla de *Interfases* LATEX que estará disponible pronto.

Al redactar el manuscrito, usar la opción de numeración automática para enumerar las páginas. Por favor, evite el uso de funciones de campo. Utilice la función de tabla, no una hoja de cálculo pegada, para hacer tablas. Si escribe su manuscrito con Word, use el editor de ecuaciones o MathType para las ecuaciones.

Tablas

Las tablas son el núcleo de los nuevos hallazgos reportados en la corriente principal de la ciencia, por lo tanto, incluya las tablas que considera son estrictamente necesarias. Todas las tablas se enumeran utilizando números arábigos (por ejemplo, Tabla 1, Tabla 2, ...) e incluyen un título que detalla la relevancia de los datos presentados.

Las tablas se mencionan en el orden en que aparecen en el manuscrito. Además del número, el título y los datos, las tablas pueden incluir una nota para detallar la fuente de información, así como explicaciones adicionales que no están incluidas en el manuscrito.

Abreviaturas

Use abreviaturas solo si son necesarias para mejorar la legibilidad de su documento. Debe definir cada abreviatura en la primera mención después de usarla de manera consistente.

Conclusiones

Recuerde que las conclusiones no son la versión narrativa y textual de las tablas incluidas en la sección Resultados. Por el contrario, las conclusiones reseñan y sintetizan los principales argumentos del artículo. Las conclusiones se extraen de los hallazgos y proporcionan una respuesta adecuada a la pregunta de investigación. Además, las conclusiones incluyen las limitaciones del estudio y sugieren nuevas preguntas y aplicaciones para futuros estudios.

Referencias

Las citas y las referencias deberán indicarse de acuerdo con las normas APA. Según la norma señalada, las referencias, enlistadas al final de la publicación, se realizarán de la siguiente forma:

a. Libros

Apellido Autor, A. A., & Apellido Autor, A. A. (Año de la publicación). *Título del libro* (número de la edición). Nombre de la editorial. DOI o URL

b. Capítulos de un libro

Apellido Autor, A. A., & Apellido Autor, A. A. (Año de publicación). Título del capítulo. En E. E. Editor (Ed.), *Título del libro* (número de edición, volumen, pp. 1-6). Nombre de la editorial. DOI o URL

c. Artículos de revistas

Apellido Autor, A. A., & Apellido Autor, A. A. (Año de publicación). Título del artículo. *Nombre de la revista, volumen*(número de la revista), número de página inicial-número de página final. DOI o URL

d. Libros electrónicos

Apellido Autor, A. A., & Apellido Autor, A. A. (Año de publicación). *Título del texto electrónico*. Nombre de la editorial. DOI o URL

e. Artículos de revistas electrónicas

Apellido Autor, A. A., & Apellido Autor, A. A. (Año de publicación). Título del artículo. *Nombre de la revista, volumen*(número de la revista), número de página inicial-número de página final. DOI o URL

f. Ponencias en congresos o simposios

Apellido Autor, A. A., & Apellido Autor, A. A. (Año de publicación, día de mes). Título de la contribución [Tipo de la contribución]. *Nombre del congreso, ubicación*. DOI o URL Material suplementario electrónico

Material suplementario electrónico

Los autores pueden incluir archivos de texto (incluyendo tablas y figuras) y hojas de cálculo como material complementario. Sin embargo, para datos de investigación, es recomendable archivarlos en repositorios de datos. Para el código de *software*, los autores pueden usar plataformas como GitHub o similares.

Si los originales contienen fotografías o reproducciones de obras pictóricas, estas se entregarán aparte en archivos TIFF o JPG, con 300 píxeles de resolución (dpi). Si contienen gráficos, cuadros, dibujos, flujogramas u otros elementos, estos deben entregarse igualmente en un archivo aparte y en el programa original en que fueron creados (por ejemplo: Excel, Illustrator, etcétera).

Lista preliminar para la preparación de envíos

Los artículos deberán respetar el siguiente formato:

- a. Formato de página A4.
- b. Título del artículo, centrado en negrita, con letra Times New Roman en doce puntos.
- c. Títulos del texto, en negrita, con letra Times New Roman en doce puntos, dejando dos líneas en blanco antes del párrafo.
- d. Texto del cuerpo con letra Times New Roman en doce puntos, con espacio y medio de interlineado.

Declaración de privacidad

Los nombres y las direcciones de correo electrónico introducidos en esta revista se usarán exclusivamente para los fines establecidos en ella y no se proporcionarán a terceros o para su uso con otros fines.

ÍNDICE

PRESENTACIÓN	15
ARTÍCULOS DE INVESTIGACIÓN	
Propuesta metodológica para la identificación de expertos en el ámbito organizacional	17
<i>Álvaro Antonio Aures García, Adelaida Estela Balvín Landeo</i>	
A Distributed Model for Computing 3D Mesh Local Descriptors Based on <i>k</i> -Rings	38
<i>Franci Suni-Lopez, Jan Hurtado, Alejandra Márquez, Leonardo Guzmán</i>	
Football Pitch Condition Analysis Based on <i>k</i> -Means Clustering	56
<i>Héctor Eduardo Ugarte Rojas, Boris Chullo Llave</i>	
Estudio de la usabilidad de una aplicación de realidad aumentada en la construcción básica de un brazo robótico	70
<i>Jose Malpartida Valverde, Jose Martin Lozano Aparicio</i>	
Aplicación de métodos de <i>deep learning</i> en series de tiempo para el pronóstico de la situación macroeconómica en América Latina	102
<i>Víctor Augusto Alegre Ibáñez, Jose Martin Lozano Aparicio</i>	
Clasificación de organismos en reinos utilizando frecuencia de codones de ADN	131
<i>Luis Beltrán Palma Ttito</i>	
ARTÍCULOS DE REVISIÓN	
Ciencia de datos en la evaluación del impacto de las políticas públicas: una revisión de la literatura	144
<i>Ganimedes T. Rosales Reyes</i>	

Aplicación de <i>machine learning</i> en la gestión de riesgo de crédito financiero: una revisión sistemática	160
<i>Juler Anderson Hermitaño Castro</i>	
DATOS DE LOS AUTORES	178

PRESENTACIÓN

Uno de los objetivos de la revista *Interfases* es promocionar la comunicación científica en áreas temáticas relacionadas a ingeniería de sistemas, ciencia de la computación, ingeniería del *software*, sistemas de información, tecnologías de información, ciberseguridad, ciencia de datos y áreas afines. En ese sentido, presentamos la edición número 15 correspondiente al periodo enero-julio del año 2022. En la presente edición se incluyen ocho artículos que corresponden al 42 % de los manuscritos recibidos entre enero y junio del presente año y lograron pasar por un riguroso proceso de revisión por pares.

En primer lugar, resaltamos el aporte de los autores Álvaro Antonio Aures García de la Universidad Peruana de Ciencias Aplicadas y Adelaida Estela Balvín Landeo de la Universidad Nacional Mayor de San Marcos, quienes proponen una metodología para la identificación de expertos en el ámbito organizacional como parte de la línea de investigación en la gestión del conocimiento. En segundo lugar, los autores Franci Suni-López, Jan Hurtado, Alejandra Márquez y Leonardo Guzmán, todos ellos de la Universidad Nacional de San Agustín de Arequipa, presentaron un modelo distribuido para calcular descriptores locales de malla 3D basados en k-rings, una contribución al campo del procesamiento de objetos 3D. En tercer lugar, se recibió la investigación de los profesores de la Universidad Nacional de San Antonio Abad del Cusco, Héctor Eduardo Ugarte Rojas y Boris Chullo Llave, quienes desarrollan una estrategia para evaluar las condiciones del estado del campo de fútbol basado en el análisis de imágenes usando algoritmos de agrupamiento de *K-means*. En cuarto lugar, recibimos un importante trabajo sobre evaluación de la usabilidad de una aplicación de realidad aumentada para contribuir al conocimiento básico de la construcción de un brazo robótico presentado por los autores Jose Malpartida Valverde y Jose Martin Lozano Aparicio de la Universidad Nacional de Ingeniería. En quinto lugar, los investigadores de la Universidad Nacional de Ingeniería, Víctor Augusto Alegre Ibáñez y Jose Martin Lozano Aparicio presentan el trabajo sobre aplicación de métodos de *deep learning* en series de tiempo para el pronóstico de la

situación macroeconómica en América Latina. Dicho trabajo evalúa el PBI de Argentina, Brasil, Chile, Colombia, México y Perú para construir modelos de predicción usando los algoritmos de perceptrón multicapa, red neuronal convolucional, LSTM, GRU y SeqToSeq. En el sexto lugar, resaltar el trabajo del profesor Luis Beltrán Palma Ttito de la Universidad Nacional de San Antonio Abad del Cusco que presenta su investigación sobre clasificación de organismos en reinos utilizando frecuencia de codones de ADN. En séptimo lugar, recibimos el trabajo de revisión sistemática sobre la ciencia de datos en la evaluación del impacto de las políticas públicas presentado por el autor Ganimedes Trimegistro Rosales Reyes, catedrático en la Universidad Nacional Mayor de San Marcos. En octavo lugar, publicamos el artículo de revisión sistemática sobre la aplicación de *machine learning* en la gestión de riesgo de crédito financiero, desarrollado por el autor Juler Anderson Hermitaño Castro de la Universidad de Lima.

Finalmente, expresamos nuestro enorme reconocimiento a todos los investigadores que sometieron a revisión sus manuscritos en la presente edición de *Interfases*. En ese sentido, valoramos profundamente vuestro esfuerzo y dedicación al desarrollo científico que en esta oportunidad nos convoca la revista *Interfases*. También agradecer a nuestro equipo editorial y revisores quienes constantemente promueven la mejora continua de la revista asegurando la calidad del material publicado.

Dr. Julio Alejandro Padilla-Solís
Director de la revista *Interfases*

Recibido: 24/2/2022 / Aceptado: 17/4/2022

doi: <https://doi.org/10.26439/interfases2022.n015.5789>

PROPUESTA METODOLÓGICA PARA LA IDENTIFICACIÓN DE EXPERTOS EN EL ÁMBITO ORGANIZACIONAL*

ÁLVARO ANTONIO AURES GARCÍA

pcsiaaur@upc.edu.pe / ORCID: 0000-0003-2433-5094

Universidad Peruana de Ciencias Aplicadas (UPC), Lima, Perú

ADELAIDA ESTELA BALVÍN LANDEO

adelaida.balvin@unmsm.edu.pe / ORCID: 0000-0002-9869-3858

Universidad Nacional Mayor de San Marcos (UNMSM), Lima, Perú

Resumen

En este artículo se presenta una metodología para la identificación de expertos que puedan ayudar a organizaciones de cualquier tipo a mejorar su competitividad y a facilitar la innovación en la gestión de proyectos. La propuesta nace de la constatación de los vacíos y entrapamientos que se generan cuando el conocimiento experto no se comunica al interior de las organizaciones. La metodología propuesta parte de la identificación, a partir de una taxonomía, de los campos de conocimiento para los cuales una organización requiere de un experto. Luego de ello, propone criterios para la identificación y valoración de los expertos a partir de conceptos de gestión del conocimiento. Esta propuesta se elaboró a partir del trabajo con diversas instituciones públicas, todas rectoras en sus respectivos sectores, y encargadas de la realización de proyectos de gran envergadura.

PALABRAS CLAVE: gestión del conocimiento / conocimiento experto / páginas amarillas de expertos

Este artículo es una extensión del artículo publicado en Aures García, A. A., & Balvín Landeo, A. E. (2021). Identificación de expertos como fuente de conocimiento en las organizaciones. *Actas del Congreso Internacional de Ingeniería de Sistemas*, 83-96. <https://doi.org/10.26439/ciis2021.5579>

Methodological proposal for the identification of experts in the organizational field

Abstract

This article presents a methodology for identifying experts who can help any kind of organization to improve their competitiveness and facilitate innovation in project management. The proposal stems from verifying the gaps and traps generated when expert knowledge is not shared within organizations. Based on a taxonomy, the proposed methodology identifies the fields of knowledge for which an organization requires an expert. After that, it proposes criteria for identifying possible experts and assessing their expertise based on knowledge management concepts. This proposal was developed from working with various public institutions, all governing in their respective sectors, and carrying out large-scale projects.

KEYWORDS: knowledge management / expert knowledge / expert yellow pages

1. INTRODUCCIÓN

La conjunción de la gestión del conocimiento y las tecnologías de la información se ha convertido en la actualidad en un factor diferenciador y de valor en la producción y la competitividad de las organizaciones porque le permiten localizar y transferir el conocimiento, entre los diversos actores que componen su cadena productiva, mejorando e innovando los productos y/o servicios que ofertan al mercado global. Esto ha conllevado a que muchas organizaciones incorporen, paulatinamente, en su gestión herramientas y métodos para identificar y organizar los dominios de conocimientos existentes en sus diferentes ámbitos funcionales con la finalidad de posibilitar que cualquier miembro y/o equipo de proyecto interactúe y transfiera el conocimiento de manera rápida y oportuna.

Con base en lo mencionado, el artículo propone una metodología para identificar expertos con coeficientes de competencias sobresalientes o altos en el marco de la gestión de proyectos de cualquier organización. La identificación de este tipo de profesional se circunscribe a través de un enfoque metodológico constituido por tres etapas: a) el diagnóstico situacional cognitivo que determina los aspectos claves que gobiernan la localización, adquisición y transferencia del conocimiento individual y grupal con el experto; b) la taxonomía que identifica las temáticas de poco dominio o de falta de conocimiento por los diversos actores de una organización, y c) la valoración del conocimiento experto con base en un estudio de coeficientes de competencias que permite a las organizaciones identificar a los expertos con trayectoria y dominio en un determinado tema, siendo reconocidos por sus pares, y cualificados para brindar información objetiva sobre una determinada consulta.

Es importante señalar que la identificación del experto está cimentada en el desarrollo de la metodología propuesta, que involucra los conceptos de la gestión del conocimiento y su relación de trabajo en el marco de la gestión de proyectos en las diversas instituciones públicas, que fueron objeto de estudio en la presente investigación.

2. MARCO TEÓRICO

Uno de los puntos focales en el desarrollo de la investigación se relaciona con el análisis del material bibliográfico referido a la gestión del conocimiento en el ámbito organizacional, desde modelos vinculados con el aprendizaje colaborativo hasta el desarrollo de redes de expertos, con la finalidad de validar la importancia de la localización del conocimiento clave en la organización para solucionar los problemas de manera efectiva, generando, de esta manera, ventajas competitivas en su espectro de actuación.

2.1 Contexto de la gestión del conocimiento

La importancia creciente de aplicar estrategias globalizadas, con base en el entorno competitivo, requiere constantemente que las organizaciones renueven sus flujos de

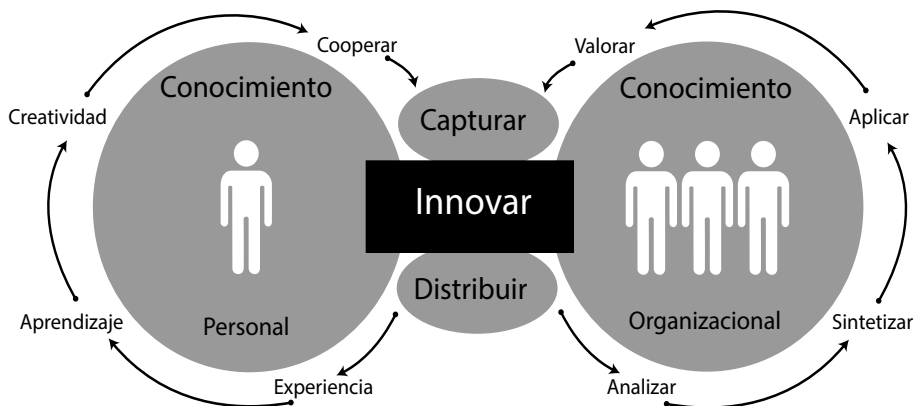
información y conocimiento en sus actividades productivas con la finalidad de innovar en su gestión desde diferentes aspectos. Según Bueno, citado por Archibold y Escobar (2015), las diferencias competitivas de las organizaciones se están cimentando en gran medida en las interacciones sociales que realizan los profesionales para obtener nuevos conocimientos con sus pares a efectos de solucionar determinados problemas en su ámbito laboral, aumentando así paralelamente su capacidad cognitiva.

Es importante manifestar que los inicios de la gestión del conocimiento se enmarcaron principalmente en las organizaciones de índole global debido a que muchas de estas utilizaron las tecnologías de la información (TI) para evolucionar y adaptarse rápidamente a las nuevas necesidades de su entorno competitivo en función al conocimiento de valor que incorporaron a los productos, bienes y servicios que ofertaban al mercado (Pérez Arrau, 2016).

Para Sveiby, Gripenberg y Segercrantz (2012), la innovación se logra a través de un conocimiento accesible en los diferentes niveles organizacionales porque a partir de este elemento intangible se generan nuevas competencias y beneficios competitivos en mayor escala. Es decir, la innovación está fuertemente condicionada por la capacidad de cooperar abiertamente, valorar la experiencia, aplicar el aprendizaje y la forma en que se tiene organizado el conocimiento dentro de una organización.

Figura 1

La colaboración en el marco de la gestión de proyectos



Nota. De *Routledge Studies in Technology, Work and Organizations: Challenging the Innovation Paradigm*, por K.-E. Sveiby, P. Gripenberg, y B. Segercrantz (Eds.), 2012, Routledge.

Del mismo modo Wiig, citado por Mohajan (2017), hace referencia que el conocimiento consiste en identificar y suscribir adecuadamente los hechos, conceptos, perspectivas y metodologías que se acumulan e integran en largos periodos de tiempo para su disponibilidad efectiva ante problemas y situaciones específicas. En el caso de una organización, el autor menciona que el conocimiento creado, utilizado y transferido adecuadamente entre los profesionales tiene un impacto significativo en sus fronteras competitivas debido a que genera toma de decisiones adecuadas para su inserción exitosa a nuevos mercados.

En términos generales, los autores enfatizan que las organizaciones actualmente están recurriendo a estrategias que sistematizan y capturan el conocimiento generado en su entorno para reducir las brechas competitivas y productivas con sus competidores, haciendo uso de las tecnologías de la información y respondiendo con efectividad las demandas de los clientes y consumidores en diversos periodos de tiempo.

2.2 Modelos de la gestión del conocimiento

Existen una variedad de modelos de gestión del conocimiento en el contexto de la literatura internacional; sin embargo, para la propuesta metodológica de identificación de expertos en las organizaciones (Aures García & Balvín Landeo, 2021) se tomaron como referencia los siguientes puntos.

2.2.1 Modelo de Karl Wiig

Según Wiig, citado por Mohajan (2017), las organizaciones deben analizar, mapear y considerar tres formas de conocimiento para mantener una competitividad sostenible en el tiempo: conocimiento público, conocimiento experto compartido y conocimiento personal. En el caso del conocimiento experto compartido, se hace indispensable que las organizaciones utilicen y transfieran el conocimiento de valor, obtenido de los expertos, a todos los ámbitos funcionales de la organización por medio de las plataformas tecnológicas, como, por ejemplo, las páginas amarillas de expertos, con la finalidad de tomar decisiones con base en una información validada y de calidad.

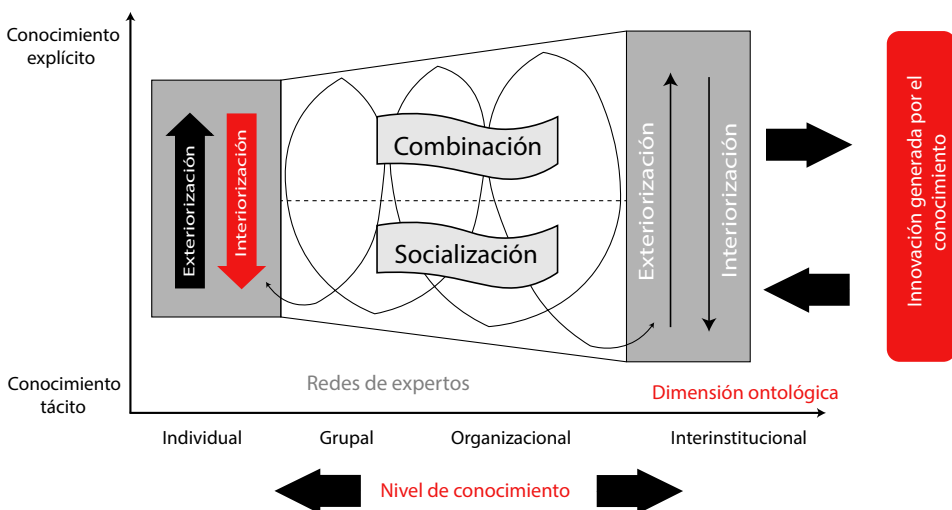
En el Perú, el conocimiento experto no es mapeado ni sistematizado adecuadamente en muchas organizaciones debido al desconocimiento de metodologías y/o conceptos propios de la gestión del conocimiento, razón por la cual se ha podido percibir en la investigación que las instituciones públicas, en general, no implementan herramientas tecnológicas específicas para gestionar e identificar el conocimiento experto. En muchos casos, desconocen al experto que labora y apoya dentro de la institución.

2.2.2 Modelo de Nonaka y Takeuchi

El modelo hace una representación de los procesos de transformación del conocimiento, en sus diferentes etapas: socialización, exteriorización, combinación e interiorización; que son vistas y articuladas generalmente por los profesionales que realizan actividades de investigación y/o proyectos con sus pares dentro y fuera de una organización (Nonaka & Takeuchi, 1995). Sin embargo, se ha observado también que las instituciones públicas, en muchos casos, no están clasificando e indexando el conocimiento a través de taxonomías y/o tesauros para su exposición y el aprendizaje colectivo.

Figura 2

Modelo de la gestión del conocimiento de Nonaka y Takeuchi



Nota. Adaptado de *The knowledge-creating company. How Japanese Companies Create the Dynamics of Innovation*, por I. Nonaka, y H. Takeuchi, 1995, Oxford University Press.

Basados en los principios de Nonaka y Takeuchi, Arambarri Basañez (2014) señala que hoy en día las organizaciones competitivas tienen los objetivos de elaborar e implementar continuamente procedimientos de gestión del conocimiento y entornos colaborativos abiertos que visen la viabilidad de trabajar en proyectos geográficamente dispersos, con equipos externos, a efectos de crear, gestionar y transferir el conocimiento en un único espacio.

2.2.3 Modelo de Arthur Andersen

Modelo citado por Flores López y Ochoa Jiménez (2016), el cual circunscribe un conjunto de factores estratégicos de la gestión del conocimiento para fortalecer el ámbito de acción de una organización que aprende y que busca la innovación dentro y fuera de sus fronteras competitivas. Asimismo, los autores hacen hincapié de que las organizaciones en primera instancia deben focalizarse en la construcción de comunidades de práctica o desarrollar, en efecto, las capacidades cognitivas entre las personas y/o equipos para generar nuevos conocimientos antes de invertir en nuevas tecnologías de la información.

Es importante mencionar, que los autores reconocen, dentro de los principios del modelo de Arthur Andersen, la aplicación de infraestructuras tecnológicas para capturar, analizar y distribuir el conocimiento a nivel organizacional con la finalidad de mejorar y fortalecer la capacidad productiva y de innovación en todos sus niveles de actuación.

2.2.4 Modelo de Choo

Para Choo (2016), el conocimiento es un componente que a través del tiempo genera valor para una organización inteligente porque resalta el uso adecuado de las lecciones aprendidas y la toma de decisiones, con base en datos e informaciones, como pilares fundamentales en la innovación. Este modelo se centra, especialmente, en la selección e incorporación de fuentes de información y de conocimientos propios de la organización, y otras asimiladas de su entorno competitivo, en repositorios tecnológicos para su explotación y uso.

Lograr que estas interrelaciones se articulen, adecuadamente, podría conducir a la organización a mejorar sus decisiones con base en un conocimiento equilibrado y actualizado, producto de una información procesada e indexada de manera eficaz.

2.2.5 Modelo de Holsapple y Joshi

El modelo de Holsapple y Joshi, citado por Marin-García y Zarate-Martínez (2008), hace hincapié en la unificación de los diferentes enfoques de la gestión del conocimiento basado en tres dimensiones fundamentales: a) los recursos del conocimiento, b) las actividades inherentes a la gestión del conocimiento y c) los factores de influencia.

En relación con la primera dimensión, los autores mencionan el conocimiento de las personas como fuentes esenciales de productividad, competitividad e innovación en las tareas funcionales de la organización y que estos pueden ser capitalizados a través de las tecnologías de la información como un elemento dominante para transferirlos a todos los ámbitos que se requieran, ya sea por portales corporativos, redes de expertos, entre otros (Alavi y Leidner, 2001).

2.3 Conocimiento experto en las organizaciones

El científico Niels Bohr definió al experto como la persona que, a lo largo de su trayectoria profesional y académica, ha cometido errores en su campo de estudio o de trabajo, pero ha salido fortalecido de ello y ha aprendido de los mismos. Por ello, la competencia experta adquirida por una persona es le permite orientar y solucionar determinados problemas en cualquier organización.

Sin lugar a dudas, el experto es aquella persona identificable por sus conocimientos que destaca entre un conjunto de profesionales y es considerado por sus pares como una fuente confiable de habilidades y experiencias para asesorar casos específicos de acuerdo a su espectro de actuación. Para Robles y Rojas (2015), los expertos se caracterizan, principalmente, por complementar habilidades y conocimientos especializados que fueron obtenidos a través de un conjunto de errores y prácticas exitosas sobre una actividad en particular.

Para las organizaciones, y en especial para los equipos de proyectos, el juicio de expertos es una de las herramientas que ayuda a definir estrategias y acciones de competitividad en relación al desarrollo de nuevos productos y servicios que desean brindar a un mercado global en toda su extensión. La fiabilidad de la opinión de una persona, reconocida como experto por sus pares, generalmente coadyuva a definir los límites de acción de una estrategia, mejorar los procesos productivos o desarrollar nuevas soluciones innovadoras en favor de la organización, manteniéndola continuamente en un espacio altamente competitivo.

Para Cabero Almenara y Llorente (2013), el conocimiento de los expertos permite que las organizaciones tengan una amplia información de valor sobre las temáticas de estudio que afectan, en muchos casos, su competitividad. Para mantenerse en un espacio altamente competitivo, las organizaciones utilizan una serie de criterios estructurados para priorizar, catalogar y sistematizar el conocimiento experto sobre determinados temas con evidencias, juicios y valoraciones que les permitan tomar decisiones adecuadas y objetivas basadas en sus aprendizajes.

Tomando como premisa lo señalado por los autores, la investigación denotó que las organizaciones públicas y privadas nacionales mantienen el conocimiento experto disperso, limitado y poco organizado, generando, en muchos casos, retrasos o fallidos proyectos que repercuten directamente en la gestión de sus lineamientos estratégicos que se orientan principalmente a buscar nuevos métodos y/o acciones de competitividad o calidad en la prestación de los servicios respectivamente.

2.4 Herramientas de soporte a la gestión del conocimiento

Una particularidad que tienen las herramientas de soporte a la gestión del conocimiento es la codificación, sistematización y transferencia del mismo hacia los profesionales que lo requieren y/o solicitan con el fin de promover y posibilitar una mejor toma de decisiones en el ámbito organizacional. Para alcanzar tales fines, las organizaciones utilizan las tecnologías de información como un elemento fundamental para gestionar y almacenar el conocimiento a través de un espacio colaborativo y abierto de oferta y demanda de información.

En el contexto de la creación y transformación del conocimiento, los expertos Nonaka y Takeuchi (1995) mencionan que existe un vínculo de conexión con las herramientas tecnológicas expuestas en las organizaciones a efectos de lograr los máximos beneficios en las actividades que demandan alto grado de conocimiento. Otro marco de referencia son los estudios generados en el proyecto europeo WISE (Chapter Review of Knowledge Management Tools) en el cual se explicita que las herramientas tecnológicas (plataforma de gestión documental, memoria organizacional, etc.) brindan un soporte adecuado y eficaz a la gestión del conocimiento a través de todo su ciclo de vida.

Tomando como referencia a los autores Nonaka y Carvalho, podemos mencionar como ejemplo algunas herramientas tecnológicas y colaborativas que brindan soporte a la gestión del conocimiento de la siguiente manera:

- Localización
 - Motores de búsqueda de información
 - Páginas amarillas de expertos
- Exploración y explotación
 - Sistemas expertos basados en inteligencia artificial
- Cocreación/descubrimiento
 - Plataformas de soporte a la generación de ideas
 - Herramienta de presentación visual de datos
- Difusión/compartición
 - Plataformas *e-learning*

Complementario al párrafo anterior, los mismos autores realizaron una tipología de las herramientas en el marco de la conversión del conocimiento, de las cuales podemos mencionar:

- De tácito a tácito
 - Páginas amarillas de expertos
- De tácito a explícito
 - Sistemas de agentes inteligentes
- De explícito a tácito
 - Motores de búsqueda
 - Plataformas *e-learning*
- De explícito a explícito
 - *Data mining*
 - Herramientas de mapas conceptuales

Es importante señalar que las ventajas que le ofrece a una organización la implementación de un mapa de conocimiento, en el marco del ciclo de la creación y conversión, está en la facilidad de ubicar el conocimiento y las experiencias del talento humano para tomar decisiones rápidas y oportunas. Zavaleta et al. (2016) señalan que la gestión del conocimiento no solamente debe ser apoyada por las tecnologías de la información, también deberán complementarse con políticas y procedimientos organizacionales que fortalezcan su implementación y difusión en todos los niveles.

2.5 Páginas amarillas de expertos en las organizaciones

Uno de los temas abordados por Drucker (2006) fue el denominado “trabajador del conocimiento” (en inglés, *knowledge worker*), que produce continuamente conocimientos e ideas para mejorar las acciones productivas de una organización, identificando a través de sus experiencias fortalezas y capacidades cognitivas, acciones relevantes en favor de la competitividad organizacional a nivel global.

Para Prada (2005), las redes sociales forjadas a través del conocimiento han venido estableciendo vínculos formales e informales de trabajo y de aprendizaje continuo entre las diferentes personas que desarrollan actividades en común con beneficio a la organización. Sin embargo, es indispensable mencionar que para alcanzar dichos objetivos las relaciones entre las personas expertas y novatas no deben ser únicamente bilaterales, sino integrales de acuerdo al rol que desempeñan dentro de la organización y soportadas por las tecnologías de la información.

Se infiere de los conceptos señalados, que la construcción, el desarrollo y la socialización de conocimientos deben ser soportados por herramientas tecnológicas, políticas y procedimientos de gestión adecuados, con la finalidad de aperturar la acción del

conocimiento hacia los objetivos estratégicos, productivos y competitivos de una organización. Para Angulo Marcial (2007), las páginas amarillas de expertos se convierten en una herramienta tecnológica importante para identificar los datos básicos de un experto, sus redes de contactos, intereses personales y la descripción de su experiencia en la ejecución de los proyectos que ha participado dentro y fuera de la organización; con la finalidad de recurrir a ellos en línea y obtener información acerca de cómo resolver una eventualidad producida en la gestión de proyectos, áreas administrativas o de investigación, destacando su experiencia y competencia sobre determinados temas a consultar.

Como se puede apreciar, la implementación de las páginas amarillas de expertos en las organizaciones son actualmente conocimientos ajenos de los responsables que lideran las diferentes direcciones de la gestión del conocimiento en el Perú, en particular, en el sector público, debido a la falta de información e investigación sobre cómo implementar dichas plataformas tecnológicas para sistematizar el conocimiento de los expertos en un único espacio virtual. Sin embargo, es importante manifestar que las instituciones públicas mantienen interés para comprender adecuadamente los ciclos de vida de la gestión del conocimiento que se circunscriben a su ámbito funcional con la finalidad de implementar diferentes plataformas tecnológicas que coadyuven a mejorar e innovar los servicios que brindan a la sociedad, con base en el conocimiento de expertos.

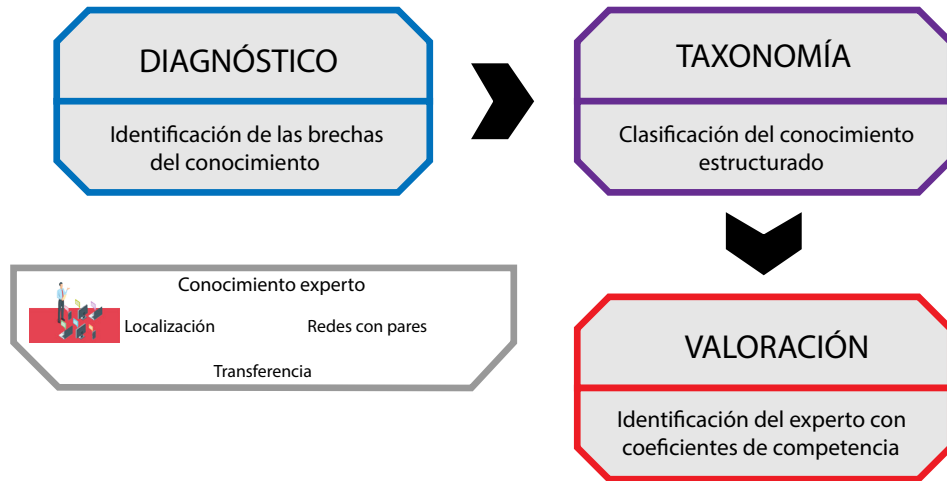
3. METODOLOGÍA DE IDENTIFICACIÓN DE EXPERTOS

La investigación, suscrita en el artículo, propone una metodología para identificar a todos aquellos expertos que realizan actividades en las organizaciones, tomando como precepto los conceptos de la gestión del conocimiento emitidos por los autores Wiig, Choo, Holsapple, Andersen, Nonaka y Takeuchi, entre otros, que en su conjunto proveen información relevante para la identificación de expertos en todos los niveles de una organización.

El artículo expone en su tenor un marco metodológico que contempla tres (3) elementos fundamentales para identificar al experto y que, a través de este proceso, permite sistematizar el conocimiento mapeado por medio de sistemas y plataformas tecnológicas.

Figura 3

Marco metodológico para identificar expertos



Nota. De "Identificación de expertos como fuente de conocimiento en las organizaciones", por A. A. Aures García, y A. E. Balvín Landeo, 2021, *Actas del Congreso Internacional de Ingeniería de Sistemas*, p. 90 (<https://doi.org/10.26439/ciis2021.5579>).

3.1 Diagnóstico

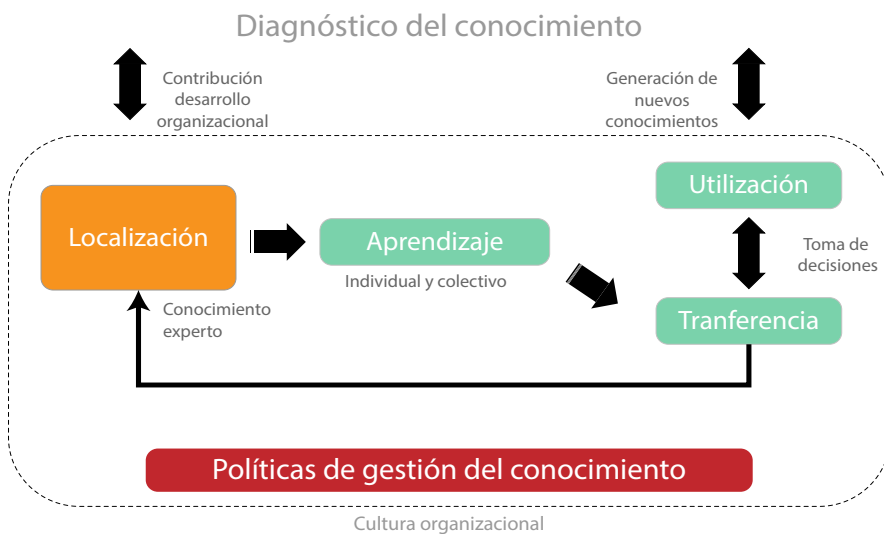
Uno de los primeros puntos a considerar, en la identificación del experto, es determinar los aspectos claves que gobiernan la localización, adquisición y transferencia del conocimiento individual y grupal con el experto. En el caso de la localización, es importante mapear los procesos por medio de los cuales se realizan los flujos de transferencia del conocimiento entre los diversos colaboradores y/o profesionales de una organización.

Otro tema a considerar, en el diagnóstico, es el análisis documental de la política organizacional, la cultura y el uso de las plataformas tecnológicas asociadas con la adquisición y transferencia del conocimiento con el experto, en caso lo hubiera. Asimismo, deberán realizarse encuestas y entrevistas para elaborar el mapa del conocimiento existente en la organización.

En el ámbito del diagnóstico, se deben considerar estrictamente los diferentes procesos (véase figura 4) que involucran el flujo de conocimiento de los expertos con el fin de determinar la transferencia de lo aprendido a nivel individual y/o grupal.

Figura 4

Procesos del diagnóstico situacional del conocimiento



Nota. Adaptado de Bellinza et al. (2011) y Probst et al. (2001)

Es importante mencionar que la investigación se circunscribió estrictamente en el diagnóstico situacional de las unidades y/u oficinas de gestión de proyectos de las instituciones públicas, rectoras en sus respectivos sectores, en relación con las políticas, procedimientos y procesos de aprendizaje que seguían los equipos de proyectos para solucionar problemas de gestión y/o intercambio de conocimientos entre sí. Asimismo, se realizaron un conjunto de entrevistas a los responsables de proyectos con la finalidad de mapear los flujos de transferencia de conocimientos con el experto y las brechas en el mismo.

De acuerdo con las preguntas de la entrevista —¿qué activos de conocimiento se generan o transfieren en cada uno de los procesos del proyecto?, ¿qué expertise se requiere para producir nuevos conocimientos en cada uno de los procesos de la gestión de proyectos? y ¿dónde reside el conocimiento experto?— se elaboraron las matrices de conocimientos de cada institución tomando en consideración lo siguiente:

- Identificación de los procesos, a través de los cuales los activos del conocimiento son generados y transferidos a los equipos de proyectos.
- Conocimiento como insumo o recurso, que hace referencia a los activos de conocimiento que los equipos de proyectos demandan para desarrollar su trabajo.
- El *expertise* requerido se refiere al conocimiento que se necesita para ejecutar las actividades de los procesos de la gestión de proyectos.

- Consulta al experto, representa el conocimiento de valor que los equipos de proyectos necesitan para gestionar sus actividades de manera efectiva.
- Conocimiento como producto generado, que identifica a los activos de conocimiento obtenidos, producidos o transformados de manera gradual en una actividad.

3.2 Taxonomía

De acuerdo con el diagnóstico, se procede a la elaboración de la taxonomía a utilizar en cualquier ámbito organizacional, la cual estará circunscrita a un modelo de red cuyo tenor conceptual se focaliza en la localización e identificación del conocimiento experto. Las taxonomías permiten el despliegue clasificado y categorizado del conocimiento mapeado a efectos de agruparlos en familias para la transferencia y el aprendizaje colectivo entre los diversos actores que se suscriben a una organización (Ezell, 2017).

La figura 5 expresa la vinculación unívoca entre diferentes grupos de trabajo, haciendo uso de redes de intercambio de experiencias e información, para socializar nuevas ideas y acciones estructuradas importantes y estratégicas para cada organización, coadyuvando de esta manera la elaboración de productos y servicios innovadores que ofrecen a la sociedad. Asimismo, se puede apreciar en la figura los flujos del conocimiento experto que se demandan y transfieren continuamente en los grupos de trabajo, con la finalidad de resolver problemas o absolver dudas relacionadas con nuevos productos, tareas, proyectos, etcétera.

Desde el ámbito de la investigación realizada, la taxonomía se diseñó de acuerdo a la relación de los procesos identificados y definidos por las instituciones que fueron objeto de estudio, en las cuales se determina qué tipo de conocimiento y/o activo se demanda con mayor frecuencia entre los equipos de proyectos y a quiénes (expertos) se recurre frecuentemente para la transferencia del conocimiento solicitado.

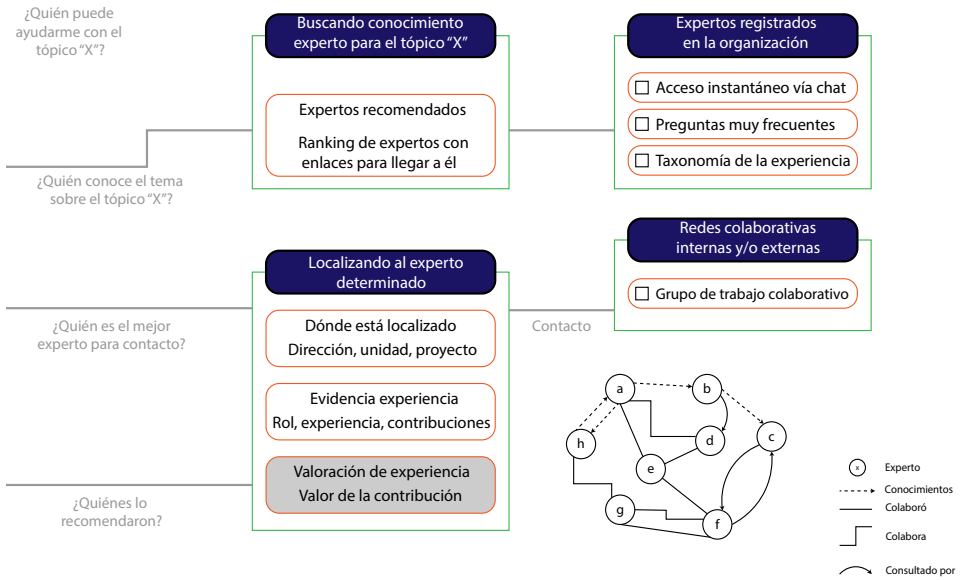
3.3 Valoración

La valoración del conocimiento experto es un instrumento útil que permite a las organizaciones identificar profesionales y/o personas con trayectoria experta en el tema, reconocidas por sus pares, y calificadas para brindar información objetiva sobre una determinada consulta. Para validar lo señalado, primero debe analizarse si los profesionales entrevistados tienen experiencia en el sector académico, público y privado con el fin de determinar los procesos de crecimiento profesional.

En segundo lugar, debe definirse con qué frecuencia se recurre al profesional entrevistado para obtener conocimientos especializados y centrados en un determinado tema, como una fuente de información válida, a efectos de que les permita a las personas y/o grupos solucionar un problema identificado a nivel organizacional.

Figura 5

Taxonomía para la identificación del experto



Nota. De "Identificación de expertos como fuente de conocimiento en las organizaciones", por A. A. Aures García, y A. E. Balvín Landeo, 2021, Actas del Congreso Internacional de Ingeniería de Sistemas, p. 92 (<https://doi.org/10.26439/ciis2021.5579>).

En el marco de la valoración, se debe proceder a evaluar el coeficiente de competencia experta (K), sobre los temas a los cuales se les ha realizado continua consulta y que están mapeados a través de una taxonomía, con la finalidad de identificar las competencias más significativas del profesional en el marco de la gestión del conocimiento. Asimismo, se deben definir los valores de las fuentes de argumentación (coeficiente de argumentación, Ka) en función de las respuestas de valor que emiten sobre determinados hechos o eventos (Cabero Almenara & Barroso Osuna, 2013).

Véase en la tabla 1, algunas fuentes de argumentación a los cuales deberán ser sometidos los expertos para determinar su Ka .

Es importante mencionar que la valoración del conocimiento permite, en una primera instancia, seleccionar y conocer a los expertos que están relacionados con los temas de dominio y, en segundo lugar, comprobar la racionalidad y coherencia de las opiniones emitidas a determinado problema haciendo uso del coeficiente de competencia experta.

Tabla 1

Valoración de las fuentes de argumentación para determinar Ka

Fuente de argumentación	Grado de influencia sobre los temas objeto de estudio		
	Alto	Medio	Bajo
Investigaciones que se realizaron en los temas expuestos en la entrevista	0,30	0,20	0,10
Conocimiento experto generado a través de la actividad profesional	0,50	0,40	0,20
Publicaciones con profesionales locales y/o nacionales	0,05	0,05	0,05
Publicaciones con profesionales extranjeros sobre las temáticas expuestas en la entrevista	0,05	0,05	0,05
Conocimiento de la problemática de manera fiable y objetiva	0,05	0,05	0,05
Uso de herramientas y/o intuición para solucionar el problema	0,05	0,05	0,05

4. RESULTADOS

Para validar la metodología propuesta, se entrevistó a un conjunto de profesionales de mando directivo que se distingue por gestionar proyectos y programas de envergadura nacional, siendo muchos de ellos directores, jefes y/o coordinadores de proyectos. Como se mencionó anteriormente, la investigación se focalizó específicamente en las instituciones públicas rectoras en su sector que tienen adscritas unidades y/u oficinas de gestión de proyectos.

En el caso específico del diagnóstico del conocimiento, se tomó como base en primera instancia la recopilación y el análisis de la documentación proveniente de estas unidades organizacionales, los mismos que describían los procesos, las metodologías y las prácticas usadas en la gestión de los proyectos. En segunda instancia, se procedió a realizar entrevistas *in situ* a los responsables de proyectos con la finalidad de obtener información relevante del estado y la forma de cómo interactuaban con el experto y los flujos de transferencia del conocimiento que se generaban a nivel individual y grupal.

Para este caso, se tomó una muestra de los proyectos más representativos de las instituciones públicas, teniendo en consideración lo siguiente:

- Se consideró únicamente a los proyectos que implicaron mayor cantidad de recursos monetarios empleados en su gestión, mayor cantidad de actores en el proyecto, mayor diversidad de información o conocimientos posibles de encontrar, etcétera.

- Se hizo una clasificación de los proyectos en función al rol que juega dentro de cada institución pública, seleccionándose para esta investigación, únicamente los que corresponden al rol de “gestor y ejecutor”, es decir, en los casos en que cada institución tenía el control absoluto del proyecto, desde el inicio hasta el cierre del mismo.

La información que se relevó por medio del diagnóstico situacional consignó que, en los procesos de la gestión de proyectos, específicamente en la formulación de ideas y de proyectos, el flujo de conocimiento generado se pierde casi en su totalidad debido a dos factores fundamentales:

- a. Las personas que participan en la conceptualización del proyecto son generalmente profesionales de la alta dirección, entre otros, que tienen la potestad y autoridad de aprobar los convenios de financiamiento o planes generales del proyecto, pero que en muchos casos por la rotación de personal y la falta de una estructura documentaria adecuada para registrar y transferir los conocimientos obtenidos durante la aprobación de un proyecto, se pierden considerablemente.
- b. El conocimiento generado por la alta dirección no es transferido adecuadamente a los nuevos miembros responsables de proyectos, por falta de tiempo o presentación de informaciones ambiguas o incompletas, perdiéndose activos del conocimiento importantes y relevantes para la planificación del mismo.

Debido a este quiebre de transferencia del conocimiento, los equipos de proyectos requirieron la opinión de expertos internos, en algunos casos identificados, para completar y definir las informaciones correspondientes a las demás fases del proyecto. En muchos casos, los responsables de proyectos (de un total de 52) de cierta manera han tenido que recurrir a empresas consultoras y/o asesorías especializadas externas para suplir una falta de *expertise* en algunas de las actividades del proyecto, ya que en su mayoría (80 % de ellos) no sabían localizar ni identificar adecuadamente el conocimiento experto institucional.

Al no identificar objetivamente al “experto institucional”, los responsables de los equipos de proyectos presentaban bajas satisfacciones en la transferencia del conocimiento que recibían (73 % de un total de 52). Esto era considerado como pérdida de tiempo y esfuerzo en el intercambio de opiniones con profesionales que desconocían la problemática expuesta y que al parecer se les denominaban “expertos”.

En el marco de la taxonomía, se mapearon y graficaron los procesos y los activos de conocimiento relacionados en el marco de la gestión del proyecto. Esto permitió visualizar los distintos tipos de conocimientos que se demandaban con mayor frecuencia entre los equipos de proyectos y a quienes se recurrían para obtener información y *expertise* sobre determinados temas.

Es importante señalar que los equipos entrevistados manifestaron frecuentemente la necesidad de localizar expertos temáticos en doce puntos críticos, que fueron registrados en una taxonomía de proyectos y entre los cuales podemos mencionar: análisis de impacto social, análisis de factibilidad, técnicas de negociación, comunicación en conflictos, identificación de interesados, matriz de riesgos, estimación de costos, métodos de comunicación con habilidades blandas, resolución de conflictos, métricas de calidad, registro y sistematización de lecciones aprendidas.

Al implementarse la identificación de expertos, en función a su nivel de coeficiente de competencia experta ($0,8 < K < 1,0$) y de acuerdo a la temática de dominio, los equipos de proyectos valoraron el aporte y el involucramiento del experto con relación a la solución de un problema determinado. El 81 % de los responsables de proyectos consideraron que el conocimiento del experto identificado le generó una solución efectiva y oportuna a un incidente propio en la gestión de sus proyectos, razón por la cual se han generado lazos de confianza institucionales para realizar trabajos de manera conjunta.

Como se puede apreciar, la no identificación y localización de un experto dentro del ámbito institucional conlleva, en muchos casos, a buscar consultores y/o profesionales cualificados, recomendados y ajenos a la institución, para gestionar proyectos que requieren *expertise* en determinados procesos del conocimiento. En muchos casos, este proceso puede coadyuvar a aumentar la insatisfacción de la transferencia del conocimiento experto (94 % de un total de 52) hacia los equipos de proyectos por el poco dominio y la capacidad de resolver un problema en un determinado tiempo.

Tomando como referencia lo señalado, la integración de una taxonomía de los activos del conocimiento de proyectos a una plataforma tecnológica puede ser de gran utilidad para identificar las clases de la localización del conocimiento experto dentro de cualquier organización, así como su registro una vez finalizada la consulta.

5. CONCLUSIONES

La validación de la metodología propuesta para identificar al experto se ha circunscrito estrictamente en las instituciones públicas nacionales que gestionan proyectos y/o programas de envergaduras notables en su sector, conllevando a realizar entrevistas y cuestionarios a un número mayor de cincuenta profesionales responsables de proyectos. La investigación ha denotado que las organizaciones no tienen implementado y/o formalizado políticas de la gestión del conocimiento y de herramientas de identificación y localización de expertos internamente, lo cual fue manifestado en las entrevistas que se realizaron a los equipos de proyectos de cada institución pública. Sin embargo, al implementarse la metodología de identificación de expertos, los integrantes de cada equipo de proyectos comenzaron a generar redes de contactos para compartir conocimientos con los expertos a fin de obtener respuestas rápidas y objetivas en su gestión.

En el marco de la taxonomía, los equipos entrevistados manifestaron frecuentemente la necesidad de identificar expertos temáticos en los diferentes puntos críticos de la gestión de proyectos tales como: análisis de impacto social, análisis de factibilidad, técnicas de negociación, comunicación en conflictos, entre otros; a efectos de renovar el conocimiento organizacional para solucionar problemas de una manera efectiva. La identificación de expertos en las instituciones públicas ha influido abiertamente en los procesos de la transferencia del conocimiento y la generación de ambientes colaborativos de manera paulatina, en distintas temáticas, orientadas a fortalecer el aprendizaje individual y colectivo para la resolución de problemas específicos en la gestión de proyectos. Con referencia a los tiempos de respuesta del experto a un determinado problema y su involucramiento en las redes colaborativas de las instituciones que fueron objeto de estudio, en una primera instancia, eran lentas porque se le hacían consultas ajenas a su dominio, generando pérdidas de tiempo y en algunos casos insatisfacción en las respuestas obtenidas. Sin embargo, al implementarse la metodología de identificación de expertos, los tiempos de respuesta mejoraron significativamente ya que muchos de los expertos fueron identificados y adscritos a una red de trabajo, que les permitiera ser consultados con objetividad de acuerdo a su *expertise*.

Otros de los aspectos que destacar es que las instituciones públicas objetos de este estudio, no tienen implementadas políticas de gestión del conocimiento por desconocimiento sobre esta temática; sin embargo, muchos de sus funcionarios manifiestan que el fortalecimiento de competencias basado en el conocimiento es una preocupación latente que está afectando la objetividad y el desempeño de los colaboradores en distintos ámbitos de cada institución. Finalmente, debemos expresar que la investigación brinda las bases necesarias para elaborar una estructura y/o arquitectura tecnológica y conceptual que permite la construcción de sistemas para la identificación sistematizada de expertos a partir de la consideración de los factores de dominio y de coeficiente de competencia experta incluidos en la metodología aquí propuesta.

REFERENCIAS

- Alavi, M., & Leidner, D. E. (2001). Review: Knowledge management and knowledge management systems: Conceptual foundations and research issues. *MIS Quarterly*, 25(1), 107-136. <http://dx.doi.org/10.2307/3250961>
- Angulo Marcial, N. (2007). Ubicando el conocimiento experto: Las páginas amarillas. *Innovación Educativa*, 7(40), 49-61. <https://www.redalyc.org/pdf/1794/179420967004.pdf>
- Andersen, A. (1999). *El management en el siglo XXI*. Ediciones Granica.

- Arambarri Basañez, J. (2014). *Metodología de evaluación y gestión del conocimiento dinámico por procesos utilizando como soporte TIC el entorno colaborativo de trabajo basado en el modelo de creación de conocimiento de Nonaka-Takeuchi. Caso de estudio en el área de Gestión de proyectos de I+D+i en institución avanzada en conocimiento* [Tesis doctoral, Universidad de Córdoba]. Biblioteca Universidad de Córdoba. <http://hdl.handle.net/10396/12114>
- Archibold, W., & Escobar, A. (2015). Capital intelectual y gestión del conocimiento en las contralorías territoriales del departamento del Atlántico. *Revista Dimensión Empresarial*, 13(1), 133-146. <http://www.scielo.org.co/pdf/diem/v13n1/v13n1a08.pdf>
- Aures García, A. A., & Balvín Landeo, A. E. (2021). Identificación de expertos como fuente de conocimiento en las organizaciones. *Actas del Congreso Internacional de Ingeniería de Sistemas*, 83-96. <https://doi.org/10.26439/ciis2021.5579>
- Bellinza, M., Guerrero-Barrera, N., Colon-Sirtori, S., & Ramírez-Mercado, W. (2011). Gestión del conocimiento: Aproximaciones teóricas. *Clío América*, 5(10), 257-271. <https://revistas.unimagdalena.edu.co/index.php/clioamerica/article/view/422>
- Cabero Almenara, J., & Llorente, M. (2013). La aplicación del juicio de experto como técnica de evaluación de las tecnologías de la información (TIC). *Eduweb: Revista de Tecnología de Información y Comunicación en Educación*, 7(2), 11-22. <http://servicio.bc.uc.edu.ve/educacion/eduweb/v7n2/art01.pdf>
- Cabero Almenara, J., & Barroso Osuna, J. (2013). La utilización del juicio de experto para la evaluación de TIC: el coeficiente de competencia experta. *Bordón. Revista de Pedagogía*, 65(2), 25-38. <https://doi.org/10.13042/brp.2013.65202>
- Carvalho, R., & Ferreira, M. (2001). Using information technology to support knowledge conversion process. *Information Research*, 7(1). <http://informationr.net/ir/7-1/paper118.html>
- Choo, W. (2016). *The inquiring organization: How organizations acquire knowledge and seek information*. Oxford University Press.
- Drucker, P. (2006). *Innovation and entrepreneurship*. Harper Business.
- Ezell, M. P. (2017). *How organizations use collaboration to create knowledge*. Dissertations available from ProQuest. AAI10255888.
- Flores López, J., & Ochoa Jiménez, S. (2016). Los modelos de gestión del conocimiento y su relación con la cultura organizacional: una revisión teórica. *Ciencia Administrativa*, (2), 179-189. <http://biblioteca.udgvirtual.udg.mx/jspui/handle/123456789/3575>
- Holsapple, C. W. (Ed.). (2004). *Handbook of knowledge management 1*. Springer. <https://doi.org/10.1007/978-3-540-24746-3>

- Marin-Garcia, J., & Zarate-Martinez, E. (2008). Propuesta de un modelo integrador entre la gestión del conocimiento y el trabajo en equipo. *Intangible Capital*, 4(4), 255-280. <https://www.intangiblecapital.org/index.php/ic/article/view/65/87>
- Mohajan, H. (2017). The impact of knowledge management models for the development of organizations. *Journal of Environmental Treatment Techniques*, 5(1), 12-33. MPRA Paper n.º 83089. <https://mpra.ub.uni-muenchen.de/id/eprint/83089>
- Nonaka, I., & Takeuchi, H. (2009). *The knowledge-creating company. How Japanese companies create the dynamics of Innovation*. Oxford University Press.
- Prada, E. (2005). Las redes de conocimiento y las organizaciones. *Revista Bibliotecas y Tecnologías de la Información*, 2(4).
- Pérez Arrau, G. (2016). Gestión del conocimiento en el sector público chileno. En F. Batista Ferreira (organizador), *Experiências Internacionais de Implementação da Gestão do Conhecimento no Setor Público*. Instituto de Pesquisa Econômica Aplicada (Ipea).
- Probst, G., Raub, S., & Romhardt, K. (2001). *Administre el conocimiento*. Prentice Hall.
- Robles, P., & Rojas, M. (2015). La validación por juicio de expertos: dos investigaciones cualitativas en lingüística aplicada. *Revista Nebrija de Lingüística Aplicada a la Enseñanza de Lenguas*, (18).
- Sveiby, K.-E., Gripenberg, P., & Segercrantz, B. (Eds.). (2012). *Routledge studies in technology, work and organizations: Challenging the innovation paradigm*. Routledge.
- Wiig, K. (2012). *People-focused knowledge management. How effective decision-making leads to corporate success*. Routledge.
- Zavaleta, M., Brito, L., & Garzón, M. (2016). Modelo de gestión del conocimiento en el área de TIC para una universidad del Caribe colombiano. *Revista Lasallista de Investigación*, 13(2), 136-150. <http://hdl.handle.net/10567/2859>

Recibido: 25/5/2022 / Aceptado: 8/6/2022

doi: <https://doi.org/10.26439/interfases2022.n015.5886>

A DISTRIBUTED MODEL FOR COMPUTING 3D MESH LOCAL DESCRIPTORS BASED ON *K*-RINGS

FRANCI SUNI-LOPEZ

fsunilo@unsa.edu.pe / ORCID: 0000-0002-4212-7910
Universidad Nacional de San Agustín de Arequipa, Peru

JAN HURTADO

jhurtadoj@unsa.edu.pe / ORCID: 0000-0003-3422-3117
Universidad Nacional de San Agustín de Arequipa, Peru

ALEJANDRA MÁRQUEZ

amarquezhe@unsa.edu.pe
Universidad Nacional de San Agustín de Arequipa, Peru

LEONARDO GUZMÁN

lguzmanz@unsa.edu.pe
Universidad Nacional de San Agustín de Arequipa, Peru

In order to facilitate 3D object processing, it is common to use high-level representations such as local descriptors that are usually computed using defined neighborhoods. *K*-rings, a technique to define them, is widely used by several methods. In this work, we propose a model for the distributed computation of local descriptors over 3D triangular meshes, using the concept of *k*-rings. In our experiments, we measure the performance of our model on huge meshes, evaluating the speedup, the scalability, and the descriptor computation time. We show the optimal configuration of our model for the cluster we implemented and the linear growth of computation time regarding the mesh size and the number of rings. We used the Harris response, which describes the saliency of the object, for our tests.

KEYWORDS: 3D local descriptor / geometry processing / distributed computing / large meshes

Un modelo distribuido para calcular descriptores locales de malla 3D basados en *k*-rings

Para facilitar el procesamiento de objetos 3D, es común utilizar representaciones de alto nivel, como los descriptores locales que generalmente se calculan utilizando vecindarios definidos. *K-rings* es una técnica para definirlos y es ampliamente utilizada por varios métodos. En este trabajo, proponemos un modelo para el cálculo distribuido de descriptores locales sobre mallas triangulares 3D, utilizando el concepto de anillos *k*. En nuestros experimentos, medimos el rendimiento de nuestro modelo en mallas enormes, evaluando la aceleración, la escalabilidad y el tiempo de cálculo del descriptor. Mostramos la configuración óptima de nuestro modelo para el clúster que implementamos y el crecimiento lineal del tiempo de cálculo con respecto al tamaño de la malla y el número de anillos. Usamos la respuesta de Harris, que describe la prominencia del objeto, para nuestras pruebas.

PALABRAS CLAVE: descriptor local 3D / procesamiento geométrico / computación distribuida / mallas grandes

1. INTRODUCTION

3D repositories are growing rapidly (Gao et al., 2015) and, consequently, there are many techniques for 3D data acquisition (e.g. 3D modeling, 3D scanners, 3D reconstruction, depth-sensing cameras, etc.). In this context, processing this type of data is an important task for computer scientists. For instance, areas such as videogames, medicine, archeology, biology, engineering, physics, and others use this kind of data for real object representation. There are some tasks that are very important for 3D data analysis: registering (Pavlakos et al., 2018; Maquart et al., 2021; O' Sullivan et al., 2022; Van Kaick et al., 2011; Li et al., 2015), symmetry detection (Mitra et al., 2013; Areias & Rabczuk, 2017), segmentation (Figueiredo et al., 2021; Zhou et al., 2017), sampling and compression. These tasks depend on the representation of the object; a 3D object can be represented by point clouds, polygon meshes, polygon soups, connectivity graphs, volumetric pixels (voxels), etcetera (low-level representations). To facilitate processing, it is common to use high-level representations such as descriptors, graphs, skeletons, key points, components, etc.

In this paper, we focus on local descriptor representation. A local descriptor is usually defined on each atomic structure of low-level representations; we use polygon mesh representations and their vertices as atomic structures (i.e., we work on triangular meshes). Usually, when computing local descriptors, it is necessary to establish the neighborhood of each vertex of the mesh. This neighborhood allows us to define local features. For example, if we want to define the curvature of a single vertex, it is not necessary to know the distribution of all vertices. Gelfand et al. (2005) proposed the Integral Volume Descriptor, which evaluates the curvature of a region of the surface immersed in a sphere (determines the neighborhood) of a given radius. Lee et al. (2005) proposed a descriptor based on the mean curvature. Gaussian filters are used in different scales to obtain the saliency value of each vertex. They determine neighborhoods using geodesic or Euclidean distance on the mesh. Then, Castellani et al. (2008) applied the Gaussian filter directly to the points of the object instead of the values.

Another way to determine a neighborhood is using the concept of rings. Zaharescu et al. (2009) proposed a descriptor based on the discrete curvature; they defined neighborhoods to obtain the approximate curvature, these neighborhoods were composed by the set of the k-level adjacent vertices of each vertex. Sipiran and Bustos (2011) implemented a technique called Harris 3D, which is based on the corner detection technique proposed by Harris and Stephens (Harris & Stephens, 1988).

They compute the Harris response for each point regarding a locality conformed by different rings; also, they approximate a continuous surface for this locality and transform it using Principal Component Analysis (PCA) to address rotation invariance. The saliency is computed evaluating the derivatives on the surface.

It is possible to generate high definition meshes with millions of vertices. For example, the digital Michelangelo project (Levoy et al., 2000), has meshes with more than 900 million faces. Using a sequential model to process large meshes involves considerable computation time. Multi-thread and GPU based algorithms are very useful when we want to speed up this task. The other problem we want to solve when working with large meshes is the insufficient memory space. Sometimes this problem is addressed using compact data structures for space optimization (Chen et al., 2009; Zamolo et al., 2022; Herath et al., 2020; Cignoni et al., 2003; Aleardi et al., 2005; Gurung et al., 2011; Luffel et al., 2014; Gurung et al., 2013; Rocca et al., 2011), but it decreases processing speed.

Distributed data processing is a computer-networking method in which multiple computers across different locations share computer-processing capability with all connected systems. That is the reason why distributed models have been widely used in different areas of computer science, taking advantage of the reduction of processing time when working on a distributed manner. The distributed paradigm has been addressed in many works concerning image processing (Le Tien et al., 2021; Verma et al., 2018; Sqyres et al., 2000; Prajapati & Vij, 2011; Warn et al., 2009), video processing (Pereira et al., 2010), molecular data analysis (Gurung et al., 2011), etc. In the case of 3D data processing, there are multiple frameworks for specific tasks. Balman (2006) proposed a distributed method for non-uniform triangular mesh refinement. Vo et al. (2011) used a MapReduce model for 3D data visualization. Cabiddu and Attene (2015) proposed a web-based system for the generally distributed processing of triangular meshes.

Determining the locality or neighborhood is commonly associated with the k nearest neighbors (k NN) problem. Sankaranarayanan et al. (2007) proposed an extension of their previous work for distributed processing; they used large point clouds for their experiments and calculated the nearest neighbors using Euclidean distance. We can use this to estimate a neighborhood, but the problem of a Euclidean range search is that we can get wrong results if two different parts of the mesh are very close regarding this metric and very far following the topological structure. That is why it is important to define a geodesic metric. The easiest way to discretize it is using adjacency level distance. This will be explained later (k -rings).

In this work, we propose a distributed model for computing 3D triangular mesh local descriptors based on k -rings. We consider a multi-process architecture to address parallel processing and a distributed disk storage for insufficient memory capacity problems. We implemented this model using MPI and applied the Harris response descriptor on it (Sipiran & Bustos, 2011). We conducted several experiments to obtain an optimal configuration for our cluster. Also, we evaluated the speedup, the scalability, and the descriptor computation time.

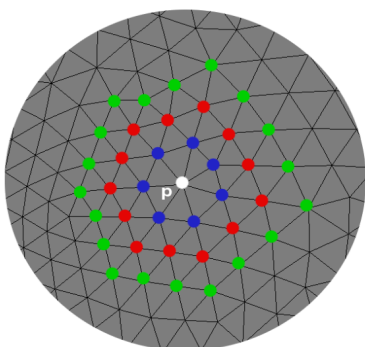
The paper is organized as follows. In section 2, we explain how to represent a 3D object, how to define a descriptor, how to define a neighborhood, and how the Harris response works. The proposed model is explained in section 3; each step is presented in detail here. We explain our experiments and show our results in section 4. Finally, in section 5, we present our conclusions and future work.

2. BACKGROUND

A 3D object can be formally represented as a 2-manifold. Discretizing, we can represent it as a triangular mesh. A triangular mesh is composed by a set of vertices and a set of faces (triangles). The faces define the connectivity of the vertices, this representation preserves geometric and topological features. A point-wise descriptor is a function, that describes all the object; these kinds of descriptors work independently on each vertex and its locality. First, we must compute the neighborhood and then the descriptor values. Each neighborhood can be defined using the concept of k -rings (an example of this concept is shown in figure 1, used by Sipiran and Bustos (2011)).

Figure 1

k -rings



Note. White: evaluated vertex. Blue: 1-ring. Red: 2-ring. Green: 3-ring. Adapted from *Harris 3D: a robust extension of the Harris operator for interest point detection on 3D meshes* by I. Sipiran & B. Bustos, 2011, *The Visual Computer*, 27 (<https://link.springer.com/content/pdf/10.1007/s00371-011-0610-y.pdf>)

A k -ring is defined as follows:

$$ring_k(p) = \{p' \in M \mid \text{minimum_path}(p, p') = k\} \quad (1)$$

We assume that the length of the edges is one. A neighborhood with radius equal to k is the set formed by the union of all rings with radius $\leq k$. Formally we have:

$$neighborhood_k(p) = \bigcup_{i=1}^k ring_i(p) \quad (2)$$

2.1 Harris 3D

In our experiments, we use the Harris response to get our results. Harris response is a local point-wise descriptor that uses neighborhoods to get the saliency value of each vertex. Given a neighborhood N of a vertex, defined using the concept of k -rings, it is possible to approximate a surface. To address rotation invariance, a fitting plane is computed using Principal Component Analysis (PCA). The neighborhood is rotated with respect to the normal of this plane. Sipiran & Bustos (2011) fit a quadratic surface to the set of transformed vertices. This surface has the following form:

$$f(x, y) = \frac{p_1}{2}x^2 + p_2xy + \frac{p_3}{2}y^2 + p_4x + p_5y + p_6 \quad (3)$$

Then, with this continuous surface, the derivatives can be calculated on the vertex by:

$$f_x = \frac{\delta f(x,y)}{\delta x} \quad (4)$$

$$f_y = \frac{\delta f(x,y)}{\delta y} \quad (5)$$

To avoid noise problems, these authors propose continuous Gaussian functions that can be expressed as follows:

$$A = p_4^2 + p_1^2 + p_2^2 \quad (6)$$

$$B = p_5^2 + p_2^2 + p_3^2 \quad (7)$$

$$C = p_4p_5 + 2p_1p_2 + 2p_2p_3 \quad (8)$$

With these terms, the matrix E is defined:

$$E = \begin{pmatrix} A & C \\ C & B \end{pmatrix} \quad (9)$$

The Harris response is computed by:

$$h = \det(E) - q(\text{tr}(E))^2 \quad (10)$$

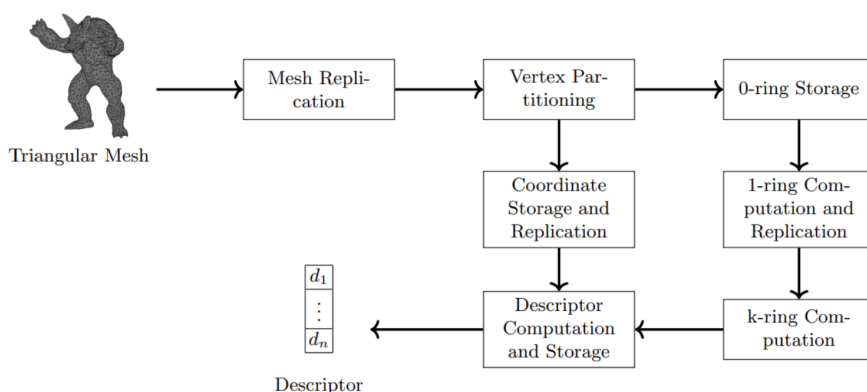
Where q is a constant.

3. DISTRIBUTED MODEL

In this paper, we introduce a distributed model for 3D mesh local descriptor computation. The cluster executes in parallel the same tasks with different data and different parameters. First, we define a neighborhood based on k -rings; using this technique, it is necessary to find different levels of adjacency (0- k levels). The adjacent vertices, which make up the neighborhood, are necessary to compute the descriptor.

Figure 2

Workflow of the proposed algorithm



Note. Adapted from *Harris 3D: a robust extension of the Harris operator for interest point detection on 3D meshes*, by I. Sipiran & B. Bustos, 2011, *The Visual Computer*, 27, (<https://link.springer.com/content/pdf/10.1007/s00371-011-0610-y.pdf>)

We address two problems: insufficient memory capacity and parallel processing. To solve the first one, our method uses disk I/O operations for final and temporal data, which is helpful for partial results storage. These operations are executed in parallel, on each node of the cluster. The second one is solved using a multi-process architecture. Our model consists of seven steps. The input is a 3D triangular mesh and the output, its respective descriptor. 1) Mesh replication: distributes the mesh file. 2) Vertex partitioning: regular partitioning using appearance order. 3) Coordinate storage and replication: extracts coordinates from mesh file and replicates them on each cluster node. 4) 0-ring storage: the 0-ring is integrated by the evaluated vertex. 5) 1-ring computation and replication: the 1-ring is integrated by the adjacent vertices, and it is replicated on all the cluster. 6) K -ring computation: the k -ring is computed using the $(k-1)$ -ring, the $(k-2)$ -ring and the 1-ring. 7) Descriptor computation and storage: using the computed rings we can load the coordinates for descriptor computation. The workflow of our model is shown in figure 2.

The mesh replication step replicates the mesh input file on each node of the cluster (a node can have multiple processes). Each process reads the same file, which has a set of vertices (with their respective coordinates) and a set of faces (triangles). The vertices are not sorted or clustered, so we can't split them by proximity. In the next step, the set of vertices is partitioned according to their order of appearance. These partitions are assigned to each process. Then, we must store two types of data: global data and local data. The global data is replicated on each node of the cluster, and the local data is stored in the corresponding node.

The 0-ring and the coordinates can be loaded at the same time. In a typical mesh file, the coordinates of each vertex are located at the top. These coordinates are replicated (global data) and will be useful for descriptor computation. The 0-ring is integrated by the vertices that have 0 distance from the source, therefore the 0-ring is conformed only by the evaluated vertex. In the mesh file each triangle is represented by the indices of the three vertices that conform it. We use that information to compute the 1-ring of each vertex and then replicate this information all over the cluster (global data). This ring is replicated because we have to access it several times for the computation of the following rings.

In our implementation, we compute and store the coordinates, the 0-ring and the 1-ring before doing the replication process. Each node of the cluster generates files for their corresponding partitions. We execute one process for each partition and each process is composed by sequential steps depending on the memory usage we want. Support data structures are used to optimize the computation time. Each node executes its processes in parallel. When this task is finished, the coordinates and the 1-ring are replicated all over the cluster.

The next step is the computation of the k -rings. We make use of the 1-ring to obtain the adjacent vertices of a specific vertex. If we are computing the k -ring, we need to access to the $(k-1)$ -ring and $(k-2)$ -ring. We use the first one in order to get the adjacent vertices of the previous ring (external ring). And the second one because the adjacent vertices we select should not be in the rings that have already been computed. We use data structures to speed up this step. As in the previous steps, we assign a process for each partition and use sequential steps depending on the memory usage. We compute the k -rings sequentially regarding the partitions. Figure 3 shows in detail the algorithm used in this step .

Figure 3

Algorithm of computation of k -ring

```
1: procedure K-RING( $k, partitions, steps$ )
2:   for all  $p \in partitions$  do
3:     split  $p$  in a set of chunks  $C$ , regarding the number of  $steps$ 
4:     for all chunk  $c \in C$  do
5:       for all vertex  $v \in c$  do
6:         load  $(k-1)$ -ring and  $(k-2)$ -ring, regarding  $v$ 
7:         for all vertex  $v \in c$  do
8:           for all vertex  $a$  such that  $(v, a)$  is an edge do
9:             if  $a \notin (k-1)$ -ring and  $a \notin (k-2)$ -ring then
10:              include  $a$  in  $k$ -ring set, for the evaluated vertex  $v$ 
11:   write  $k$ -ring partition
```

Note. Adapted from *Harris 3D: a robust extension of the Harris operator for interest point detection on 3D meshes* by I. Sipiran & B. Bustos, 2011, *The Visual Computer*, 27, (<https://link.springer.com/content/pdf/10.1007/s00371-011-0610-y.pdf>)

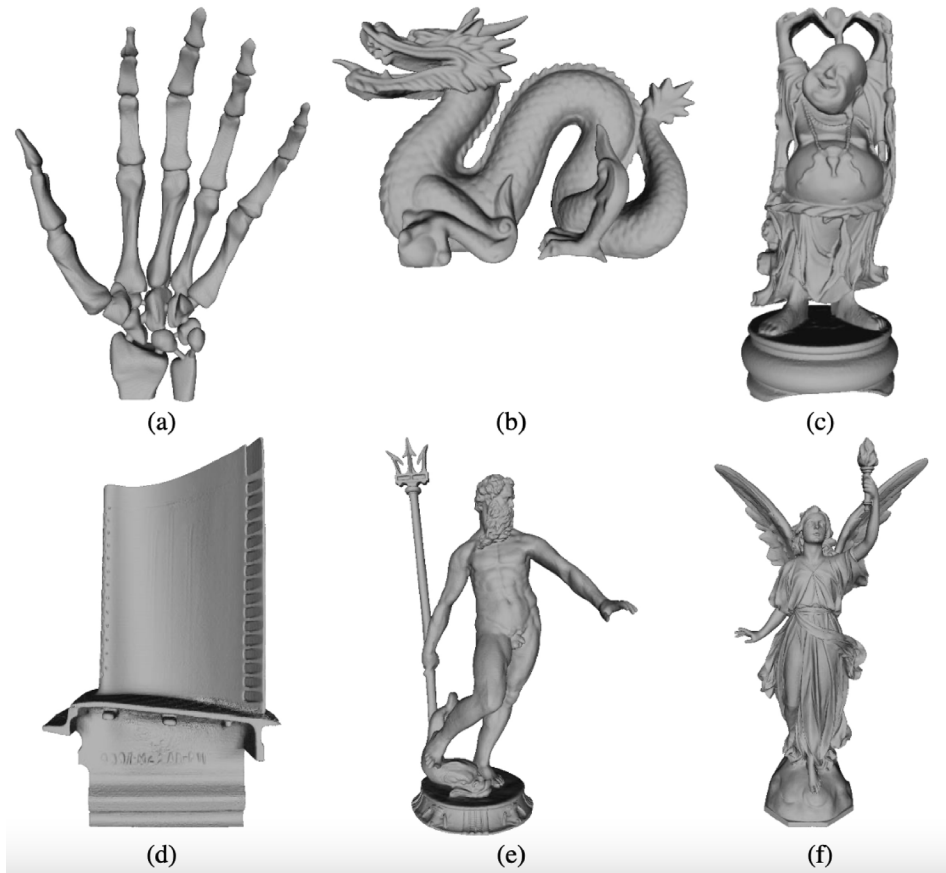
Given the set of rings and the coordinates, we can define a neighborhood and the descriptor values. In the same way as in a single ring computation of the k -rings, we load the coordinates instead of the adjacent vertices. For each partition, the corresponding descriptor is computed regarding the vertices involved. The result consists of files storing the descriptor values of each vertex of the mesh.

4. EXPERIMENTS AND RESULTS

We implemented a cluster of ten computers, each with the following specifications: Ubuntu 14.04 LTS OS, Intel Core i7-4770 CPU @ 3.40GHz x8 processor architecture and 8 Gb of RAM. We did several tests to obtain an optimal configuration of the number of processes and steps. For testing reasons, we generated uniform plane meshes of different sizes. Also, we used 3D models which were directly obtained from the Stanford 3D Scanning Repository¹, the AIM@SHAPE Shape Repository² and the GIT Large Geometry Models Archive³. These models are shown in figure 4.

Figure 4

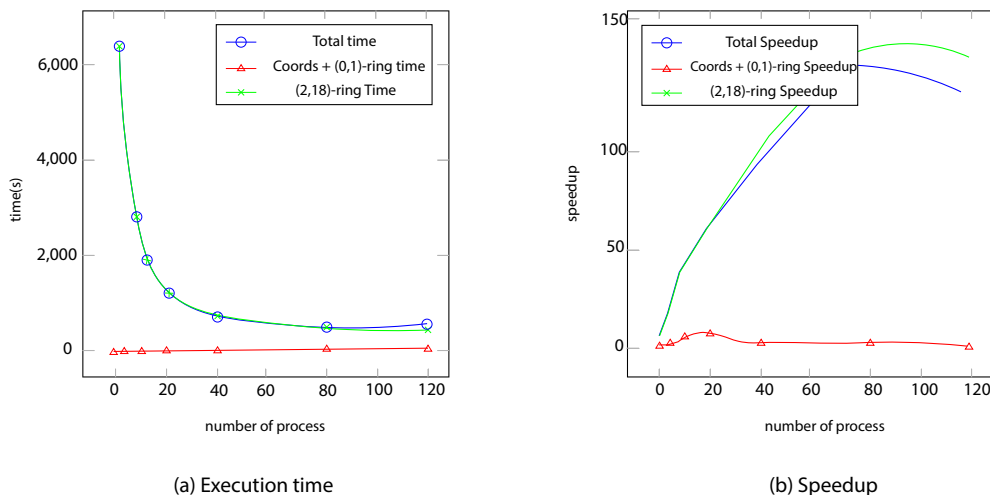
Models used for the experiments



Note. The models were obtained from the Stanford 3D Scanning Repository, the AIM@SHAPE Shape Repository, and the GIT Large Geometry Models Archive. (a) Skeleton Hand. (b) Dragon. (c) Happy Buddha. (d) Turbine Blade. (e) Neptune. (f) Lucy.

Figure 5

Execution time and speed up obtained over a mesh with 16 million vertices using different number of processes



Note. Blue: total time/speedup of k-rings computation. Red: time/speedup of coordinates and (0,1)-ring computation. Green: time/speedup of (2-k)-ring computation.

4.1 Speedup

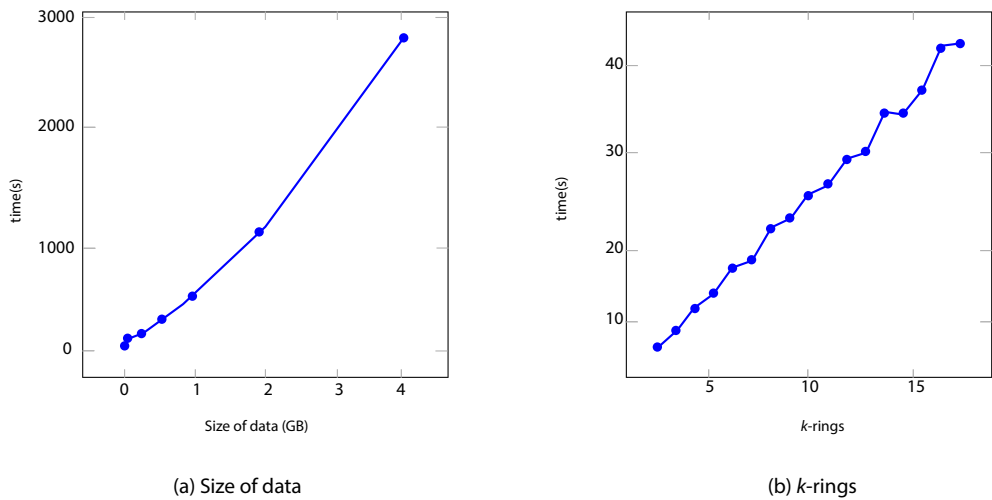
First, using a mesh conformed by 16 million vertices we calculated the execution time of the coordinates and (0-k)-ring computation for different number of processes. This task is the critical part in our implementation. In the case of the coordinates and (0,1)-ring computation, we obtained better results using one step for all number of processes except when we used only one. This occurred since we didn't have enough memory capacity. For the (2-k)-ring computation we calculated an optimal number of steps depending on the number of processes used. Here we did not consider the computation time of the descriptor because the process of extracting the neighborhoods is very similar to a single ring computation and calculating the values for each neighborhood will depend on the descriptor we choose. In table 1 we show the total and partial execution times for different number of processes. Also, we show the speedup obtained regarding the sequential implementation. In figure 5 (a) and 5 (b) we show the graphs obtained from the execution time values and the speedup values respectively. The optimal number of processes for the total computation time is 80. But in the case of coordinates and (0,1)-ring computation, the optimal number of processes is 20. This occurs because these steps don't have too much processing cost.

4.2 Scalability

Using the optimal number of processes (80), we tested the task mentioned on the previous section on plane regular meshes of different sizes. The number of rings that we calculated was 18. The number of steps we used were selected according to the size of the mesh. In table 2 we show the number of vertices and the size of the meshes, the number of steps of each process, the size of the computed data (rings and coordinates), and the total computation time. In figure 6 (a) we show the respective behavior. The graph shows that the growth is linear with respect to the number of vertices.

Figure 6

K-rings computation time using plane meshes of different sizes



Note. (b) shows the computation time of k -rings with 16 million of vertices and 80 processes.

We calculated the computation time for each ring over a regular plane mesh of 16 million vertices, using 80 processes. In table 3 we show the results from 2-Ring to 18-ring, and in figure 6(b) its respective graph. The growth regarding the number of rings is also linear.

Table 1

Results describing the execution time and speedup obtained using different number of processes over a mesh with 16 million vertices

Proc.	Steps	Coords + (0,1)- ring (s)	(2-18)-ring (s)	Total (s)	Speedup
1	25	130,083	50 849,972	50 980,055	1
2	18	43,023	7203,385	7246,408	7,035
5	16	44,918	3093,867	3138,785	16,242
10	15	29,118	1448,57	1477,688	34,5
20	14	17,057	931,066	948,123	53,769
40	14	31,782	588,037	619,819	82,25
80	13	34,834	371,161	405,995	125,568
120	11	63,030	385,741	448,771	113,559

Table 2

K-ring computation time using plane meshes of different sizes

Vertices	Size (GB)	Steps	Data (GB)	Time (s)
1 million	0,052	1	0,666	28,404
2 million	0,109	2	1,5	52,492
4 million	0,224	4	3,1	90,751
8 million	0,454	7	4,8	151,881
16 million	0,948	13	14	405,995
32 million	2	20	28	1090,801
64 million	4	30	56	2744,154

Table 3

Computation time in seconds of (1-18) – ring using 80 processes in a 16 million vertex plane mesh

Ring	Time (s)	Ring	Time (s)	Ring	Time (s)
2-ring	4,856	8-ring	19,367	14-ring	33,587
3-ring	6,927	9-ring	20,677	15-ring	33,567
4-ring	9,649	10-ring	23,472	16-ring	36,394
5-ring	11,523	11-ring	24,840	17-ring	41,592
6-ring	14,564	12-ring	27,583	18-ring	42,052
7-ring	15,557	13-ring	29,920		

4.3 Descriptor

In this test we calculated the execution times for all the tasks involved in the descriptor computation. We used the Harris response as a descriptor, which is in Sipiran & Bustos, (2011).

Table 4*K-ring based descriptor computation*

Mesh	# Vertices	Size (GB)	Replication (s)	(2-10) – ring (s)	Descriptor (s)	Data size (GB)	Total (s)
Skeleton Hand	327 323	0,020	1,084	3,800	1,101	0,091	9,7
Dragon	437 645	0,029	1,504	4,109	1,939	0,153	14,449
Happy Buddha	543 652	0,037	2,082	4,499	2,469	0,199	17,689
Turbine Blade	882 954	0,037	2,267	5,771	2,806	0,254	19,74
Neptune	2 003 932	0,150	6,736	10,445	8,774	0,626	55,552
Lucy	1 402 7872	0,969	36,641	65,878	45,061	4,9	285,682
Plane (16)	16 000 000	0,948	34,834	72,603	34,5	4,8	238,607
Plane (32)	32 001 649	1,957	68,781	156,604	140,193	10,5	582,54
Plane (64)	64 000 000	3,977	135,969	324,011	419,891	21,5	1533,345

The neighborhoods are conformed by rings with $k \leq 10$. In the test we used six meshes obtained from the repositories we previously mentioned, and three plane meshes with different sizes. The results we obtained are shown in table 4. We show the name of the mesh, the number of vertices, the size of the mesh, the coordinates and (0,1)-ring execution time, the replication time, the (2-10)-ring execution time, the descriptor computation time, the size of the generated data, and the total time of the whole computation. Our implementation works better when the partition vertices are near, since less disk operations are needed to obtain the adjacent vertices and the coordinates.

5. CONCLUSION AND FUTURE WORK

The main contribution of this paper is the distributed computation of 3D mesh local descriptors based on the concept of k -rings. In our implementation, we work with three main resources: processors, RAM and hard disk. We have parameters that allow us to fix the usage of these resources. Getting an optimal configuration will also depend on the cluster features. Automating the selection of these parameters is a future work.

The partitioning method used on this model is very simple (appearance order), Using a smart partitioning, such as a Euclidean distance-based partitioning, would optimize the execution time. In that case, a preprocessing step is needed in order to obtain balanced partitions.

We showed that our method is scalable, so increasing the cluster capacity will decrease the execution time. Also, there are some operations that can be executed using GPU. This will reduce considerably the execution time of the descriptor computation. We can extend our model for the computation of local descriptors over point clouds. In that case, we would have to use spatial data structures and a Euclidean distance-based partitioning with overlapping.

REFERENCES

- Aleardi, L., Devillers, O., & Schaeffer, G. (2005). Succinct representation of triangulations with a boundary. In F. Dehne, A. López-Ortiz, & J.-R. Sack (Eds.), *Algorithms and data structures*. WADS 2005. Lecture Notes in Computer Science, vol 3608 (pp. 134-145). Springer. https://doi.org/10.1007/11534273_13
- Areias, P., & Rabczuk, T. (2017). Steiner-point free edge cutting of tetrahedral meshes with applications in fracture. *Finite Elements in Analysis and Design*, 132, 27-41. <https://doi.org/10.1016/j.finel.2017.05.001>
- Balman, M. (2006). Tetrahedral mesh refinement in distributed environments. In T. M. Pinkston & F. Ozguner (Eds.), *Proceedings of the 2006 International Conference on Parallel Processing Workshops (ICPPW'06)* (pp. 498-504). IEEE Computer Society. <https://doi.org/10.1109/ICPPW.2006.72>
- Cabiddu, D., & Attene, M. (2015). Distributed processing of large polygon meshes. In A. Giachetti, S. Biasotti, & M. Tarini (Eds.), *Smart tools and apps for graphics—Eurographics Italian chapter conference* (pp. 139-148). The Eurographics Association. <https://doi.org/10.2312/stag.20151301>
- Castellani, U., Cristani, M., Fantoni, S., & Murino, V. (2008). Sparse points matching by combining 3D mesh saliency with statistical descriptors. *Computer Graphics Forum*, 27(2), 643-652. <https://doi.org/10.1111/j.1467-8659.2008.01162.x>

- Chen, X., Golovinskiy, A., & Funkhouser, T. (2009). A benchmark for 3D mesh segmentation. *ACM Transactions on Graphics*, 28(3), 1-12. <https://doi.org/10.1145/1531326.1531379>
- Cignoni, P., Montani, C., Rocchini, C., & Scopigno, R. (2003). External memory management and simplification of huge meshes. *IEEE Transactions on Visualization and Computer Graphics*, 9(4), 525-537. <https://doi.org/10.1109/TVCG.2003.1260746>
- Figueiredo, L., Ivson, P., & Celes, W. (2021). Deep learning-based framework for Shape Instance Registration on 3D CAD models. *Computers & Graphics*, 101, 72-81. <https://doi.org/10.1016/j.cag.2021.08.012>
- Gao, L., Cao, Y.-P., Lai, Y.-K., Huang, H.-Z., Kobbelt, L., & Hu, S.-M. (2015). Active exploration of large 3D model repositories. *IEEE Transactions on Visualization and Computer Graphics*, 21(12), 1390-1402. <https://doi.org/10.1109/TVCG.2014.2369039>
- Gelfand, N., Mitra, N. J., Guibas, L. J., & Pottmann, H. (2005). Robust global registration. In M. Desbrun & H. Pottmann (Eds), *Eurographics Symposium on Geometry Processing* (pp. 197-206). Alvey Vision Club. http://vecg.cs.ucl.ac.uk/Projects/SmartGeometry/global_registration/paper_docs/global_registration_sgp_05.pdf
- Gupta, O., & Rani, S. (2013). Accelerating molecular sequence analysis using distributed computing environment. *International Journal of Scientific & Engineering Research-IJSER*, 4(10), 262-265. <https://www.ijser.org/onlineResearchPaperViewer.aspx?Accelerating-Molecular-Sequence-Analysis-using-Distributed-Computing-Environment.pdf>
- Gurung, T., Laney, D., Lindstrom, P., & Rossignac, J. (2011). SQuad: Compact representation for triangle meshes. *Computer Graphics Forum*, 30(2), 355-364. <https://doi.org/10.1111/j.1467-8659.2011.01866.x>
- Gurung, T., Luffel, M., Lindstrom, P., & Rossignac, J. (2013). Zipper: A compact connectivity data structure for triangle meshes. *Computer-Aided Design*, 45(2), 262-269. <https://doi.org/10.1016/j.cad.2012.10.009>
- Harris, C., & Stephens, M. (1988). A combined corner and edge detector. In C. J. Taylor (Ed.), *Proceedings of the Alvey Vision Conference* (pp. 23.1-23.6). <http://dx.doi.org/10.5244/C.2.23>
- Herath, U., Tavazde, P., He, X., Bousquet, E., Singh, S., Muñoz, F., & Romero, A. H. (2020). PyProcar: A Python library for electronic structure pre/post-processing. *Computer Physics Communications*, 251, 107080. <https://doi.org/10.1016/j.cpc.2019.107080>
- Lee, C. H., Varshney, A., & Jacobs, D. W. (2005). Mesh saliency. *ACM Transactions on Graphics*, 24(3), 659-666. <https://doi.org/10.1145/1186822.1073244>

- Le Tien, M., Tan, K. N., & Raffin, R. (2021, 15-16 December). Analysis of geometrical features of 3D model based on the surface curvature of a set of point cloud. In *The 5th International Conference on Future Networks & Distributed Systems* (pp. 17-23). The Association of Computing Machinery. <https://doi.org/10.1145/3508072.3508076>
- Levoy, M., Pulli, K., Curless, B., Rusinkiewicz, S., Koller, D., Pereira, L., Ginzton, M., Anderson, S., Davis, J., Ginsberg, J., Shade, J., & Fulk, D. (2000). The digital Michelangelo project: 3D scanning of large statues. In *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques* (pp. 131-144). SIFFGRAPH. <https://doi.org/10.1145/344779.344849>
- Li, B., Lu, Y., Li, C., Godil, A., Schreck, T., Aono, M., Burtscher, M., Chen, Q., Chowdhury, N. K., Fang, B., Fu, H., Furuya, T., Li, H., Liu, J., Johan, H., Kosaka, R., Koyanagi, H., Ohbuchi, R., Tatsuma, A., Wan, Y, Zhang, C., & Zou, C. (2015). A comparison of 3D shape retrieval methods based on a large-scale benchmark supporting multimodal queries. *Computer Vision and Image Understanding*, 131, 1-27. <https://doi.org/10.1016/j.cviu.2014.10.006>
- Luffel, M., Gurung, T., Lindstrom, P., & Rossignac, J. (2014). Grouper: A Compact, Streamable Triangle Mesh Data Structure. *IEEE Transactions on Visualization and Computer Graphics*, 20(1), 84-98. <https://doi.org/10.1109/TVCG.2013.81>
- Maquart, T., Elguedj, T., Gravouil, A., & Rochette, M. (2021). 3D B-Rep meshing for real-time data-based geometric parametric analysis. *Advanced Modeling and Simulation in Engineering Sciences*, 8, 8. <https://doi.org/10.1186/s40323-021-00194-5>
- Mitra, N. J., Pauly, M., Wand, M., & Ceylan, D. (2013). Symmetry in 3D geometry: Extraction and applications. *Computer Graphics Forum*, 32(6), 1-23. <https://doi.org/10.1111/cgf.12010>
- O' Sullivan, E., Van de Lande, L. S., Papaioannou, A., Breakey, R. W. F., Jeelani, N. O., Ponniah, A., Duncan, C., Schievano, S., Khonsari, R. H., Zafeiriou, S., & Dunaway, D. J. (2022). Convolutional mesh autoencoders for the 3-dimensional identification of FGFR-related craniosynostosis. *Scientific Reports*, 12, 2230. <https://doi.org/10.1038/s41598-021-02411-y>
- Pavlakos, G., Zhu, L., Zhou, X., & Daniilidis, K. (2018). Learning to estimate 3D human pose and shape from a single color image. In L. O'Conner (Ed.), *Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 459-468). IEEE Computer Society; Conference Publishing Services. <https://doi.org/10.1109/CVPR.2018.00055>
- Pereira, R., Azambuja, M., Breitman, K., & Endler, M. (2010). An architecture for distributed high performance video processing in the cloud. In R. Bilof (Ed.), *Proceedings of*

- the *2010 IEEE 3rd International Conference on Cloud Computing* (pp. 482–489). IEEE Computer Society; Conference Publishing Services. <https://doi.org/10.1109/CLOUD.2010.73>
- Prajapati, H. B., & Vij, S. K. (2011). Analytical study of parallel and distributed image processing. In R Siddavatam & S. P. Ghreya (Eds.), *2011 International Conference on Image Information Processing* (pp. 1-6). The Institute of Electrical and Electronics Engineers. <https://doi.org/10.1109/ICIIP.2011.6108870>
- Rocca, L., De Giorgis, N., Panozzo, D., & Puppo, E. (2011). Fast neighborhood search on polygonal meshes. In A. F. Abate, M. Nappi, & G. Tortora (Eds.). *Eurographics Italian Chapter Conference 2011* (pp.15-21). The Eurographics Association. <https://doi.org/10.2312/LocalChapterEvents/ItalChap/ItalianChapConf2011/015-021>
- Sankaranarayanan, J., Samet, H., & Varshney, A. (2007). A fast all nearest neighbor algorithm for applications involving large point-clouds. *Computers & Graphics*, 31(2), 157-174. <https://doi.org/10.1016/j.cag.2006.11.011>
- Sipiran, I., & Bustos, B. (2011). Harris 3D: a robust extension of the Harris operator for interest point detection on 3D meshes. *The Visual Computer*, 27, 963. <https://doi.org/10.1007/s00371-011-0610-y>
- Squyres, J. M., Lumsdaine, A., McCandless, B. C., & Stevenson, R. L. (2000). *Parallel and distributed algorithms for high speed image processing*. University of Notre Dame. <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.939.2869&rep=rep1&type=pdf>
- Van Kaick, O., Zhang, H., Hamarneh, G., & Cohen-Or, D. (2011). A survey on shape correspondence. *Computer Graphics Forum*, 30(6), 1681-1707. <https://doi.org/10.1111/j.1467-8659.2011.01884.x>
- Verma, N., Boyer, E., & Verbeek, J. (2018). FeaStNet: Feature-steered graph convolutions for 3d shape analysis. In L. O'Conner, (Ed.), *Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 2598-2606). IEEE Computer Society; Conference Publishing Services. <https://doi.org/10.1109/CVPR.2018.00275>
- Vo, H. T., Bronson, J., Summa, B., Comba, J. L. D., Freire, J., Howe, B., Pascucci, V., & Silva, C. T. (2011). Parallel visualization on large clusters using MapReduce. In D. Rogers & C. T. Silva. *Proceedings of the 2011 IEEE Symposium on Large Data Analysis and Visualization* (pp. 81-88). IEEE Computer Society Press. <https://doi.org/10.1109/LDAV.2011.6092321>
- Warn, S., Emeneker, W., Cothren, J., & Apon, A. (2009, 31 de Agosto-4 de setiembre). *Accelerating SIFT on parallel architectures* [written presentation]. 2009 IEEE

International Conference on Cluster Computing and Workshops, Nueva Orleans, Louisiana, USA.. <https://doi.org/10.1109/CLUSTR.2009.5289155>

Zaharescu, A., Boyer, E., Varanasi, K., & Horaud, R. (2009, 20-25 de junio). *Surface feature detection and description with applications to mesh matching*. [written presentation] 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, Florida, Estados Unidos. <https://doi.org/10.1109/CVPR.2009.5206748>

Zamolo, R., Miotti, D., & Nobile, E. (2022). Numerical analysis of thermo-fluid problems in 3D domains by means of the RBF-FD meshless method. *Journal of Physics: Conference Series*, 2177, 012007. <https://doi.org/10.1088/1742-6596/2177/1/012007>

Zhou, R., Song, Z., & Lu, Y. (2017). 3D mesoscale finite element modelling of concrete. *Computers & Structures*, 192, 96–113. <https://doi.org/10.1016/j.compstruc.2017.07.009>

FOOTBALL PITCH CONDITION ANALYSIS BASED ON K-MEANS CLUSTERING

Héctor Eduardo Ugarte Rojas
hector.ugarte@unsaac.edu.pe / ORCID: 0000-0002-5338-2140
Departamento Académico de Ingeniería Informática
Universidad Nacional de San Antonio Abad del Cusco, Peru

Boris Chullo Llave
boris.chullo@unsaac.edu.pe / ORCID: 0000-0002-6844-7680
Departamento Académico de Ingeniería Informática
Universidad Nacional de San Antonio Abad del Cusco, Peru

Abstract

Football, a highly popular sport all over the world, requires that professional footballers practice it on a field of play in ideal conditions, which, among other things, includes the usage and maintenance of healthy natural grass. In this study, we present an unsupervised allocator strategy for image analysis of football pitches that uses *k*-means clustering and color comparison to assess whether a playing field is in good or bad condition. Our approach considers proportions of dominant RGB colors for automatized decision-making. We developed a prototype and tested it with a series of images; this paper offers a comparison between the findings of this test and our expected results.

KEYWORDS: image analysis / *k*-means algorithm / dominant colors / clustering / football

ANÁLISIS DE LA CONDICIÓN DEL CAMPO DE FÚTBOL BASADO EN EL AGRUPAMIENTO DE *K-MEANS*

Resumen

RESUMEN. El fútbol, un deporte muy popular en todo el mundo, requiere que los futbolistas profesionales lo practiquen en un campo de juego en condiciones ideales, lo que, entre otras cosas, incluye el uso y mantenimiento de un césped natural saludable. En este estudio, presentamos una estrategia de asignación sin supervisión para el análisis de imágenes de campos de fútbol que utiliza agrupamiento *k-means* y comparación de colores para evaluar si un campo de juego está en buenas o malas condiciones. Nuestro enfoque considera las proporciones de los colores RGB dominantes para automatizar la toma de decisiones. Para tal fin, se desarrolló un prototipo que se probó con una serie de imágenes; los resultados obtenidos se compararon con los esperados.

PALABRAS CLAVE: análisis de imágenes / algoritmo *k-means* / colores dominantes / clustering / fútbol

1. INTRODUCTION

The world's most popular game is football (usually called soccer, mainly in the United States) (Szymanski, 2014). It is a team sport involving two squads of eleven players. The game's objective is to score a ball into the opposing goal using any part of the body except the hands and the arms. Only the goalkeepers may use their hands or arms (inside a pitch area called the penalty area) to prevent the opposing team from scoring. The team scoring the greater number of goals is the match's winner, which usually lasts 90 minutes. It is usually played on a rectangular grass field but can also be played on artificial turf.

A natural grass field is highly preferred to artificial turf for the professional practice of the sport. Artificial turfs should only be used when grass maintenance is problematic due to adverse weather (very wet, extremely dry, or frequent snow). The International Federation of Football Association (a.k.a. FIFA - Fédération Internationale de Football Association) recommends field of play dimensions fo 105 meters in length and 68 meters in width. An optimal playing surface must be flat and even with complete grass coverage; it must also be free of potholes. The playing field is marked with lines following pitch dimensions (Football NSW Limited, 2015).

Although FIFA highly recommends playing on an optimal sports ground, it is common in regions with developing countries such as Latin America to play professional football on inadequate pitches. The condition of a playing field is important not simply for aesthetic reasons: it can also affect players' safety and the quality of play. Normally, the condition of a football ground is evident to the eye. The color of healthy grass should always be green it should be maintained by professionals and the sward sould be constantly cut short. Figure 1 shows an artificial turf that might easily be mistaken for a natural field in excellent condition.

Figure 1

A football field at Brastad arena, Brastad, Lysekil Municipality, Sweden



Note. Carter (2019)

Figure 2

The football field at Zelengaj, Dugave neighbourhood, Zagreb, Croatia



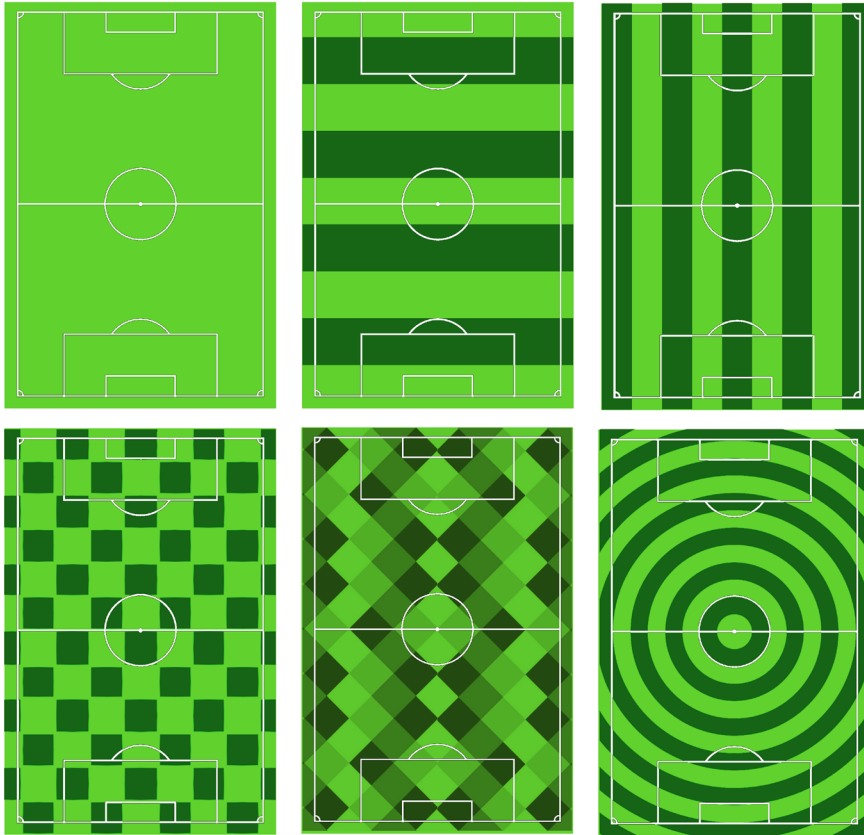
Note. Radovanović (2011)

Pitches in bad condition show brownish or yellowish colored patches of dead grass, dirt field spots, and others, as is evident in figure 2. Football fans and journalists know the importance of pitches in good condition, so it is a common practice to paint bad grass fields green. For instance, *El País* (2018) informed that the field of the historic Uruguayan Centenario stadium was painted for a final match after being damaged by the stage set up for a rock concert.

As we see, pitches in good condition are all green with white marking painted lines. Different patterns can be created depending on how the sward is cut. In figure 3, we present some popular patterns where different shades of green are present.

Figure 3

Popular football playing field grass patterns



Solomon and Breckon (2011) define image analysis as extracting useful information from images, mostly digital, using image processing techniques. It has a wide range of applications in astronomy, medicine, and robotics, to name a few. The process of splitting a digital image into multiple image segments is known as image segmentation. This technique makes the representation of an image more meaningful and easier to examine by simplifying and/or changing it (Stockman & Shapiro, 2001).

This research paper presents a technique to establish the condition of a football pitch based on RGB image analysis. *K*-means clustering was chosen for this analysis since it is a popular technique to obtain dominant colors used by different researchers (Pavan Kumar et al., 2020; Sharma, 2021). After comparing each dominant color with shades of green, a decision has to be made according to an analysis. For validation, a prototype has been implemented and tested with a set of pictures, where the computer analysis results obtained are compared with those of simple visual examination.

2. BACKGROUND

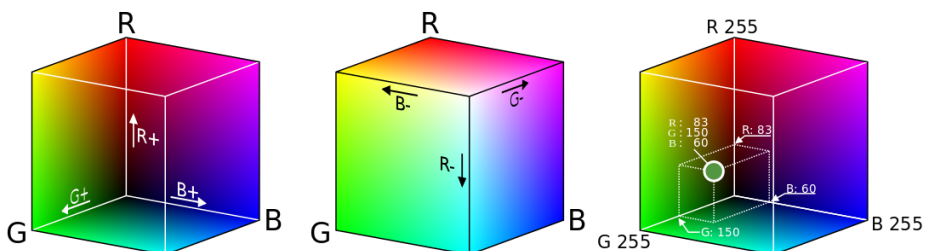
2.1 RGB color model

A color model is a systematic method for generating a wide range of colors from a small number of primary colors (Rhyne, 2016). Red, Green, and Blue are the main colors in the RGB additive color model. It is used in many applications such as digital images, color scanners, digital cameras, and digital screens (Loesda et al., 2014). Each color is expressed as a triple from 0 (least intense) to 255 (most intense). Therefore, there are 16,777,216 ($256 \times 256 \times 256$) possible combinations. Some representations for popular colours are: red (255,0,0), green (0,255,0), blue (0,0,255), black (0,0,0), white (255,255,255), yellow (255,255,0), cyan (0,255,255), magenta (255,0,255).

As a result of representing each color in a 3D space, we obtain a cube where we get a wider visual spectrum of the RGB color space, as seen in figure 4. R, G, and B axes are used.

Figure 4

RGB color cube



Note. The R, G, and B axes and their corresponding positive directions are shown on the left. In the center we see the cube's external faces and the negative R, G, and B directions, while on the right, we see how an inside point is represented in the RGB color model. Source: Maklaan (2015).

2.2 K-means clustering

K-means is also known as Lloyd–Forgy algorithm (Lloyd, 1982; Forgy, 1965). An unsupervised clusterization system divides objects into *k* groups depending on their attributes. Each of the data points will be allocated to a cluster with the closest mean, and the “centroid” or “center” of any cluster is its mean. Na et al. (2010) explain that the algorithm consists of two phases. First, select *k* centers randomly and later take each data object to the nearest center, commonly using Euclidian distance. Once all objects belong to one cluster, an iterative process of recalculating the average of the earlier clusters is done.

After the criterion function becomes the minimum, the algorithm finishes. We can formalize such a process with the following formula:

$$E = \sum_{j=1}^k \sum_{i=1}^n a_{ij} \|x_i - c_j\|_2^2 \quad (1)$$

Where:

If $x_i \in j$ Cluster

$a_{ij}=1$

else:

$a_{ij}=0$

E is the criterion function that is minimized, k is the number of clusters, n is the number of objects, x_i is the object, and c_j is the center of cluster j. The pseudocode is:

K-means clustering algorithm pseudocode

Input:

X = {x1,x2,...,xn} //Set of n data objects

k //Number of desired clusters

Output:

A set of k clusters.

Step:

Begin

1. Randomly choose k objects from X as initial centroids.

2. Repeat

 Assign each item x_i to the cluster with the nearest centroid.

 Calculate the new mean for each cluster

 Until convergence criteria are met.

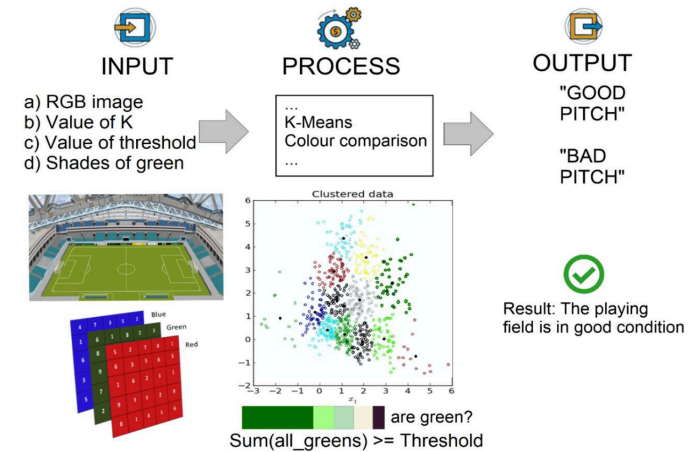
End

3. PROPOSED ALGORITHM

A general overview of our approach is presented in figure 5.

Figure 5

General overview of our proposed approach



As far as we know, our work is the first to focus on football pitch condition analysis. Our technique to determine if an image shows a football field in good condition is mainly based on identifying dominant colors and the percentage of pixels assigned to them for later determining if they are considered a shade of green. If the total percentage of shades of green is greater than a threshold value, the pitch is considered in good condition; otherwise, it isn't. Each centroid in k clusters represents a dominant color in the input RGB image. Only the football pitch region should be processed, avoiding other regions; that is why a method to crop only the pitch is needed.

Proposed algorithm pseudocode to determine the condition of a football playing field

Input:

Image //Input image as an RGB array containing a football pitch

K //Number of clusters

Threshold //Value of Threshold

Shades_of_green //RGB array containing all shades of green

Output:

Result //Result after applying our approach

Step:

Begin

1. Apply K -means clustering to the Image using the RGB color model.

2. Obtain proportions of pixels assigned to each dominant color.
 3. Create a variable Percentage = 0.
 4. Repeat
 - FOR EACH i=dominant color: IF i is in Shades_of_green.
Percentage = Percentage + Proportions(i).
//Proportions(i) returns the percentage of pixels assigned to i
 5. IF Percentage > Threshold //Threshold value
Result = "Pitch in good condition"
 6. ELSE
Result = "Pitch in bad condition"
- End

Optionally, we can apply a color correction, sharpening technique, and noise removal to the image before processing it. A set of RGB colors can be chosen in advance to decide if a color is a shade of green; however, this decision is somewhat subjective because not everyone agrees on what constitutes green. The image should ideally display an empty pitch (no players, no stands, etc.). It is preferable to use pictures that display as much of the field as possible, like airborne images in optimal resolution.

4. RESULTS

4.1 Prototype

We developed a prototype coded in Python. The modules that we used are: opencv-python: not official pre-built CPU-only OpenCV (Computer Vision library) packages; NumPy: the core Python module for scientific computing; scikit-learn: tools for predictive data analysis that are simple and efficient; tkinter: ("Tool kit interface") is the standard Python interface to the Tool Command language/Tk graphical user interface toolkit; Pillow: The Python Imaging Library extends the capabilities of the Python interpreter to include image processing. In table 1 we present the usage that we are doing with these modules:

Table 1

Python modules usage

Module name	Usage
opencv-python	<ul style="list-style-type: none"> • For reading the input image • Resizing the image • Creating a polygon • Creating a mask • Sharpening the image.
NumPy	To remove black pixels from the mask and array manipulation.
scikit-learn	For applying <i>k</i> -means directly to RGB pixels.
Tkinter	To create frames, canvas, file dialogs, buttons, and text regions.
Pillow	To create and modify Tkinter BitmapImage and PhotoImage objects.

We manually delineate the football pitch region, drawing a polygon with the mouse as explained below. The resulting image is processed, ignoring the black color (color of the mask).

In our prototype, we consider a value of 6 for *k*, and there are different ways to determine this value. We are using the “values of *K* equated to the number of classes” method (Pham, S., & Nguyen, 2005). This seems a reasonable choice for this purpose, considering possible classes such as a few shades of green, dirt detection, dead grass, white marking lines, and other possible objects. We consider a pitch in good condition if at most one-fifth is not green. This value is just for testing purposes. Therefore, our threshold value is equal to 80.

Our prototype is available in the GitHub repository¹. A normal execution flow of our software solution presented in figure 6 is:

1. Choose a file image (JPG, PNG, etc.) (1).
2. The image is loaded in the visualization area (2).
3. Draw a polygon using the mouse over the region to delimit where the football playing field is present, considering:
 - Left-click to add points to the polygon.
 - Double left-click to finish the polygon.
 - Right-click to remove the polygon.
4. Once the cropping polygon is finished and shown in red, press Perform *K*-means button.

¹ <https://github.com/hedugaro/Football-pitch-condition-analyzer>

- 5. Obtained results are shown, including the final verdict about the condition of the pitch. The 6 colors of the centroids are also shown in (3) and their proportions in a histogram (4).

Figure 7 shows an example of the result when a pitch in bad condition is processed.

Figure 6

The interface of our developed prototype



Note. (1) Buttons and results frame, (2) Visualization area frame, (3) The six dominant colors frame, (4) Dominant colors histogram frame.

4.2. Experiments

Our experiments were conducted using the Python Language (version 3.8.5) on a 2.70GHz Intel Xeon E-2176M CPU with 22.0 GB of RAM memory and running Windows 10 version 21H2 as the operating system.

Even though our approach is based on unsupervised learning, we need a useful statistic for validation. We decided to use “external cluster validation” and compare our results with known (labeled) images. (Kassambara, n.d.). We used 80 different images of football pitches from around the world that we found using internet search engines. These photos were captured at different times of the day and from various viewpoints. As previously noted, we prefer images that solely show the playing field and no other objects for greater accuracy. By simple human visual inspection, 40 images show pictures in good condition, and 40 show pictures in bad condition. Table 2 summarizes the findings.

Table 2

Experiments results for our approach and prototype

Condition	Expected quantity	Obtained Quantity	Accuracy
Good playing field	40	38	95 %
Bad playing field	40	33	82,5 %

5. CONCLUSIONS AND FUTURE WORK

The proposed method using *k*-means clustering and color comparison with shades of green was able to identify the condition of a football pitch. Based on the results, we examined our allocator's accuracy and can conclude that our prototype performs well with high accuracy in both cases (95 % and 82,5 %): good and bad pitch conditions.

Some work that we plan to do in the future:

- a. Research other possible learning techniques and methods such as supervised learning.
- b. Test with other parameter values such as *k*, threshold, and shades of green and measure results from accuracy.
- c. Try with color models different from RGB that could be closer to how humans perceive colors.
- d. Recognize the football pitch region automatically from an image and possible external elements that should not be processed.
- e. Consider more possible outputs than "regular condition pitches".
- f. Process images in high resolution much faster using GPU processing power and CUDA parallel programming and permit batch image processing.

REFERENCES

- Carter, W. (2020). Corner flag in the soccer field at Brastad arena [Photograph]. Wikimedia Commons. https://commons.wikimedia.org/wiki/File:Soccer_field_at_Brastad_arena_6.jpg
- El País. (2018, November 9). Así fue como pintaron el césped del Centenario de verde. <https://www.elpais.com.uy/ovacion/futbol/asi-pintaron-cesped-centenario-verde.html>

- Football NSW Limited. (2015, November 6). Field marking & equipment. A guide to preparing your field for football. <https://footballnsw.com.au/wp-content/uploads/2017/06/Field-Markings-and-Equipment.pdf>
- Forgy, E. W. (1965). Cluster analysis of multivariate data: efficiency versus interpretability of classifications. *Biometrics*, 21, 768-780.
- Kassambara, A. (n.d.). Cluster validation statistics: Must know methods. *Datanovia*. <https://www.datanovia.com/en/lessons/cluster-validation-statistics-must-know-methods/>
- Lloyd, S. (1982). Least squares quantization in PCM. *IEEE Transactions on Information Theory*, 28(2), 129-137. <https://doi.org/10.1109/TIT.1982.1056489>
- Loesdau, M., Chabrier, S., & Gabillon, A. (2014). Hue and Saturation in the RGB Color Space. In A. Elmoataz, O. Lezoray, F. Nouboud, & D. Mammass, (Eds.), *Image and Signal Processing. ICISP 2014. Lecture Notes in Computer Science, vol 8509*. Springer. https://doi.org/10.1007/978-3-319-07998-1_23
- Maklaan. (2015) A RGB color cube explained with three diagrams [Diagram]. Wikimedia Commons. https://commons.wikimedia.org/wiki/File:RGB_color_cube.svg
- Na, S., Xumin, L., & Yong, G. (2010). Research on k -means clustering algorithm: An improved k -means clustering algorithm. In F. Yu, X. Peng, H. Liu, J. Shiu, & R. Ng (Eds.), *Proceedings of the Third International Symposium on Intelligent Information Technology and Security Informatics* (pp. 63-67). IEEE Computer Society; Conference Publishing Services. <https://doi.org/10.1109/IITSI.2010.74>
- Pavan Kumar, I., Hara Gopal, V. P., Ramasubbareddy, S., Nalluri, S., & Govinda, K. (2020). Dominant color palette extraction by k -means clustering algorithm and reconstruction of image. In K. Raju, R. Senkerik, S. Lanka, & V. Rajagopal (Eds.), *Data engineering and communication technology, vol 1079* (pp. 921-929). Springer Singapore. https://doi.org/10.1007/978-981-15-1097-7_78
- Pham, D. T., Dimov, S. S., & Nguyen, C. D. (2005). Selection of K in K -means clustering. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, 219(1), 103-119. <https://doi.org/10.1243/095440605X8298>
- Radovanović, B. (2011). *NK Zelengaj football pitch in Dugave neighborhood, Zagreb, Croatia* [Photography]. Wikimedia Commons. https://commons.wikimedia.org/wiki/File:NK_Zelengaj_football_pitch_20110918_3186.jpg
- Rhynne, T.-M. (2016). *Applying color theory to digital media and visualization*. CRC Press. <https://doi.org/10.1145/2776880.2792696>

- Sharma, A. (2021, December 9). How to find the most dominant colors in an image using *kmeans* clustering —with source code— interesting project. *Towards Dev.* <https://towardsdev.com/how-to-find-the-most-dominant-colors-in-an-image-in-python-using-kmeans-clustering-with-source-527ef3e6775f>
- Solomon, C. & Breckon, T. (2011). *Fundamentals of digital image processing: A practical approach with examples in Matlab*. John Wiley & Sons, Ltd. <https://doi.org/10.1002/9780470689776>
- Stockman, G., & Shapiro, L. G. (2001). *Computer vision*. Pearson.
- Szymanski, S. (2014). It's football not soccer. <http://ns.umich.edu/Releases/2014/June14/lts-football-not-soccer.pdf>

ESTUDIO DE LA USABILIDAD DE UNA APLICACIÓN DE REALIDAD AUMENTADA EN LA CONSTRUCCIÓN BÁSICA DE UN BRAZO ROBÓTICO

JOSE MALPARTIDA VALVERDE
jose.malpartida.v@uni.pe / orcid: 0000-0003-3696-5827
Universidad Nacional de Ingeniería (UNI), Lima, Perú

JOSE MARTIN LOZANO APARICIO
jlozano@uni.edu.pe / orcid: 0000-0003-1086-5799
Universidad Nacional de Ingeniería (UNI), Lima, Perú

Resumen

La realidad aumentada (RA) es una tecnología que puede ayudar en diferentes áreas de la investigación científica, como en la manipulación de equipos costosos, fáciles de malograr o peligrosos. Los componentes para la construcción de un robot pueden llegar a tener un costo considerable. La realidad aumentada puede ser aplicada en la fabricación de robots para reducir costos, ya que permite experimentar el manejo de componentes robóticos sin necesidad de comprarlos. El estudio tiene como objetivo medir la calidad de la experiencia de usuario con una aplicación de realidad aumentada en el ámbito de la robótica. Para ello se creó una aplicación que utiliza los gestos de las manos para ensamblar un brazo robótico. En este artículo se presentan los resultados obtenidos, así como los índices de eficiencia de los usuarios en el ensamblaje del brazo robótico y la usabilidad de la aplicación.

PALABRAS CLAVE: realidad aumentada / brazo robótico / rastreo de mano / imágenes 3D / usabilidad

STUDY OF THE USABILITY OF AN AUGMENTED REALITY APPLICATION IN THE BASIC CONSTRUCTION OF A ROBOTIC ARM

Abstract

Augmented reality (AR) can help in different areas of scientific research, such as handling expensive, easy-to-break or dangerous equipment. Since the components for the construction of a robot can be considerably expensive, augmented reality can be applied in the manufacture of robots to reduce costs, allowing users to experience handling robotic components without having to buy them. This study aims to measure the quality of user experience with an augmented reality application in robotics. To do this, we created an application that uses hand gestures to assemble a robotic arm. This article presents the results obtained, the user's efficiency rates in assembling the robotic arm, and an assessment of the application's usability.

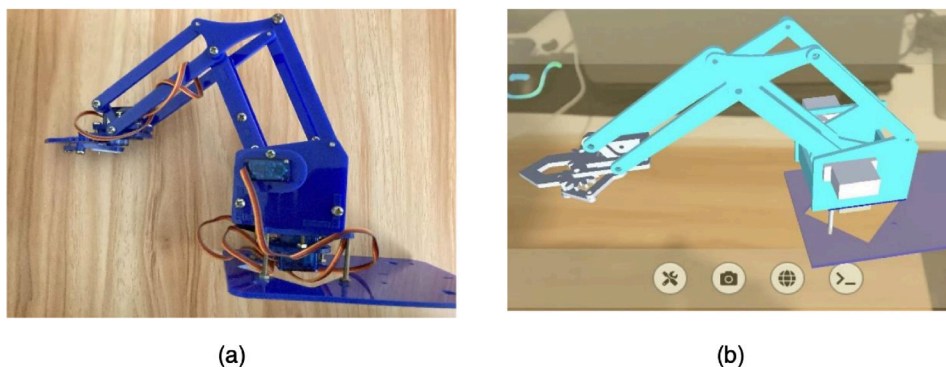
KEYWORDS: augmented reality / robotic arm / hand tracking / usability

1. INTRODUCCIÓN

En el campo industrial, el brazo robótico permite generar soluciones viables y polivalentes, pues logra una automatización mediante tareas continuas que implican movilizar objetos de distintos pesos combinados con presión. Un brazo robótico está compuesto por eslabones y articulaciones (figura 1a). Sus montajes se realizan en diversos ángulos y posiciones. Para esto se debe hallar una forma en la cual el brazo robótico realice las tareas que la persona desea ejecutar. A partir de esta necesidad, los humanos crearon interfaces como medio de comunicación para lograr que los robots elaboren las asignaciones que se les ordenan. Debido a ello, la comunicación entre humano-robot debe darse de una manera sencilla y de fácil comprensión para una mayor eficiencia y eficacia con respecto al trabajo que se realizará (Chacón Malasquez & Tuiro Achulle, 2018).

Figura 1

Brazo robótico



Nota. Ejemplos de brazo robótico físico (a) y vista virtual (bB).

Sin embargo, el diseñar y elaborar de manera industrial un brazo robótico puede llegar a ser costoso y poco rentable si se quieren realizar pequeñas mejoras posteriormente. Además, para fabricarlo se deben usar máquinas y/o herramientas que pueden poner en riesgo la integridad física de la persona en el transcurso de este proceso. Una solución beneficiosa y segura es utilizar modelos 3D de los componentes del brazo robótico.

El término "3D" proviene de la palabra "tridimensional", que conjuga la utilización de los tres ejes de coordenadas X, Y, Z. Con el avance de la tecnología hubo una mejora en efectuar los cálculos para la creación de una imagen en tres dimensiones con respecto a su altura, anchura y profundidad. El proceso de la creación de una imagen 3D se basa en múltiples fórmulas matemáticas que juntan objetos poligonales, texturas, sombras,

reflejos, entre otros. Todo este grupo de elementos hace posible la creación de un modelo 3D. Al modelo 3D se le aplica una transformación llamada "renderización" para producir una imagen 3D. Para el presente proyecto, la imagen 3D fue sometida a la función de traslación y a la función de rotación para el objetivo principal del sistema, la fabricación interactiva del brazo robótico (véase figura 1b).

Para la fabricación interactiva usaremos MediaPipe, que es una librería para *machine learning* que tiene un método para el reconocimiento de manos, el cual es el modelo de *hand landmark*. Este modelo consiste en obtener el área reconocida como la mano a través de la cámara y poder ubicar los veintiún (21) puntos de referencia de la imagen dada por la detección de la mano.

De esta manera, se creará la aplicación de realidad aumentada (RA) con interacciones mediante los gestos de la mano y se analizará el impacto que lleva consigo utilizar este tipo de programa en usuarios inexperimentados en el tema. En este sentido, en la sección 2 describimos MediaPipe y la realidad aumentada. Luego, en la sección 3 presentamos un breve resumen de trabajos parecidos. La construcción de la aplicación se presenta en la sección 4. La evaluación de la aplicación es descrita en la sección 5 y, finalmente, las discusiones y conclusiones.

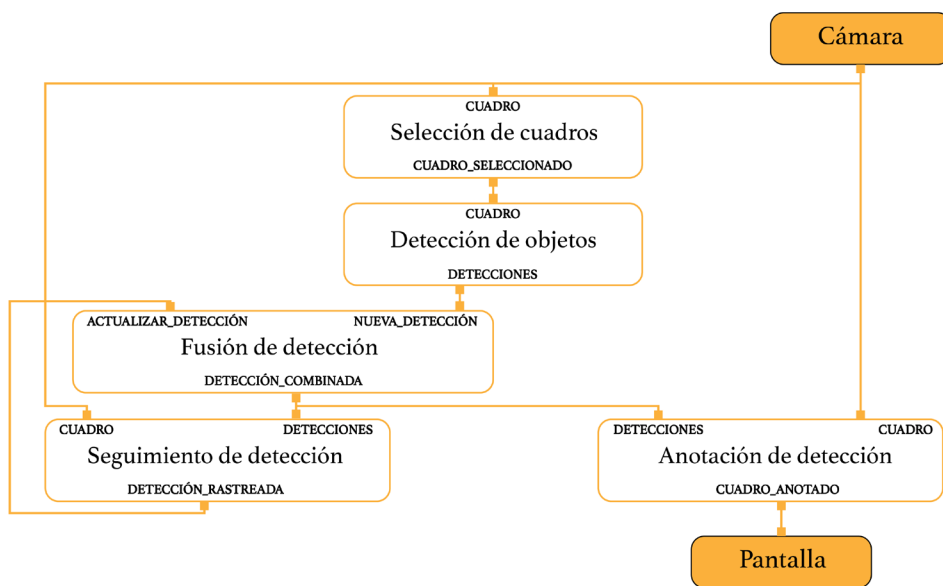
2. CONCEPTOS PREVIOS

2.1 MediaPipe

Es una librería para inteligencia artificial (IA) de Google disponible en dispositivos móviles Android, aunque también posee APIs para iOS, Python y JavaScript. MediaPipe a través de la cámara del dispositivo móvil percibe los movimientos y la estructura de las manos por medio del aprendizaje automático, permitiendo gesticular el lenguaje de señas para comunicar a Google lo que se desea transmitir (Lugaresi et al., 2019). Tiene un módulo de rastreo de la palma y dedos de la mano que detecta 21 puntos claves 3D de un solo cuadro (*frame*). Este proceso se inicia obteniendo un *frame* captado por la cámara, y se le designa a tres procesos: selección de cuadros, seguimiento de detección, anotación de detección (véase figura 2).

Figura 2

Arquitectura de MediaPipe

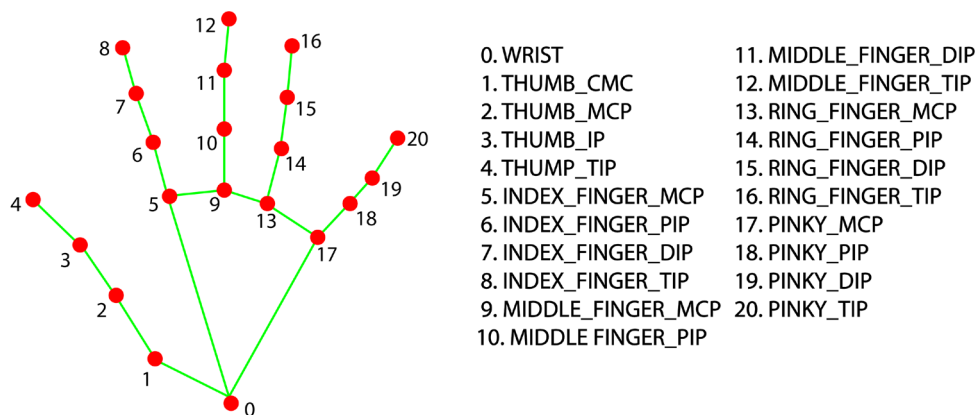


Nota. Adaptado de "MediaPipe: A Framework for Perceiving and Processing Reality", por C. Lugaresi, J. Tang, H. Nash, C. McClanahan, E. Uboweja, M., Hays, F. Zhang, C.-L. Chang, M. Yong, J. Lee, W.-T.Chang, W. Hua, M. Georg, y M. Grundmann, 2019, *Third Workshop on Computer Vision for AR/VR at IEEE Computer Vision and Pattern Recognition (CVPR) 2019*, p. 1 (https://mixedreality.cs.cornell.edu/s/NewTitle_May1_MediaPipe_CVPR_CV4ARVR_Workshop_2019.pdf).

Entonces, según la figura 2 sobre el proceso de selección de cuadros, se envían los *frames* al proceso de detección de objetos si existe algún objeto. En el proceso de detección de objetos, se ubican y seleccionan todos los objetos que se encuentren en el *frame* y se envía la información al proceso de fusión de detección. También, en el proceso de seguimiento de detección, se actualiza el rastreo de la solución indicada en cada *frame* recibido por la cámara y siempre se actualiza con la solución de rastreo de manos. En el proceso de fusión de detección, se unen los procesos de seguimiento de detección y el de detección de objetos, es decir, solo se muestran los objetos con la solución proporcionada por seguimiento de detección. Debido a esto, la aplicación construida solo mostrará los objetos interpretados como las manos. Finalmente, en el proceso de anotación de detección, en el *frame* recibido por la cámara, se le añaden los *landmarks* como esferas rojas y las uniones entre los *landmarks* como segmentos verdes y, posteriormente, se muestra en la pantalla del dispositivo (véase figura 3).

Figura 3

Landmark de la mano



Nota. Descrita por MediaPipe (2020).

2.2 Realidad aumentada

La realidad aumentada (RA) (Azuma, 1997) está constituida por tres características principales: mezcla objetos reales y virtuales en un entorno del mundo real, interactividad que se presenta en tiempo real, y registra y alinea elementos virtuales con objetos físicos en un entorno 3D. A inicios del desarrollo de la realidad aumentada se usaban *displays* montados en la cabeza del usuario. Sin embargo, en la actualidad, la implementación de la realidad aumentada se expande a dispositivos móviles, como, por ejemplo, los teléfonos inteligentes (*smartphones*). A continuación, presentamos técnicas de interacción en realidad aumentada.

2.2.1 Técnicas de interacción que utilicen movimientos corporales

Este tipo de técnica se basa en la interacción de los gestos de la mano a partir del reconocimiento y seguimiento. Las tecnologías usadas para estas funcionalidades se realizan mediante el seguimiento inercial, seguimiento magnético, sistemas basados en visión computacional, entre otros. En este trabajo se utilizó la visión computacional mediante MediaPipe.

2.2.2 Dispositivos como herramientas de interacción

Los dispositivos más frecuentes en sistemas o aplicaciones de realidad aumentada son los celulares inteligentes. Son usados por medio de reconocimiento de las manos y los gestos que realiza para determinar ciertas interacciones.

2.2.3 Interacción multimodal

En términos generales, los sistemas multimodales se constituyen como aquellos que procesan como entradas combinaciones de diversos formatos y estímulos (auditivos, táctiles, gestuales, entre otros). La implementación de estos sistemas a la realidad aumentada genera muchas posibilidades para que el usuario interactúe con los elementos que proporciona la RA. En el presente proyecto se puede observar como las interacciones con los gestos de las manos dan lugar a las diversas funcionalidades de la aplicación desarrollada.

3. TRABAJOS RELACIONADOS

Pérez López (2021), en su tesis de grado, que afianza la investigación de Pasaréti et al. (2011), describe como se utiliza la realidad aumentada para la docencia. Uno de sus objetivos es modificar los elementos 3D de manera que el alumno pueda visualizarlos con mayor claridad. Esto se realiza mediante ciertos gestos particulares de las manos que son captados por una cámara web. De ello se pueden rescatar los métodos de interacción que realiza el docente para la manipulación de los modelos 3D. En el presente trabajo la aplicación realiza la traslación, la rotación y el escalamiento de objetos tridimensionales de forma similar a lo desarrollado en el trabajo antes mencionado.

Vogel et al. (2017) plantean en su artículo que la persona trabaje con un robot de realidad aumentada en un proceso de manufactura y los movimientos realizados por las manos sean captados por sensores. Estos sensores enviarán la información al robot real y ejecutará sus funciones en el proceso de manufactura.

Fang et al. (2014) plantean una interfaz basada en realidad aumentada para la planificación de trayectorias de robots. Las designaciones de las trayectorias están dadas por un objeto real en donde se señalan puntos referenciales que son captados por la cámara. Esta unión de puntos crea un camino por donde el robot ejecutará su movimiento. En nuestro caso, el objeto referencial es la mano cuyo movimiento es usado para el desplazamiento del robot.

De la Cruz et al. (2020) dan a conocer la simulación de trayectorias con algoritmos avanzados para ser visualizadas con la animación del brazo robótico en RA. También se puede manipular cada grado de libertad según el ángulo ingresado. Esta función se agregará en un trabajo futuro; es decir, según el gesto de la mano captada por la cámara del dispositivo, se manipulará cada grado de libertad y será visualizada en el robot de realidad aumentada.

El estudio de Ziden et al. (2022) tuvo como objetivo determinar el impacto de su aplicación de realidad aumentada en el rendimiento y la motivación de los estudiantes. Para ello, se usó la investigación de métodos mixtos que busca determinar la influencia

de las tarjetas de aprendizaje AR-NutricARd en la motivación y el rendimiento de los estudiantes de ciencias. De la misma manera, para el estudio de la usabilidad de la aplicación, se hizo uso del método mixto que combina la perspectiva cuantitativa y cualitativa, con el objetivo de darle profundidad al análisis de la investigación.

4. CONSTRUCCIÓN DE LA APLICACIÓN BASADA EN REALIDAD AUMENTADA

La aplicación se desarrolló con el entorno de trabajo (*framework*) AR Foundation en Unity, que es una multiplataforma para el desarrollo de videojuegos. Este *framework* es compatible con dispositivos que posean los *kits* de desarrollo de *software* ARCore (Android) o ARKit (iOS), por lo que se implementa en un iPhone SE. Además, se utilizó la solución del rastreo de mano proporcionado por MediaPipe. Por un lado, AR Foundation se encarga de que se le añada la realidad aumentada a la aplicación. Esto se da en mezclar el modelo 3D del brazo robótico, que es el objeto virtual, y los gestos de la mano en un entorno del mundo real. El modelo de brazo robótico utilizado es el presentado en la figura 1A. Por otro lado, las interacciones mediante gestos de la mano se encargan de realizar distintas operaciones con el objeto 3D; es decir, se comunica con el robot mediante los gestos de la mano para realizar distintas operaciones. Entre ellas están la traslación, rotación, selección e instanciación.

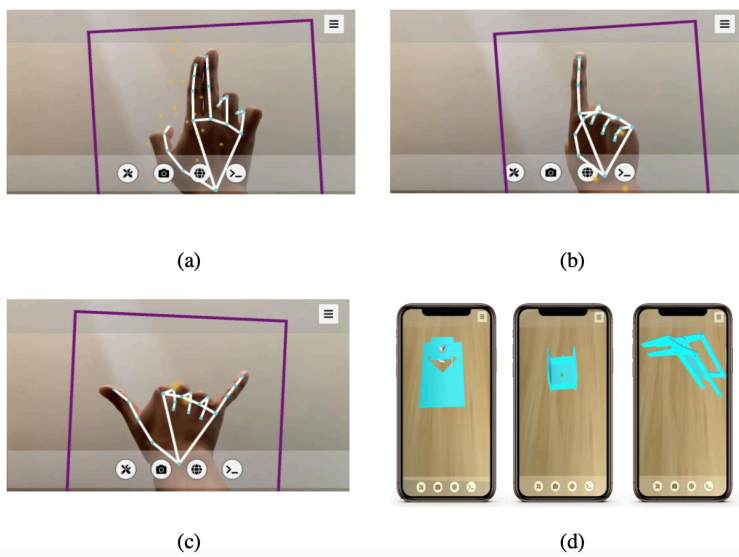
Instanciar. El gesto realizado es el de levantar solo el dedo índice y el dedo medio, que al juntarlos efectúan la instancia del objeto. Se decidió que este gesto sería el más adecuado porque son los dos dedos utilizados en el *mouse* y con la distancia que existe entre ellos se logra tener un mayor control sobre la función (véase figura 4a). La figura 4d presenta los componentes instanciados del robot.

Seleccionar un objeto virtual. El gesto a realizar es el de levantar solo el dedo índice, que al señalar al objeto que se quiere seleccionar cambiará de color. Esto se debe a que es el gesto común de las personas para señalar o apuntar un objeto (véase figura 4b).

Deseleccionar un objeto virtual. El gesto a realizar es el de levantar solo el dedo pulgar y el dedo meñique, que al efectuar la función, los objetos volverán a su color original (véase figura 4c).

Figura 4

Instanciación, selección y desección



Trasladar un objeto virtual. Para la función de traslación se realizaron dos gestos. Un primer gesto, ya mencionado, como se muestra en la figura 4a, para trasladar los objetos por el plano XZ según el algoritmo de la figura 5a. Se determinó de esta manera que la función de instanciar un objeto es similar a la traslación por este plano. También para que el usuario no tenga que memorizar más gestos. El otro gesto es el de levantar el dedo pulgar y el dedo índice (figura 7a, figura 7b). Este gesto trasladará el objeto hacia arriba y si el dedo índice baja, el objeto virtual se trasladará en sentido contrario siguiendo los pasos del algoritmo de la figura 5b. La razón por la que se tomó este gesto es la similitud en la acción de deslizar con el dedo hacia arriba y hacia abajo en un *smartphone*.

Figura 5

Algoritmo de traslación horizontal y vertical

```
Procedimiento TraslacionHorizontal()  
  // Comentario:  $L = (L_1, L_2, \dots, L_n)$  es una lista de eslabones //  
  Repetir  
    Si gesto_instanciar Entonces  
      Si  $\exists L_i$  Entonces  
        InstanciarEstabon( $L_i$ , gesto_instanciar.posicion, gesto_instanciar.rotacion)  
      Sino  
         $L_i$ .AsignarPosicionYRotacion(gesto_instanciar.posicion, gesto_instanciar.rotacion)  
    Fin Si  
  Fin Si  
Hasta Que FinalizarApp()  
Fin Procedimiento
```

(a)

```
Procedimiento TraslacionVertical()  
  // Comentario:  $L = (L_1, L_2, \dots, L_n)$  es una lista de eslabones //  
  Repetir  
    Si gesto_trasladar_arriba Entonces  
       $L_i$ .Trasladar(Vector3.Up * Tiempo)  
    Sino Si gesto_trasladar_abajo Entonces  
       $L_i$ .Trasladar(Vector3.Down * Tiempo)  
    Fin Si  
  Hasta Que FinalizarApp()  
Fin Procedimiento
```

(b)

Rotar un objeto virtual. El gesto a realizar es similar a la acción de desplazarse hacia la derecha y hacia la izquierda en un *smartphone*. No obstante, este gesto ya es usado para la traslación, por lo que se le agregó el dedo medio para diferenciarlo (véase figura 7c y figura 7d).

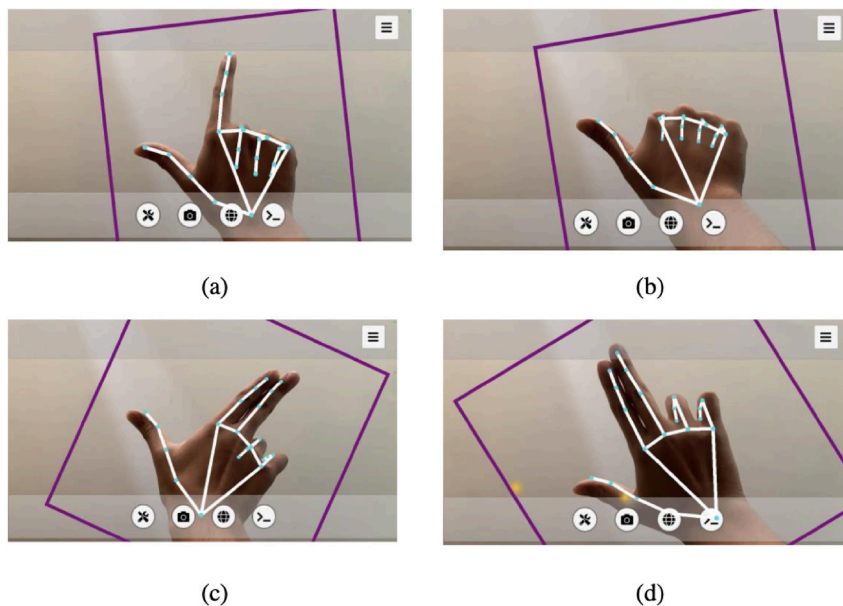
Figura 6

Función de rotación en el eje Y

```
Procedimiento Rotacion()  
  // Comentario:  $L = (L_1, L_2, \dots, L_n)$  es una lista de eslabones //  
  Repetir  
    Si gesto_rotar_izq Entonces  
       $L_i$ .Rotar(Vector3.Up * Tiempo)  
    Sino Si gesto_rotar_der Entonces  
       $L_i$ .Rotar(Vector3.Down * Tiempo)  
    Fin Si  
  Hasta Que FinalizarApp()  
Fin Procedimiento
```

Figura 7

Traslación y rotación



5. RESULTADOS

Se realizaron pruebas en 20 usuarios para evaluar la usabilidad de la aplicación. Las pruebas están compuestas por dos partes. En la primera parte, el usuario realizó todas las operaciones con los gestos de la mano previamente enseñadas. Esta parte tuvo como propósito familiarizar al usuario con este tipo de técnica de interacción, ya que era la primera vez en todos en probar una aplicación con una interacción multimodal. En la segunda parte, el usuario realizó la fabricación del brazo robótico usando los gestos de la mano, ya que previamente se entrenó en movimientos básicos. Esta fabricación consistía en unir tres eslabones, la primera unión con la base y el eslabón 1; y la segunda unión con el eslabón 1 y el eslabón 2. Si la primera unión se realizaba correctamente, entonces en el eslabón 1 aparecían dos servomotores que daban lugar a las articulaciones del brazo robótico; y si la segunda unión se realizaba correctamente, entonces en el eslabón 2 aparecían las pinzas del brazo robótico. Ambas partes se dieron en un tiempo máximo de 15 minutos cada uno, para no exigir demasiado al *hardware* utilizado, ya que consume muchos recursos, principalmente de la CPU.

Para medir la usabilidad de la aplicación se escogió un conjunto de 20 personas para descubrir todos los problemas de usabilidad (Nielsen & Landauer, 1993), sin tomar como limitación el sexo y la edad. Además, se dividió al azar en dos grupos a los usuarios para evaluar la usabilidad de la aplicación. En el primer grupo, los usuarios no usaron el visor,

teniendo que valerse de una mano para sujetarlo; en el segundo grupo, utilizaron un visor como sujetador del celular, teniendo libre las dos manos. Para la experimentación se tomaron métricas cualitativas, como la comodidad en el uso de los gestos de la mano, el aprendizaje del uso de los gestos de la mano en el transcurso de la experimentación (asimilación de la relación gesto-operación), selección y manipulación del objeto 3D; y métricas cuantitativas, como el tiempo que se toma el usuario en la experimentación y los números de intentos para realizar las operaciones con los gestos de la mano (Samini & Ludin, 2017).

5.1 Grupo de usuarios sin visor

El primer usuario tenía educación superior completa y una capacidad tecnológica baja. Por otra parte, manifestó que no se sentía cómodo al manipular el celular con una mano, ya que le era difícil controlar el objeto 3D y los gestos de la mano, a la vez. Además, mencionó que era cuestión de acostumbrarse con la disposición del celular y los gestos de la mano al ser su primera experiencia con una interfaz como la que ha sido diseñada. Con respecto al aprendizaje de las interacciones, destacó que con algunas pruebas más podría mejorar y realizar un mejor desempeño. También añadió que el cambio de color al seleccionar un objeto le facilitó su movimiento y que la interacción más sencilla que realizó fue la traslación por el plano XZ. Por último, se desplazó alrededor del ambiente de prueba para poder visualizar mejor los objetos virtuales, ya que no percibía la profundidad en la cámara. Al finalizar la prueba, no pudo realizar ninguna unión entre los eslabones. La tabla 1 resume las métricas cualitativas aplicadas al usuario 1.

Tabla 1

Evaluación cualitativa del usuario 1 del grupo 1

Comodidad en el uso de los gestos de la mano	Baja
Aprendizaje en el transcurso de la experimentación	Ni baja ni alta
Selección y manipulación del objeto 3D	Ni baja ni alta

El segundo usuario tenía educación superior completa y una capacidad tecnológica alta. Por otra parte, manifestó que se sentía medianamente cómodo, ya que tenía que estar atento en mover el celular, para enfocar un objeto 3D, y la mano sin que los demás elementos interfirieran. Por ello, decidió levantarse y dispersar los componentes para poder visualizar y controlar mejor cada elemento. Asimismo, mencionó que al aislar los objetos pudo aprender con más facilidad a interactuar con cada uno de ellos y logró

controlarlos con una mayor precisión. Además, añadió que le fue sencillo seleccionar y manipular los componentes. Por último, se desplazó alrededor del ambiente de prueba para poder manipular mejor los objetos virtuales, ya que acercaba el celular a cada objeto y las operaciones se realizaron satisfactoriamente. Al finalizar la prueba, pudo realizar la primera unión, pero no la segunda, precisó que requería una mayor precisión. La tabla 2 resume las métricas cualitativas aplicadas al usuario 2.

Tabla 2

Evaluación cualitativa del usuario 2 del grupo 1

Comodidad en el uso de los gestos de la mano	Ni baja ni alta
Aprendizaje en el transcurso de la experimentación	Ni baja ni alta
Selección y manipulación del objeto 3D	Alta

El tercer usuario tenía educación superior completa y una capacidad tecnológica media. Por otra parte, manifestó que algunos gestos eran sencillos de hacer, ya que eran similares al usar el celular. Sin embargo, otros gestos, como el deseleccionar, no son comunes y se necesita interactuar más tiempo con la aplicación para poder aprenderlo. Asimismo, recalcó que la posición de la cámara, la luz y los factores externos generaban más de un plano, por lo que al manipular los objetos 3D saltaban de un plano a otro, generando dificultad. Por último, mencionó que podía efectuar las operaciones mejor en la segunda parte que en la primera parte, debido a que recordaba con mayor fluidez los gestos de las manos generando menos confusión en la operación que quisiera realizar. Al finalizar la prueba, pudo realizar la primera unión y la segunda unión satisfactoriamente.

Tabla 3

Evaluación cualitativa del usuario 3 del grupo 1

Comodidad en el uso de los gestos de la mano	Ni baja ni alta
Aprendizaje en el transcurso de la experimentación	Alta
Selección y manipulación del objeto 3D	Ni baja ni alta

El cuarto usuario tenía educación superior completa y una capacidad tecnológica alta. Por otra parte, manifestó que no se sentía tan cómodo al usar los gestos de la mano, ya que para él no era frecuente realizar dicha acción en aplicaciones o sistemas.

Sin embargo, añadió que fue sencillo familiarizarse con los gestos de la mano, ya que algunos eran intuitivos, y con los que no, eran fáciles de recordar. Por último, mencionó que, al acercar la cámara a cada uno de los objetos, era más sencillo de seleccionar y poder moverlos en donde él quería que se posicionaran. Al finalizar la prueba, pudo realizar la primera unión y la segunda unión satisfactoriamente.

Tabla 4

Evaluación cualitativa del usuario 4 del grupo 1

Comodidad en el uso de los gestos de la mano	Ni baja ni alta
Aprendizaje en el transcurso de la experimentación	Ni baja ni alta
Selección y manipulación del objeto 3D	Alta

El quinto usuario tenía educación superior completa y una capacidad tecnológica baja. Por otra parte, manifestó que era incómodo hacer algunos gestos de las manos que requerían el dedo gordo, ya que se le dificultaba enderezar o retraer el dedo de tal manera que el gesto se efectúe correctamente. Además, comentó que le costaba seleccionar el objeto que quería, por lo que trasladar o rotar se ejecutaba en los objetos que no quería. Por último, recalcó que se consideraba un principiante en este tipo de aplicaciones con interacciones no convencionales. Al finalizar la prueba, pudo realizar la primera unión y la segunda unión satisfactoriamente.

Tabla 5

Evaluación cualitativa del usuario 5 del grupo 1

Comodidad en el uso de los gestos de la mano	Baja
Aprendizaje en el transcurso de la experimentación	Ni baja ni alta
Selección y manipulación del objeto 3D	Baja

El sexto usuario tenía educación superior completa y una capacidad tecnológica baja. Por otra parte, manifestó que algunos gestos los efectuaba fácilmente al ser intuitivos, pero en otros, como la rotación, se le complicaba acomodar los dedos. Asimismo, mencionó que le resultaba complicado relacionar el patrón de los gestos de la mano con la operación que quería realizar. Por último, dio a entender que el espacio de prueba era importante para la experimentación, debido a que un espacio lleno de objetos generaba

muchos planos en donde colocar los objetos 3D y dificultaba el desplazamiento en el lugar para visualizar mejor los objetos. Al finalizar la prueba, no pudo realizar ni la primera ni la segunda unión.

Tabla 6

Evaluación cualitativa del usuario 6 del grupo 1

Comodidad en el uso de los gestos de la mano	Ni baja ni alta
Aprendizaje en el transcurso de la experimentación	Baja
Selección y manipulación del objeto 3D	Muy baja

El séptimo usuario tenía educación superior completa y una capacidad tecnológica media. Por otra parte, manifestó que el uso de los gestos de la mano le fue cómodo, pero tener que sujetar el celular cansaba la muñeca y tenía que intercambiar, cada cierto tiempo, la mano derecha con la mano izquierda y viceversa. Además, comentó que los gestos de las manos se debían realizar con precisión para poder ejecutar correctamente la operación deseada y que al desplazarse por el espacio de prueba se le facilitaba la visualización de la profundidad del plano. Por tanto, no tuvo dificultades para seleccionar y manipular los objetos 3D. Al finalizar la prueba, pudo realizar la primera unión y la segunda unión satisfactoriamente.

Tabla 7

Evaluación cualitativa del usuario 7 del grupo 1

Comodidad en el uso de los gestos de la mano	Alta
Aprendizaje en el transcurso de la experimentación	Alta
Selección y manipulación del objeto 3D	Ni baja ni alta

El octavo usuario tenía educación superior completa y una capacidad tecnológica media. Por otra parte, manifestó que el uso de los gestos de la mano fue regularmente cómodo. Además, añadió que no fue tan fácil recordar qué gesto de la mano correspondía a la operación que quería realizar y que el estar sujetando el celular le causaba cansancio en la mano. Por último, no le fue tan fácil ni tan difícil el seleccionar y manipular los objetos 3D. Al finalizar la prueba, pudo realizar la primera unión y la segunda unión satisfactoriamente.

Tabla 8

Evaluación cualitativa del usuario 8 del grupo 1

Comodidad en el uso de los gestos de la mano	Ni baja ni alta
Aprendizaje en el transcurso de la experimentación	Ni baja ni alta
Selección y manipulación del objeto 3D	Ni baja ni alta

El noveno usuario tenía educación superior completa y una capacidad tecnológica media. Por otra parte, manifestó que el uso de los gestos de la mano fue regularmente cómodo. Asimismo, mencionó que fue fácil recordar los gestos de la mano con respecto a la operación relacionada, pero fue complicado efectuarlo. Esto fue debido a que en el espacio de prueba se generaron muchos planos y fue un obstáculo para ubicar el objeto 3D en el plano en donde se quería ensamblar el brazo robótico. Por último, recalcó que, aunque tenía aquel problema, no fue tan fácil ni tan difícil el seleccionar y manipular el objeto 3D. Al finalizar la prueba, pudo realizar la primera unión satisfactoriamente, pero no la segunda unión.

Tabla 9

Evaluación cualitativa del usuario 9 del grupo 1

Comodidad en el uso de los gestos de la mano	Ni baja ni alta
Aprendizaje en el transcurso de la experimentación	Alta
Selección y manipulación del objeto 3D	Ni baja ni alta

El décimo usuario tenía educación superior completa y una capacidad tecnológica media. Por otra parte, manifestó que algunos gestos de las manos eran fáciles de usar, pero otros no. Debido a que algunos necesitaban realizarse con mayor precisión, es decir, llegar a exagerar un poco el enderezar o retraer los dedos lo necesario para realizar el gesto. Aunque tenía aquel problema, no fue tan fácil ni tan difícil el seleccionar y manipular los objetos 3D. Por último, mencionó que fue sencillo recordar los gestos de la mano según la operación que deseaba realizar. Al finalizar la prueba, pudo realizar la primera unión satisfactoriamente, pero no la segunda unión.

Tabla 10

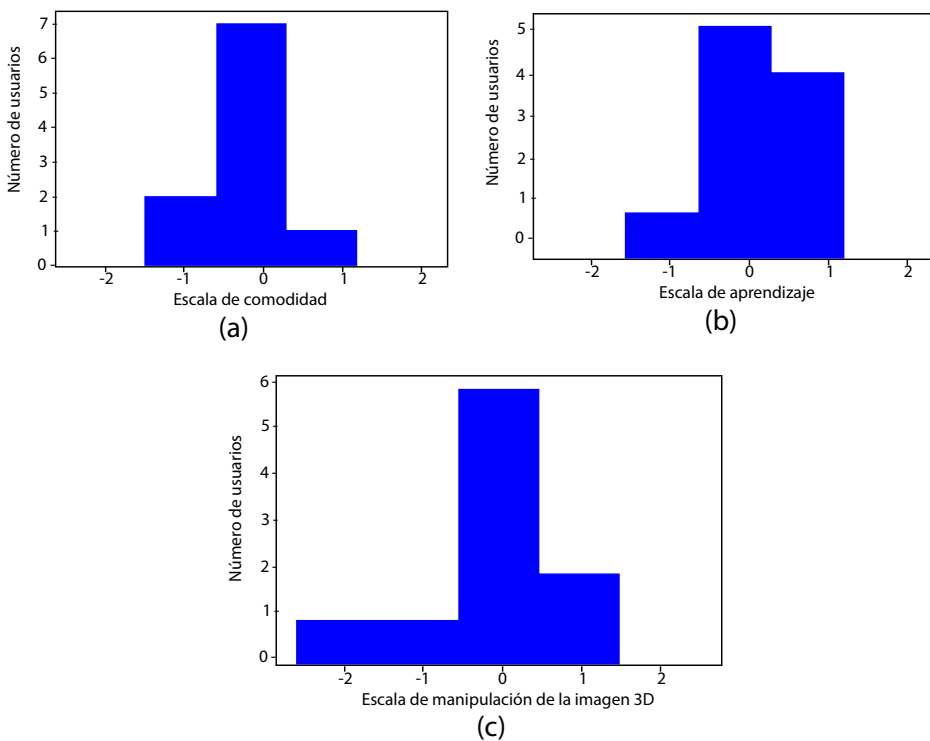
Evaluación cualitativa del usuario 10 del grupo 1

Comodidad en el uso de los gestos de la mano	Ni baja ni alta
Aprendizaje en el transcurso de la experimentación	Alta
Selección y manipulación del objeto 3D	Ni baja ni alta

Se elaboraron histogramas para las métricas cualitativas sobre la comodidad del uso de los gestos de la mano (véase figura 8A), el aprendizaje en el transcurso de la experimentación (véase figura 8B) y la selección y manipulación del objeto 3D (véase figura 8C) para el grupo 1. Las métricas son representadas a manera de escala (del -2 al 2) en lugar de los valores usados en las tablas anteriores con el fin de visualizar mejor los resultados obtenidos.

Figura 8

Histogramas de las métricas cualitativas del grupo 1



5.2 Grupo de usuarios con visor

El primer usuario tenía educación superior completa y una capacidad tecnológica alta. Por otra parte, se sentía cómodo con el visor, aunque tenía poca experiencia en el uso de este tipo de aplicaciones, al enfocar el objeto 3D y la mano para realizar los gestos. Además, separó los componentes del brazo robótico para aprender los gestos de translación y rotación con cada objeto 3D. Por último, manifestó que fue sencillo la primera unión del robot, pero que para la segunda se requería una mayor precisión en trasladar y rotar el objeto 3D para que se ensamblaran las piezas. Al finalizar la prueba, pudo realizar la primera unión satisfactoriamente, pero no la segunda unión.

Tabla 11

Evaluación cualitativa del usuario 1 del grupo 2

Comodidad en el uso de los gestos de la mano	Alta
Aprendizaje en el transcurso de la experimentación	Alta
Selección y manipulación del objeto 3D	Alta

El segundo usuario tenía educación superior completa y una capacidad tecnológica media. Por otra parte, afirmó que se sentía medianamente cómodo al usar por primera vez un visor y hacer gestos de la mano para realizar ciertas funciones con objetos en realidad aumentada. Por ello, decidió realizar la experimentación de pie para tener una mayor visión de la mesa (espacio de prueba) y mayor movilidad para trasladar los elementos en distintos puntos del plano horizontal. Además, mencionó que los gestos de seleccionar y trasladar eran interacciones sencillas de entender, pero necesitó más tiempo para aprender los gestos de deseleccionar y rotar. Por último, manifestó que le costó realizar la primera unión del robot y para la segunda unión le resultó mucho más difícil porque requería una mayor precisión. Al finalizar la prueba, pudo realizar la primera unión satisfactoriamente, pero no la segunda unión.

Tabla 12

Evaluación cualitativa del usuario 2 del grupo 2

Comodidad en el uso de los gestos de la mano	Ni baja ni alta
Aprendizaje en el transcurso de la experimentación	Ni baja ni alta
Selección y manipulación del objeto 3D	Alta

El tercer usuario tenía educación superior completa y una capacidad tecnológica media. Por otra parte, manifestó que en una primera instancia le fue incómodo, pero en el transcurso de la prueba se fue acostumbrando y realizaba los gestos de la mano con naturalidad. Además, mencionó que necesitaba más práctica para aprender todos los gestos de la mano, ya que se olvidaba y hacía otras operaciones que no quería. Por último, añadió que no fue ni tan fácil ni tan difícil el seleccionar y manipular los objetos en realidad aumentada. Al finalizar la prueba, pudo realizar la primera unión y la segunda unión satisfactoriamente.

Tabla 13

Evaluación cualitativa del usuario 3 del grupo 2

Comodidad en el uso de los gestos de la mano	Alta
Aprendizaje en el transcurso de la experimentación	Ni baja ni alta
Selección y manipulación del objeto 3D	Ni baja ni alta

El cuarto usuario tenía educación superior completa y una capacidad tecnológica media. Por otra parte, manifestó que se sentía muy cómodo con los gestos de la mano, pero que se podían mejorar para que fuesen más sencillos e intuitivos. También recalcó que fue bueno su aprendizaje en el transcurso de la prueba. Por último, mencionó que el uso del visor le facilitaba el seleccionar y manipular los objetos en realidad aumentada, ya que podía enfocar los objetos como si estuviera mirándolos. Al finalizar la prueba, pudo realizar la primera unión y la segunda unión satisfactoriamente.

Tabla 14

Evaluación cualitativa del usuario 4 del grupo 2

Comodidad en el uso de los gestos de la mano	Ni baja ni alta
Aprendizaje en el transcurso de la experimentación	Ni baja ni alta
Selección y manipulación del objeto 3D	Alta

El quinto usuario tenía educación secundaria completa y una capacidad tecnológica alta. Por otra parte, manifestó que solo el gesto de la mano para girar hacia la derecha un objeto 3D le resultaba incómodo, pero los demás fueron regularmente cómodos. Asimismo, mencionó que no fue tan fácil ni tan difícil el seleccionar y manipular los objetos

3D, pero a veces trasladaba el objeto 3D cuando quería rotarlo. Por último, recalcó que fue sencillo recordar todos los gestos de la mano. Al finalizar la prueba, pudo realizar la primera unión y la segunda unión satisfactoriamente.

Tabla 15

Evaluación cualitativa del usuario 5 del grupo 2

Comodidad en el uso de los gestos de la mano	Ni baja ni alta
Aprendizaje en el transcurso de la experimentación	Alta
Selección y manipulación del objeto 3D	Ni baja ni alta

El sexto usuario tenía educación superior y una capacidad tecnológica alta. Por otra parte, manifestó que los gestos de selección y traslación fueron cómodos, pero el de rotación a la derecha fue más difícil. Sin embargo, recalcó que gracias al uso del visor podía cambiar de mano fácilmente y superar el obstáculo de rotar hacia la derecha. Por último, mencionó que fue sencillo y rápido de recordar, especialmente los gestos de selección y traslación le parecieron intuitivos. Al finalizar la prueba, pudo realizar la primera unión y la segunda unión satisfactoriamente.

Tabla 16

Evaluación cualitativa del usuario 6 del grupo 2

Comodidad en el uso de los gestos de la mano	Ni baja ni alta
Aprendizaje en el transcurso de la experimentación	Muy alta
Selección y manipulación del objeto 3D	Alta

El séptimo usuario tenía educación superior completa y una capacidad tecnológica media. Por otra parte, manifestó que se sintió cómodo con el uso de los gestos de la mano y el uso del visor. Además, añadió que al inicio le costó recordar y aprender los gestos de la mano. Sin embargo, pudo realizar todas las pruebas sin confundirse en qué operación utilizar. Por último, mencionó que con la práctica realizada en la experimentación sí fue fácil poder manipular los objetos en realidad aumentada con excepción del gesto de traslación. Al finalizar la prueba, pudo realizar la primera unión y la segunda unión satisfactoriamente.

Tabla 17

Evaluación cualitativa del usuario 7 del grupo 2

Comodidad en el uso de los gestos de la mano	Alta
Aprendizaje en el transcurso de la experimentación	Muy alta
Selección y manipulación del objeto 3D	Alta

El octavo usuario tenía educación superior y una capacidad tecnológica media. Por otra parte, manifestó que estuvo cómodo con el uso de los gestos de la mano excepto el gesto para trasladar, recalcó que fue el más dificultoso. También mencionó que comprendió los gestos rápidamente y con el transcurso de la experimentación pudo reforzar el aprendizaje del uso de los gestos. Por último, comentó que fue sencillo el seleccionar y manipular los objetos 3D gracias al visor, en cambio, si tuviese el celular en la mano, no sentiría tan simple el realizar dichas operaciones. Al finalizar la prueba, pudo hacer la primera unión y la segunda unión satisfactoriamente.

Tabla 18

Evaluación cualitativa del usuario 8 del grupo 2

Comodidad en el uso de los gestos de la mano	Ni baja ni alta
Aprendizaje en el transcurso de la experimentación	Alta
Selección y manipulación del objeto 3D	Muy alta

El noveno usuario tenía educación superior completa y una capacidad tecnológica media. Por otra parte, manifestó que tuvo complicaciones con el uso de los gestos de la mano, ya que la aplicación no reconocía bien los gestos de la mano por la sombra que se generaba en el espacio de prueba. Por lo tanto, al inicio se le complicaba el seleccionar y manipular los objetos en realidad aumentada, pero conforme se iba adaptando a la situación se le hacían más sencillas dichas operaciones. Por último, sintetizó cuanto había comprendido y aprendido los gestos de la mano con una nota de quince sobre veinte. Al finalizar la prueba, pudo realizar la primera unión satisfactoriamente, pero no la segunda unión.

Tabla 19

Evaluación cualitativa del usuario 9 del grupo 2

Comodidad en el uso de los gestos de la mano	Ni baja ni alta
Aprendizaje en el transcurso de la experimentación	Ni baja ni alta
Selección y manipulación del objeto 3D	Alta

El décimo usuario tenía educación superior y una capacidad tecnológica alta. Por otra parte, manifestó que al comienzo de la prueba le fue algo complicado, pero poco a poco pudo usar todos los gestos de la mano con más naturalidad. Además, añadió que el gesto más complicado fue el de deseleccionar el objeto y que los gestos fueron fáciles de entender, pero fue un poco difícil aplicarlos en la aplicación. Por último, mencionó que fue fácil seleccionar y mover los objetos 3D, pero más difícil fue deseleccionarlos. Al finalizar la prueba, pudo realizar la primera unión y la segunda unión satisfactoriamente.

Tabla 20

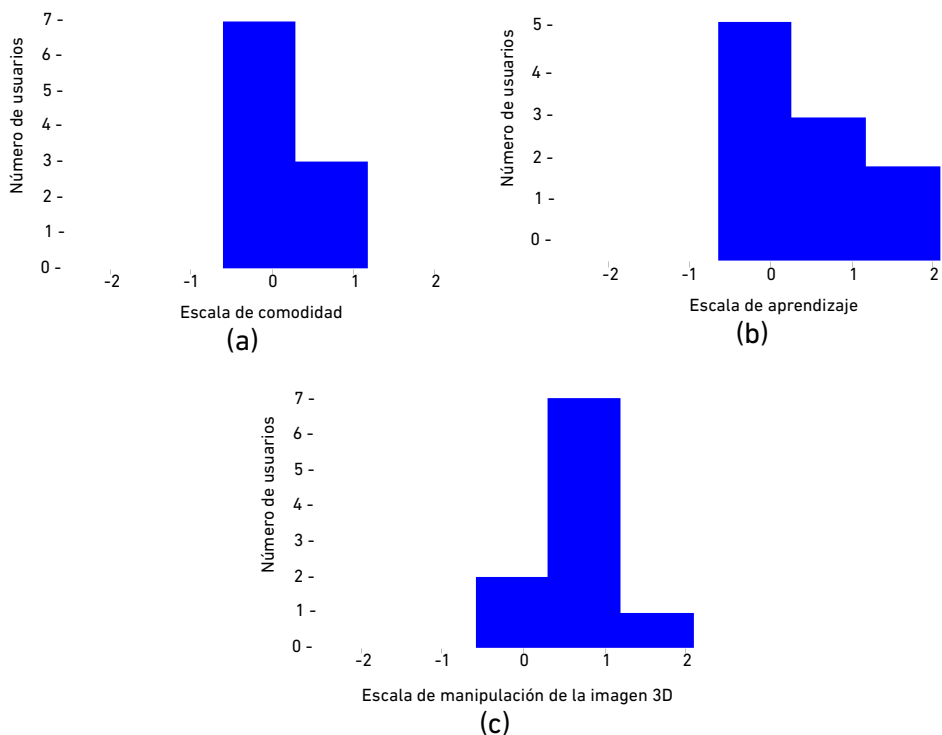
Evaluación cualitativa del usuario 10 del grupo 2

Comodidad en el uso de los gestos de la mano	Ni baja ni alta
Aprendizaje en el transcurso de la experimentación	Ni baja ni alta
Selección y manipulación del objeto 3D	Alta

Se elaboraron histogramas para las métricas cualitativas sobre la comodidad del uso de los gestos de la mano (véase figura 9A), el aprendizaje en el transcurso de la experimentación (véase figura 9B) y la selección y manipulación del objeto 3D (véase figura 9C) para el grupo 2. Las métricas son representadas a manera de escala (del -2 al 2) en lugar de los valores usados en las tablas anteriores con el fin de visualizar mejor los resultados obtenidos.

Figura 9

Histogramas de las métricas cualitativas del grupo 2



Durante las diversas pruebas que se realizaron a cada usuario, se pueden extraer ciertas limitaciones del proyecto. Por un lado, se encuentra la limitación con respecto al espacio de la experimentación. Dicho de otra manera, estar en un lugar en donde alrededor hay diversos objetos, la aplicación genera diversos planos lo que perjudica la precisión en donde se quiere colocar la imagen 3D del eslabón y entorpece la visualización del usuario. Por el otro lado, se encuentra la limitación con respecto al rastreo de la mano que se realiza en la aplicación. En otras palabras, en ciertos momentos de la experimentación, el reconocimiento de la mano arrojaba falsos positivos (reconocer una mano en donde no había) por factores externos como la iluminación, la sombra, el daño en la cámara, entre otros.

Las métricas cuantitativas realizadas en ambos grupos según las tablas 21 y 22 muestran que inicialmente el tiempo elegido para la experimentación era de diez minutos, pero los resultados mostraron que no era tiempo suficiente para que el usuario pudiera aprender sobre los gestos de la mano y cómo utilizarlos para la fabricación del brazo robótico. Por lo que el usuario del grupo sin visor con métricas más bajas es el

usuario 1 y por parte del grupo con visor, el usuario 2. Por eso, se decidió aumentar el tiempo a quince minutos por cada parte de la prueba. Dando como resultado al usuario del grupo sin visor con métricas más altas al usuario 4, ya que redujo los números de intentos de traslación y rotación considerablemente y obtuvo el mejor tiempo, de 2 minutos y 35 segundos, para poder fabricar el robot, lo cual refleja su alta manipulación con los objetos en realidad aumentada. Por el lado del grupo con visor, el usuario con métricas más altas fue el usuario 6, debido a que disminuyó los números de intentos de traslación y rotación significativamente y alcanzó el mejor tiempo, de 4 minutos, para poder fabricar el robot, lo cual evidencia su alta manipulación con los objetos en realidad aumentada.

Tabla 21

Evaluación cuantitativa del grupo sin visor

		Grupo sin visor							
U s u a r i o	E d a d a	Primera parte			Segunda parte				
		Tiempo (minutos)	N.º de intentos para trasladar	N.º de intentos para rotar	Tiempo (minutos)	N.º de intentos para trasladar	N.º de intentos para rotar	N.º de intentos para la primera unión	N.º de intentos para la segunda unión
1	67	10	2	3	10	2	3	-	-
2	30	10	1	2	10	1	2	5	-
3	24	15	1	2	15	2	2	2	10
4	26	15	5	4	2:35	2	1	4	3
5	26	15	5	5	4:40	4	5	7	10
6	26	15	5	3	15	4	2	8	-
7	26	15	10	3	6:54	6	2	3	8
8	26	15	5	3	14:50	3	2	5	10
9	26	15	5	3	15	5	3	10	2
10	26	15	7	4	15	5	3	10	-

Tabla 22

Evaluación cuantitativa del grupo con visor

		Grupo con visor							
U s u a r i o	E d a d	Primera parte			Segunda parte				
		Tiempo (minutos)	N.º de intentos para trasladar	N.º de intentos para rotar	Tiempo (minutos)	N.º de intentos para trasladar	N.º de intentos para rotar	N.º de intentos para la primera unión	N.º de intentos para la segunda unión
1	26	10	1	1	10	1	4	4	10
2	28	10	1	3	10	1	3	5	12
3	31	15	3	2	6:04	5	2	2	8
4	37	15	4	2	12:52	4	2	7	5
5	24	15	3	6	6:06	2	2	2	5
6	24	15	2	6	4	1	3	1	6
7	29	15	3	1	8	1	2	2	4
8	23	15	5	1	15	5	1	6	3
9	33	15	7	2	15	4	1	3	5
10	25	15	5	3	10	5	3	5	10

6. DISCUSIONES

De la experimentación, se pudo observar que la comodidad en el uso de los gestos de la mano, en la mayoría de los usuarios, no fue ni tan cómoda ni tan incómoda, es decir, fue intermedia. Esto se debe a que la mayoría de usuarios era inexperta en el uso de una aplicación en realidad aumentada y con una interfaz con interacciones a través de los gestos de la mano. También, se vio reflejado en el número de intentos para realizar una operación, como, por ejemplo, el número de intentos para realizar una rotación que, en general, fue entre dos a cinco. Sin embargo, se observó que la inexperiencia de los usuarios no fue impedimento para entender y comprender sobre estos gestos. Los valores del aprendizaje en el transcurso de la experimentación oscilaban entre intermedia a alta, en ambos grupos de usuarios, debido a que los gestos de la mano eran intuitivos según mencionaron algunos de ellos.

Lo que sí se pudo visualizar de ambos grupos es que difieren en la selección y manipulación de objetos en realidad aumentada. En el desarrollo de las pruebas, los usuarios sin el uso del visor tuvieron, en el mayor de los casos, un desempeño intermedio; en cambio, los usuarios con el uso del visor tuvieron un desempeño alto. Asimismo, se observó en el número de resultados satisfactorios en la primera unión y en la segunda unión en la fabricación del brazo robótico.

El grupo de usuarios que utilizó el visor obtuvo un mayor número de ensambles o uniones entre los eslabones del brazo robótico con respecto al grupo de usuarios sin el uso de visor. Además, entre los usuarios que pudieron realizar la primera unión y la segunda unión resaltaron que realizar la segunda unión fue más complicado. Esto se puede verificar en el alto número de intentos para realizar dicha unión con respecto a la otra en la mayoría de los usuarios.

7. CONCLUSIONES Y TRABAJOS FUTUROS

En conclusión, se puede asegurar que la aplicación que ayuda a efectuar un diseño previo a la fabricación de un brazo robótico de marea interactiva usando realidad aumentada fue un reto para todos los usuarios. Esto se debe a que, para la mayoría de los usuarios, no es usual utilizar aplicaciones con realidad aumentada o con interacciones no convencionales; es decir, el uso de gestos de la mano. También se visualiza que en métrica cualitativa de la comodidad en el uso de los gestos de la mano, los resultados tienden a ser ni bajos ni altos. Asimismo, se puede inferir mediante los resultados que los usuarios pudieron entender y comprender los gestos de la mano de manera favorable. Esto se refleja en la métrica cualitativa del aprendizaje en el transcurso de la experimentación que tiende a ser alta, ya que los gestos de la mano, en la mayoría, eran intuitivos y sencillos de memorizar con su respectiva operación en la aplicación. Además, de las pruebas en el procedimiento de ensamblaje se puede deducir que el grupo de usuarios que utilizó el visor obtuvo un valor alto en el seleccionar y manipular los objetos en realidad aumentada. Por lo tanto, los usuarios de dicho grupo pudieron realizar más ensamblajes satisfactoriamente con respecto al otro grupo de usuarios.

Finalmente, existen diversas aplicaciones que se podrían llevar a cabo con este proyecto. Como se observa, es posible asociar el rastreo de mano y la realidad aumentada con objetos modelados en 3D, por lo que da lugar a nuevos proyectos con diferentes rastreos mediante la herramienta de MediaPipe y la realidad aumentada (Díaz Álvarez, 2021; Cáceres Sánchez, 2019; Delgado Barrera & Duchi Farez, 2021). Asimismo, en un futuro proyecto se podría aumentar la cantidad de modelos de brazos robóticos a fabricar y decidir cuál construir, dando mayor diversidad en interacciones con los gestos de la mano o manos, dependiendo de la dificultad. Es decir, se podría realizar una operación usando ambas manos o realizar una operación por cada mano. Además, otro proyecto futuro sería elaborar un simulador sobre las diferentes trayectorias que puede realizar el brazo robótico siendo guiado por los gestos de la mano o manos. Esto se podría realizar al controlar los grados de libertad y las articulaciones del brazo robótico mediante los gestos de la mano o manos, añadiendo también una operación en el elemento terminal como pinzas o herramientas.

REFERENCIAS

- Azuma, R. T. (1997). A survey of augmented reality. *Presence: Teleoperators & Virtual Environments*, 6(4), 355-385. <https://doi.org/10.1162/pres.1997.6.4.355>
- Cáceres Sánchez, J. S. (2019). *Diseño de prototipo de aplicación para visualización de productos en realidad aumentada* [Tesis de bachiller, Universidad San Francisco de Quito, Colegio de Ciencias e Ingenierías]. <http://repositorio.usfq.edu.ec/handle/23000/8893>
- Chacón Malasquez, N., & Tuiro Achulle, J. M. (2018). *Aplicación móvil de realidad aumentada, utilizando la metodología Mobile - D, para el entrenamiento de técnicos de mantenimiento de maquinaria pesada en la empresa Zamine Service Perú S.A.C.* [Tesis de grado, Universidad Autónoma del Perú]. <https://hdl.handle.net/20.500.13067/581>
- de la Cruz, E. I., Salazar, E. R., Romero, J. A., Jiménez, L. M., & Rodríguez, J. J. (2022). Control to manipulate robotic arms using augmented reality. En X.-S. Yang, S. Sherratt, N. Dey & A. Joshi (Eds.), *Proceedings of Sixth International Congress on Information and Communication Technology. Lecture Notes in Networks and Systems* (Vol. 216, pp. 101-112). Springer. https://doi.org/10.1007/978-981-16-1781-2_11
- Delgado Barrera, J. D., & Duchi Farez, D. D. (2021). *Desarrollo de una interfaz humano-computador mediante la animación de avatares generadores a partir de fotogrametría* [Tesis de grado, Universidad Politécnica Salesiana Sede Cuenca]. <https://dspace.ups.edu.ec/handle/123456789/21214>
- Díaz Álvarez, J. (2021). *Visión por computador para el uso de realidad aumentada en Unity3D* [Tesis de grado, Universidad Politécnica de Madrid]. <https://oa.upm.es/68007/>
- Fang, H. C., Ong, S. K., & Nee, A. Y. C. (2014). A novel augmented reality-based interface for robot path planning. *International Journal on Interactive Design and Manufacturing (IJIDeM)*, 8, 33-42. <https://doi.org/10.1007/s12008-013-0191-2>
- Lugaresi, C., Tang, J., Nash, H., McClanahan, C., Uboweja, E., Hays, M., Zhang, F., Chang, C.-L., Yong, M., Lee, J., Chang, W.-T., Hua, W., Georg, M., & Grundmann, M. (2019). MediaPipe: A framework for perceiving and processing reality. *Third Workshop on Computer Vision for AR/VR at IEEE Computer Vision and Pattern Recognition (CVPR) 2019*. https://mixedreality.cs.cornell.edu/s/NewTitle_May1_MediaPipe_CVPR_CV4ARVR_Workshop_2019.pdf
- MediaPipe (2020). MediaPipe Hands. *MediaPipe on GitHub*. <https://google.github.io/mediapipe/solutions/hands.html>
- Nielsen, J., & Landauer, T. K. (1993, 24-29 de abril). A mathematical model of the finding of usability problems. *Proceedings of the INTERACT '93 and CHI '93 Conference*

- on Human Factors in Computing Systems (CHI '93)*. Association for Computing Machinery, 206-213. <https://doi.org/10.1145/169059.169166>
- Pérez López, B. (2021). *Aplicación de realidad aumentada para la docencia online* [Tesis de grado, Universidad de Alicante]. https://rua.ua.es/dspace/bitstream/10045/115955/1/Sistema_de_reconocimiento_de_manos_y_realidad_virtual_Perez_Lopez_Borja.pdf
- Pasaréti, O., Hajdú, H., Matusaka, T., Jámbori, A., Molnár, I., & Turcsányi-Szabó, M. (2011). Augmented reality in education. *INFODIDACT 2011 Informatika Szakmódszertani Konferencia*. https://people.inf.elte.hu/szlavi/InfoDidact11/Manuscripts/PO_HH_MT_JA_MI_TSzM.pdf
- Samini, A., & Palmerius, K. L. (2017). Popular performance metrics for evaluation of interaction in virtual and augmented reality. *2017 International Conference on Cyberworlds (CW)*, 206-209. <https://doi.org/10.1109/CW.2017.25>
- Vogel, C., Walter, C., & Elkmann, N. (2017). Safeguarding and supporting future human-robot cooperative manufacturing processes by a Projection- and Camera-based Technology. *Procedia Manufacturing*, 11, 39-46. <https://doi.org/10.1016/j.promfg.2017.07.127>
- Ziden, A. A., Ziden, A. A. A., & Ifedayo, A. E. (2022). Effectiveness of augmented reality (AR) on students' achievement and motivation in learning science. *Eurasia Journal of Mathematics, Science and Technology Education*, 18(4), em2097. <https://doi.org/10.29333/ejmste/11923>

ANEXOS

Figura 10

Instanciar un objeto usando la aplicación



Figura 11

Selección y deselección de un objeto usando la aplicación

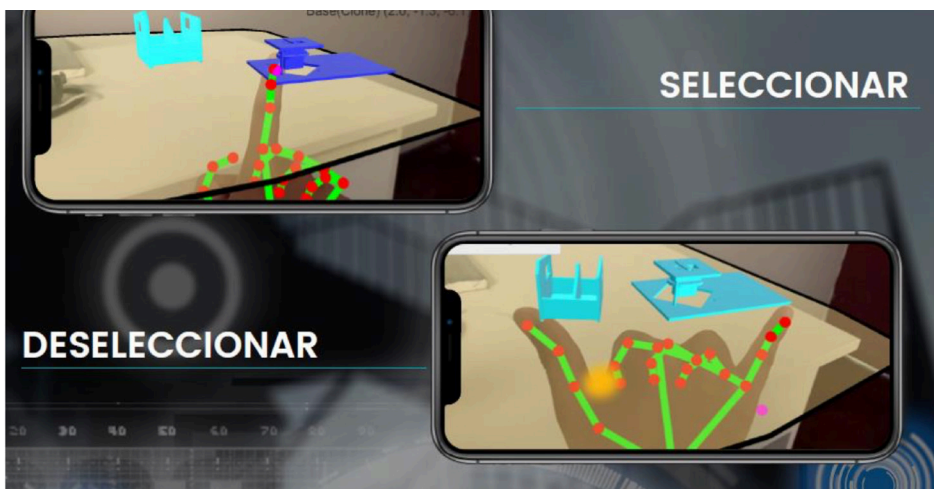


Figura 12

Traslación de un objeto durante la evaluación

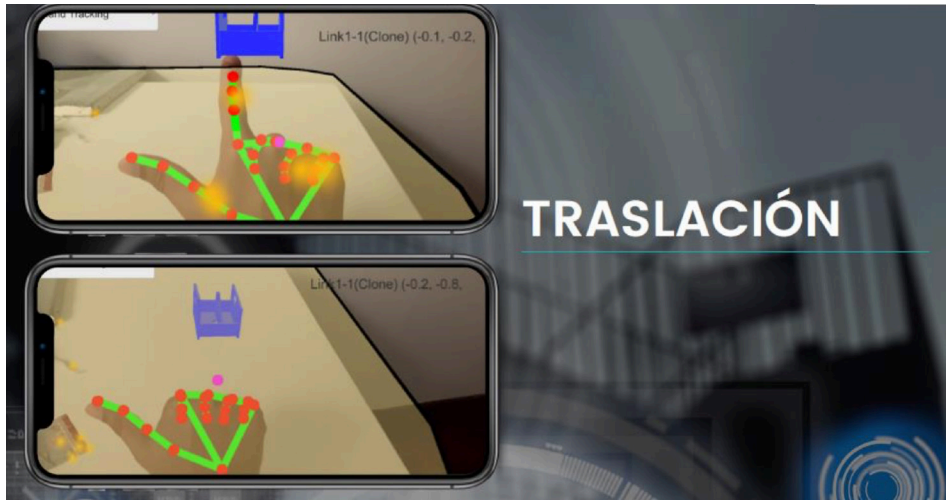


Figura 13

Rotación de objetos durante la evaluación

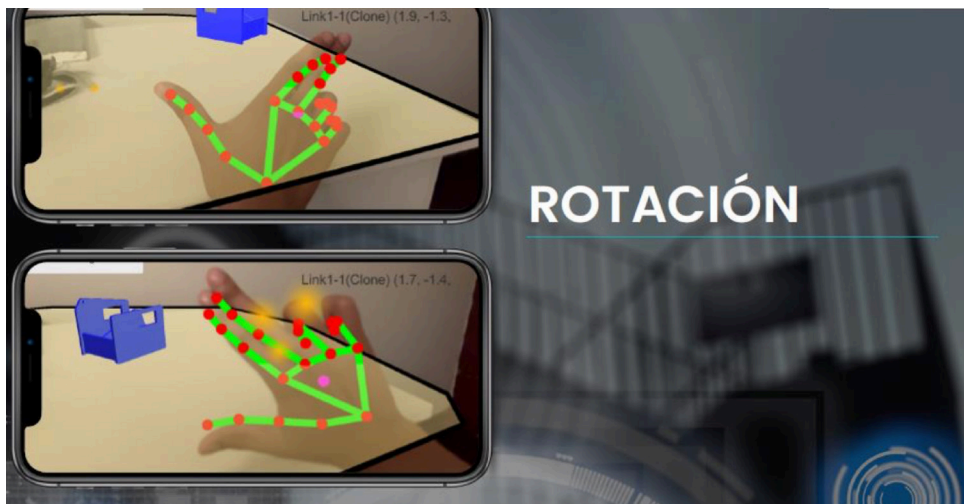


Figura 14

Interacción del usuario 5, del grupo con visor, con la aplicación de realidad aumentada rotando en sentido horario el eslabón del brazo robótico seleccionado (base del brazo robótico)

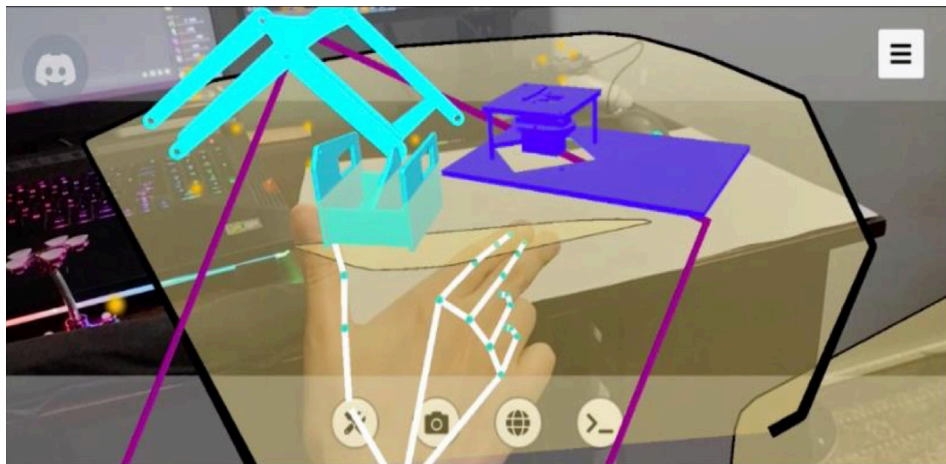


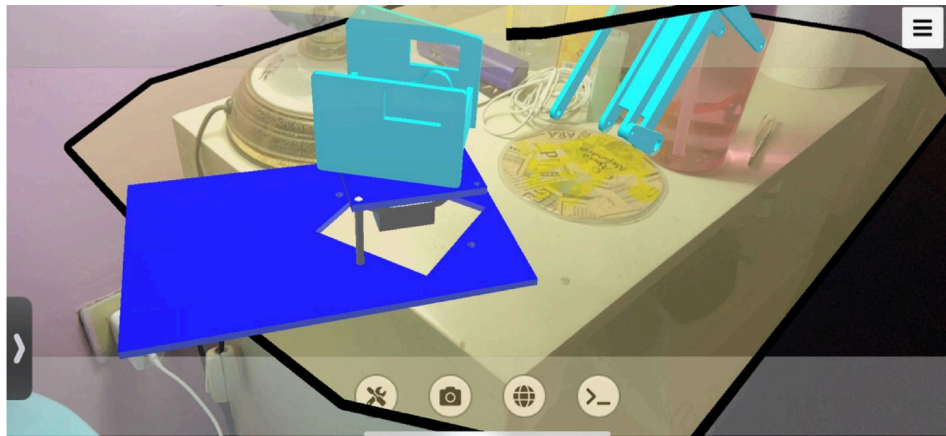
Figura 15

Interacción del usuario 6, del grupo con visor, con la aplicación de realidad aumentada trasladando hacia arriba el eslabón del brazo robótico seleccionado



Figura 16

Interacción del usuario 6, del grupo sin visor, con la aplicación de realidad aumentada intentando realizar la primera unión del brazo robótico



APLICACIÓN DE MÉTODOS DE *DEEP LEARNING* EN SERIES DE TIEMPO PARA EL PRONÓSTICO DE LA SITUACIÓN MACROECONÓMICA EN AMÉRICA LATINA

VÍCTOR AUGUSTO ALEGRE IBÁÑEZ
victoralegre@uni.pe / ORCID: 0000-0002-1456-8065
Universidad Nacional de Ingeniería (UNI), Lima, Perú

JOSE MARTIN LOZANO APARICIO
jlozano@uni.edu.pe / ORCID: 0000-0003-1086-5799
Universidad Nacional de Ingeniería (UNI), Lima, Perú

Resumen

Los métodos de *deep learning* pueden ser aplicados para generar modelos de pronóstico. Nosotros trabajamos con el producto bruto interno (PBI) de seis países de América Latina: Argentina, Brasil, Chile, Colombia, México y Perú empleando indicadores macroeconómicos anuales y trimestrales, del Banco Mundial y la Comisión Económica para América Latina y el Caribe (CEPAL), respectivamente. Para el preprocesamiento de los datos, a las series trimestrales se agregaron como características adicionales la descomposición de estas en tendencia, estacionalidad y residuo, con la finalidad de aportar más información a los modelos. Además, se reemplazaron datos atípicos producto del impacto de la pandemia del COVID-19 en la economía mundial. Se construyeron modelos de Perceptrón Multi Capa, Red Neuronal Convolucional, LSTM, GRU y SeqToSeq para cada país y frecuencia de sus series, y luego se evaluaron mediante validación cruzada continua y métricas MAE, RMSE y MAPE. Los modelos óptimos varían por cada caso

PALABRAS CLAVE: aprendizaje profundo / pronóstico de PBI / CEPAL / redes neuronales

APPLICATION OF DEEP LEARNING METHODS IN TIME SERIES FOR THE FORECAST OF THE MACROECONOMIC SITUATION IN LATIN AMERICA

Abstract

Deep learning methods can be applied to generate predictive models. We worked with the gross domestic product (GDP) of six Latin American countries: Argentina, Brazil, Chile, Colombia, Mexico, and Peru, using annual and quarterly macroeconomic indicators from the World Bank and the Economic Commission for Latin America and the Caribbean (ECLAC), respectively. For the pre-processing of the data, we decomposed the quarterly series into trend, seasonality, and residual and used them as additional characteristics to provide more information to the models. In addition, outliers resulting from the impact of the COVID-19 pandemic on the world economy were replaced. Multilayer perceptron, convolutional neural networks, LSTM, GRU, and SeqToSeq models were built for each country and their series' frequency, then evaluated by continuous cross-validation and MAE, RMSE, and MAPE metrics. The optimal models vary for each case.

KEYWORDS: deep learning / GDP forecasting / CEPAL / neural network

1. INTRODUCCIÓN

El *deep learning* (DL), parte del *machine learning* (ML), se caracteriza por sus arquitecturas de redes neuronales de varias capas ocultas a lo que se denomina profundidad (de ahí el nombre). Este tiene aplicaciones como el reconocimiento de patrones, el procesamiento de lenguaje natural y la predicción, de la cual trataremos en este trabajo.

La predicción puede ser muy útil para la toma de decisiones sobre eventos a futuro. Esto es respaldado por datos de sucesos anteriores como los datos históricos o series de tiempo. Ejemplos de esto son los pronósticos de precios de artículos, *commodities*, acciones, o los valores de indicadores económicos. Estos son importantes para una empresa o para un Gobierno, para vigilar y tomar medidas para mejorar su situación.

Los indicadores macroeconómicos nos describen la situación económica de un país o región, como también el bienestar de su población. El indicador que se pronosticará es el producto bruto interno (PBI) el cual mide el ingreso de todos los miembros de la economía, que es igual al gasto total en la producción de bienes y servicios finales. Esto aplica para un país en cierto periodo (Mankiw, 2014).

Para el pronóstico de series temporales, ya existen modelos estadísticos como las medias móviles y las autorregresiones. Pero en este trabajo nos enfocaremos en modelos de *deep learning* que también tienen buen desempeño con las series de tiempo, y los aplicaremos para predecir indicadores macroeconómicos de países de América Latina. Esta investigación comienza describiendo los conceptos principales en la sección 2. Los trabajos relacionados en la sección 3. La metodología usada es presentada en la sección 4. Los resultados y las discusiones son mostradas en la sección 5. Finalmente, presentamos nuestras conclusiones.

2. MARCO TEÓRICO

Las *series de tiempo* son datos secuenciales cuyos registros tienen un tiempo asociado (hora, fecha, año, entre otros). Por ejemplo, transacciones de compras. Una serie de tiempo es un tipo de dato secuencial en la que sus registros consisten en datos numéricos medidos sobre el tiempo (Tan et al., 2006). Estos pueden ser generados por procesos naturales y económicos como los mercados de acciones, observaciones científicas, médicas o fenómenos naturales (Han, 2012).

El *aprendizaje profundo* (*deep learning*) es un subgrupo del aprendizaje automático (*machine learning*) que pone énfasis en el aprendizaje de capas sucesivas de representaciones cada vez más significativas. El número de capas que contribuyen al modelo se llama profundidad, y estos modelos mayormente son redes neuronales (Chollet, 2018). Las redes neuronales de aprendizaje profundo usadas en este trabajo son:

- *Perceptrón multicapa (MLP)*, esta red neuronal más básica es usada para regresión lineal, en la cual los nodos de cada capa están conectados a los nodos de la siguiente capa (Skansi, 2018).
- *Redes neuronales convolucionales (CNN)* tienen una o más capas convolucionales que realizan la operación de convolución. Esta consiste en aplicar kernels o filtros por medio de producto escalar, para obtener mapas de características (Skansi, 2018).
- *Redes neuronales recurrentes (RNN)*, son una clase de redes neuronales donde las conexiones entre unidades forman un ciclo dirigido. Este crea un estado interno de la red, lo cual permite exponer un comportamiento temporal dinámico. Las RNN aprovechan su memoria interna para manejar secuencias de entradas, en vez de solo mapear entradas y salidas, son capaces de aprovechar una función de mapeo para las entradas sobre el tiempo, hacia una salida (Lazzeri, 2021).
- *Long short-term memory (LSTM)* modifica las RNN para resolver el problema del desvanecimiento de gradiente, haciéndolo capaz de aprender a largo plazo. La LSTM tiene mecanismos internos llamados puertas y estados de célula que pueden regular el flujo de información. La célula toma decisiones sobre qué, cuánto y cuándo guardar y liberar información. Estos aprenden cuándo permitir a la información entrar, salir o eliminar, a lo largo del proceso iterativo de hacer conjeturas, retropropagación de error y ajuste de pesos vía gradiente descendente (Lazzeri, 2020). Una variante de esta red neuronal es *gate recurrent unit (GRU)*.
- *Seq2Seq (sequence-to-sequence)* es aquel que toma como entrada a una secuencia de objetos (ya sea de palabras, letras, series de tiempo, etc.) y da otra secuencia de objetos como salida. El modelo se compone de un codificador y decodificador. Codificador consiste en una RNN (también puede ser LSTM o GRU para evitar el desvanecimiento de gradiente). Esta toma la secuencia de entrada hasta generar el último estado oculto. En el caso de un codificador LSTM, también se considera el último estado de la célula. Decodificador consiste en una RNN, el último estado oculto del decodificador se usa como entrada de todas las células, y si fuera LSTM, el último estado de célula se usa en la primera célula del decodificador (Sutskever et al., 2014).

3. TRABAJOS RELACIONADOS

Cook y Smalter Hall (2017) predicen la tasa de desempleo la cual fue recolectada mensualmente por la Agencia Estadounidense de Trabajo y Estadísticas (US Bureau of

Labor and Statistics). Se usaron modelos basados en diferentes arquitecturas de redes neuronales (perceptrón multicapa [MLP], red neuronal convolucional [CNN], *long short-term memory* [LSTM] y *encoder-decoder* [con LSTM]), y se compararon con el modelo estadístico DARM. La métrica empleada fue el error medio absoluto (MAE). Se resalta que la arquitectura *encoder-decoder* supera a los modelos de comparación en todos los horizontes de predicción (hasta cuatro trimestres).

Jung et al. (2018) predicen el crecimiento del PBI real a corto plazo para 7 países de diferentes geografías y desarrollo económico (Alemania, México, Filipinas, España, Reino Unido, Estados Unidos y Vietnam). La fuente fue la base de datos World Economic Outlook (WEO) del Fondo Monetario Internacional. También se agregaron datos de índices de mercados de acciones, precios de energías, entre otros, proveídos por Bloomberg, y datos procedentes de la Guía Internacional de Riesgo País (ICRG). Se emplearon técnicas de *machine learning* como *elastic nets*, *super learner* y RNN. La meta del artículo fue evaluar si estas técnicas podían mejorar la precisión en la predicción, concluyendo que superan a los empleados por la WEO.

Viswanath et al. (2019) proponen pronosticar los monzones en la región central de la India por medio de clasificación, usando métodos de *deep learning* como LSTM y Seq2Seq, y comparando con métodos tradicionales como SVM y KNN. Los datos fueron mediciones espaciotemporales de la región central de la India, la cual considera las lluvias diarias durante junio a septiembre, desde los años de 1948 al 2014, cuya fuente fue el Departamento Meteorológico de la India. El trabajo concluye que los modelos propuestos se desempeñan mejor que los modelos de clasificación SVM y KNN.

Nguyen y Nguyen (2020) proponen un modelo de pronóstico del PBI por medio de *transfer learning*. Para este propósito se implementaron modelos basados en LSTM y *encoder-decoder*, con y sin *transfer learning*. Los datos se obtuvieron de JSTdatasetR4 y están comprendidos de 17 variables macroeconómicas de 17 países desarrollados, lo cual dividen en 2 grupos: datos de 16 países para el entrenamiento, y el restante para afinamiento. Se evaluaron los modelos con las métricas: RMSE, nRSME, MAPE, sMAPE; que, según los resultados, los modelos con *transfer learning* tienen mejor desempeño en el preentrenamiento y en el afinamiento.

Kelany et al. (2020) emplean modelos basados en LSTM para pronosticar precios futuros para acciones de bajo, mediano y alto riesgo, y los comparan con regresión logística y *random forest*. Se usaron las métricas MAE y RMSE para las evaluaciones, de las cuales el modelo LSTM supera a los otros dos modelos.

Zyatkov y Krivorotko (2021) construyen un indicador para pronosticar el comienzo de recesiones en la economía de EE. UU. usando métodos de *machine learning*. Para este propósito se emplearon variables socioeconómicas las cuales son típicas en

periodos precrisis. Los métodos empleados fueron KNN, SVM, *random forest*, perceptrón multicapa y LSTM. Los horizontes para el indicador son los siguientes 6, 12 y 24 meses.

4. METODOLOGÍA

En esta sección, se explican los procesos llevados a cabo para construir y entrenar los modelos de predicción.

4.1 Recolección de datos

Se recolectaron series de tiempo de indicadores macroeconómicos de 6 países Latinoamericanos: Argentina, Brasil, Chile, Colombia, México y Perú.

- Anuales. Consisten en 14 series y 360 observaciones (en total 5040 datos), las cuales son el PBI real, componentes de consumo del PBI, valores agregados de industria y agricultura, precios de *commodities*. Están medidos en dólares americanos (USD) y a precios constantes de 2010, para que los indicadores no estén afectados por la inflación. Las observaciones empiezan desde el año 1960 hasta 2020. Las series se obtuvieron de la base de datos del Banco Mundial (2021). La descripción de las características se muestra en la tabla 1.
- Trimestrales. Son 11 series y 512 observaciones (en total 5632 datos), conformadas por el PBI real y por sectores. Medidos en volumen de moneda nacional y por año base para que no sean afectados por la inflación. El año base y periodo de observación depende de cada país: Argentina año base 2004 y periodo del 2004-2021, Brasil año base 2010 y periodo de 1996-2021, Chile año base 2013 y periodo 1996-2021, Colombia año base 2015 y periodo 2005-2021, México año base 2013 y periodo 1993-2021, y Perú año base 2007 y periodo 2007-2021. Los datos fueron recolectados de la web de la Comisión Económica para América Latina y el Caribe (2021). Las características se describen en la tabla 2.

Tabla 1

Características de las series anuales

Característica	Descripción
GDP	Producto bruto interno real (en USD, 2010).
CONS_TOTAL	Gasto de consumo final (consumo privado más Gobierno) (en USD, 2010).
INVEST	Formación bruta de capital (inversión) (en USD, 2010).
EXPORT	Exportaciones de bienes y servicios (en USD, 2010).
IMPORT	Importaciones de bienes y servicios (en USD, 2010).
AGRICUL	Agricultura, valor agregado (en USD, 2010).

(continúa)

(continuación)

INDUST	Industria, valor agregado (en USD, 2010).
INFLA_GDP_DEF	Inflación, deflactor del PBI (en porcentaje).
ENERGY	Índice anual real de energía (carbón, crudo de petróleo y gas natural) (en USD, 2010).
AGRICULTURE	Índice anual real de agricultura (alimentos, bebidas, materias primas agrícolas).
FERTILIZERS	Índice anual real de fertilizantes (en USD, 2010).
METMIN	Índice anual real de metales y minerales (aluminio, cobre, hierro, plomo, níquel, zinc) (en USD, 2010).
PRECIOUSMET	Índice anual real de metales preciosos (oro, plata y platino) (en USD, 2010).
HMUUV	Índice deflactor de precios de <i>commodities</i> (en porcentaje, 2010: 100 %).

Nota. Series obtenidas del Banco de Datos del Banco Mundial, 2021 (Recuperado el 24 de octubre del 2021, de <https://databank.bancomundial.org/>).

Tabla 2

Características de las series trimestrales

Característica	Descripción
GDP	Producto bruto interno (en moneda local).
AGRO	Agricultura, ganadería, caza, silvicultura y pesca (en moneda local).
MINERIA	Explotación de minas y canteras (en moneda local).
INDUS	Industrias manufactureras (en moneda local).
SUMELECAGUA	Suministro de electricidad, gas y agua.
CONSTR	Construcción.
COMER	Comercio al por mayor y al por menor, reparación de bienes, hoteles y restaurantes.
TRANSPCOM	Transporte, almacenamiento y comunicaciones.
INMOBIL	Intermediación financiera, actividades inmobiliarias, empresariales y de alquiler.
GOBIERNO	Administración pública, defensa, seguridad social obligatoria, enseñanza, servicios sociales.
IMPUESTOS	Impuestos.

Nota. Series obtenidas de CEPALSTAT: Bases de Datos y Publicaciones Estadísticas, por la Comisión Económica para América Latina y el Caribe, 2021 (Recuperado el 14 de diciembre del 2021 de <https://statistics.cepal.org/portal/cepalstat/dashboard.html>).

4.2 Preprocesamiento

El siguiente preprocesamiento de los datos nos permitirá tener mejores resultados, el cual se divide en cuatro fases:

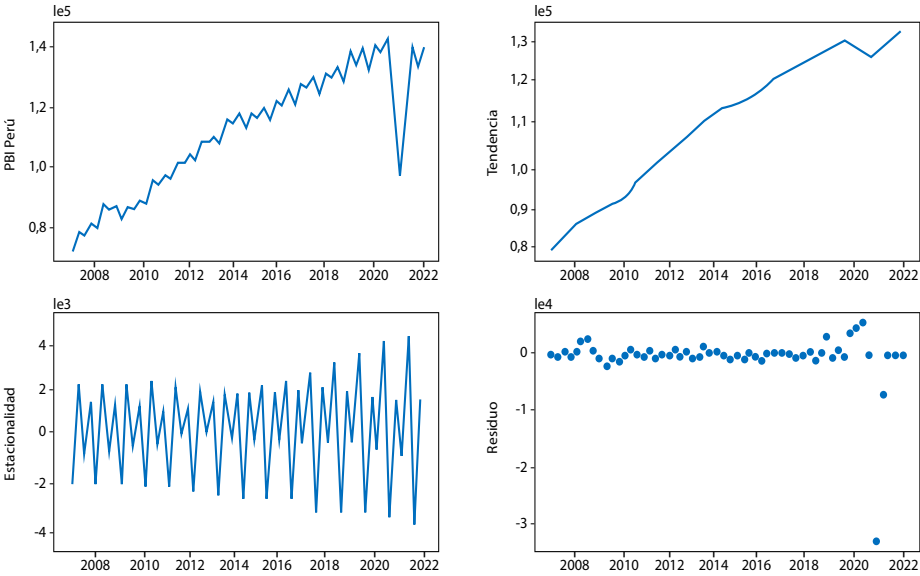
4.2.1 Descomposición de series temporales

Se descomponen las series trimestrales (y_t) en sus tres componentes: tendencia (T_t), estacionalidad (S_t) y residuo (r_t) (ecuación 1). Estas se adicionarán al conjunto de datos como características. Para este propósito se aplica la técnica de descomposición STL (*seasonal and trend decomposition using loess*) debido a que es robusta para la detección de datos atípicos (Cleveland et al., 2021). En la figura 1 se ejemplifica este proceso.

$$y_t = T_t + S_t + r_t \tag{1}$$

Figura 1

Ejemplo de descomposición de serie trimestral



Nota. Resultado de descomponer la serie del PBI trimestral de Perú aplicando la técnica de descomposición STL.

4.2.2 Reemplazo de datos atípicos

Se detectan los datos atípicos de los residuos aplicando la regla del rango inter-cuartil (IQR') y luego reemplazándolos por cero, como lo vemos en las ecuaciones 2 a 5. Con los residuos resultantes (r'), se recomponen las series (y'_t) según la ecuación 6. Esta nueva serie reemplazará a la original en el conjunto de datos. Con este proceso retiramos el impacto negativo de la pandemia del COVID-19. Un ejemplo de este proceso lo observamos en la figura 2.

$$IQR^r = Q_3^r - Q_1^r \quad (2)$$

$$l = Q_1^r - 1.5 \times IQR^r \quad (3)$$

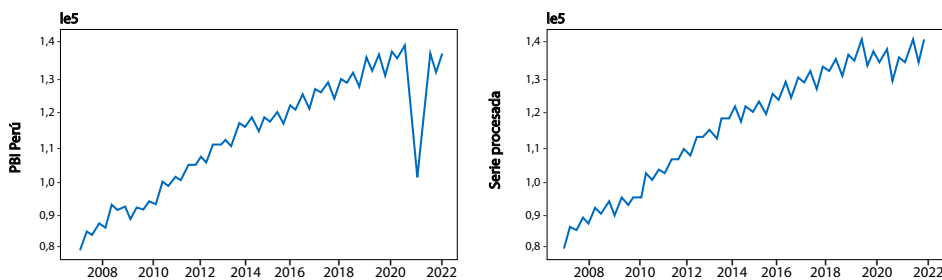
$$u = Q_3^r + 1.5 \times IQR^r \quad (4)$$

$$r' = f(r) = \begin{cases} 0, & (r < l) \vee (r > u) \\ r, & l \leq r \leq u \end{cases} \quad (5)$$

$$y'_t = T_t + S_t + r'_t \quad (6)$$

Figura 2

Ejemplo de reemplazo de datos atípicos en series trimestrales



Nota. Resultado de aplicar la técnica a la serie del PBI trimestral de Perú.

4.2.3 Escalamiento mínimo-máximo

Los datos de entrada y salida se escalan por mínimo y máximo hacia el rango (0, 1) para el correcto funcionamiento de los algoritmos de aprendizaje. La ecuación 7 muestra cómo se calcula.

$$x' = \frac{x - x_{\text{mín.}}}{x_{\text{máx.}} - x_{\text{mín.}}} \quad (7)$$

4.2.4 Generación de características pasadas y futuras

Se generan características de observaciones de L pasos de tiempo anteriores (*lag features*) a partir de los datos de entrada X que se usarán para pronosticar H pasos a futuro (*future features*) de los datos de salida Y . Para el estudio, se pronosticarán solo

$H = 1$ paso de tiempo a futuro y el dato de salida Y será la serie del PBI, esto será de un año para las series anuales, y de un trimestre para el otro caso. El número de pasos a pasado L se encontrarán por ajuste de hiperparámetros. La ecuación 8 muestra cómo serán los datos de entrada y salida.

$$(X_t, Y_t) \rightarrow (X_{t-L}, \dots, X_{t-1}, X_t, Y_{t+H}) \quad (8)$$

4.3 Modelos e hiperparámetros

Para la construcción de los modelos anuales se tomaron 13 características de entrada y una de salida (el PBI a pronosticar), para los trimestrales son 43 de entrada y 1 de salida. Se fijaron los siguientes hiperparámetros para todos los modelos: la función de activación ReLU, el optimizador Adam, la función de pérdida que corresponde al error cuadrático medio (MSE) y la condición de parada de 50 épocas después de encontrar el menor error de validación. Estos parámetros fueron considerados según lo aplicado en los trabajos de Zyatkov y Krivorotko (2021) y Viswanath et al. (2019).

Respecto a los hiperparámetros con espacios de búsqueda presentes en las tablas 3 a 7 se detalla lo siguiente: se consideraron los números de pasos pasados en un rango de 4 a 20 para encontrar el óptimo de forma experimental; de la misma forma se consideraron los números para capas ocultas. Para el número de nodos, kernels y células se eligieron múltiplos de 32, y la ratio de aprendizaje son números pequeños (potencias entre 10^{-2} y 10^{-4}), basándose en el trabajo de Guimarães (2022).

Para la búsqueda de los hiperparámetros óptimos se empleó la técnica de hiperbanda, propuesto por Li et al. (2018), donde recursos predefinidos como iteraciones, muestras o características son asignados a muestras aleatorias de configuraciones de hiperparámetros, de los cuales se buscan los óptimos por medio de *successive halving*, técnica propuesta por Jamieson y Talwalkar (2016), la cual descarta las configuraciones con peores métricas y asigna más recursos a las mejores por medio de iteraciones. A continuación, se detalla la construcción de los modelos y las tablas 4 a 7, las cuales indican el espacio de búsqueda y los valores óptimos encontrados para las series anuales y trimestrales.

4.3.1 Perceptrón multicapa (MLP)

En la tabla 3 se muestran los hiperparámetros a buscar, como el número de pasos pasados necesarios para pronosticar, el número de capas ocultas, el número de nodos por capa oculta y la ratio de aprendizaje.

Tabla 3

Hiperparámetros para modelos basados en MLP

Hiperparámetro	Espacio de búsqueda	Óptimo anual	Óptimo trimestral
Número de pasos pasados	[4, 6, 8, 10, 12, 14, 16, 18, 20]	6	4
Número de capas ocultas	[1, 2, 3, 4, 5]	5	3
Número de nodos por capa	[32, 64, 96, 128, 160]	64	96
Ratio de aprendizaje	[10^{-2} , 10^{-3} , 10^{-4}]	10^{-2}	10^{-2}

4.3.2 Red neuronal convolucional (CNN)

La tabla 4 indica que se buscan el número de pasos pasados óptimo para este modelo, el número de kernels para la capa convolucional (la cual será una sola capa y unidimensional), el tamaño de kernel y *max pooling*, como también el número de nodos de capa densa, que procesarán la salida de la capa convolucional. También se busca la ratio de aprendizaje.

Tabla 4

Hiperparámetros para modelos basados en CNN

Hiperparámetro	Espacio de búsqueda	Óptimo anual	Óptimo trimestral
Número de pasos pasados	[4, 6, 8, 10, 12, 14, 16, 18, 20]	4	4
Número kernels capa convolucional	[32, 64, 96, 128, 160]	128	128
Tamaño de kernel	[2, 3]	2	2
Tamaño de <i>max-pooling</i>	-	2	2
Número de nodos capa densa	[32, 64, 96, 128, 160]	128	96
Ratio de aprendizaje	[10^{-2} , 10^{-3} , 10^{-4}]	10^{-2}	10^{-3}

4.3.3 Long short-term memory (LSTM)

La tabla 5 indica que se buscarán el número de pasos pasados, el número de células por capa, la ratio de aprendizaje. Se usarán dos capas LSTM.

Tabla 5

Hiperparámetros para modelos basados en LSTM

Hiperparámetro	Espacio de búsqueda	Óptimo anual	Óptimo trimestral
Número de pasos pasados	[4, 6, 8, 10, 12, 14, 16, 18, 20]	4	4
Número de capas LSTM	-	2	2
Número de células por capa	[32, 64, 96, 128, 160]	128	160
Ratio de aprendizaje	[10^{-2} , 10^{-3} , 10^{-4}]	10^{-2}	10^{-3}

4.3.4. Gated recurrent unit (GRU)

En la tabla 6 se muestrans los hiperparámetros a buscar: número de pasos pasados, número de células por capa y ratio de aprendizaje. Se fija el número de capas GRU a dos.

Tabla 6

Hiperparámetros para modelos basados en GRU

Hiperparámetro	Espacio de búsqueda	Óptimo anual	Óptimo trimestral
Número de pasos pasados	[4, 6, 8, 10, 12, 14, 16, 18, 20]	6	4
Número de capas GRU	-	2	2
Número de células por capa	[32, 64, 96, 128, 160]	32	160
Ratio de aprendizaje	[10^{-2} , 10^{-3} , 10^{-4}]	10^{-2}	10^{-3}

4.3.5 Sequence to sequence (Seq2Seq)

La tabla 7 indica el número de pasos pasados a buscar, como también el número de células por capa, ya sea para el codificador y decodificador, y por último la ratio de aprendizaje.

Tabla 7

Hiperparámetros para modelos basados en Seq2Seq

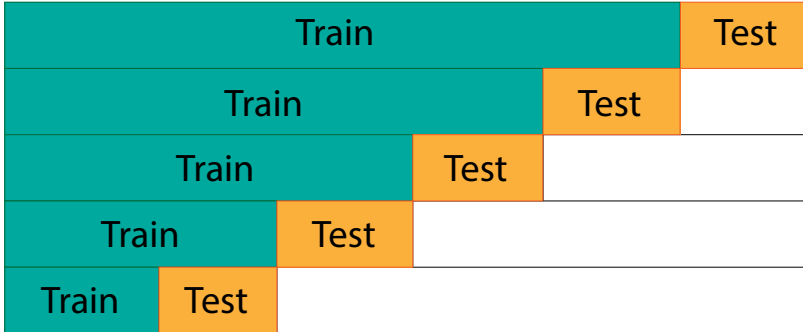
Hiperparámetro	Espacio de búsqueda	Óptimo anual	Óptimo trimestral
Número de pasos pasados	[4, 6, 8, 10, 12, 14, 16, 18, 20]	16	4
Número de células por capa (codificador/decodificador)	[32, 64, 96, 128, 160]	64	64
Ratio de aprendizaje	[10^{-2} , 10^{-3} , 10^{-4}]	10^{-2}	10^{-3}

4.4 Entrenamiento y evaluación de modelos

Para el entrenamiento y la evaluación de los modelos, se emplea la técnica de validación cruzada continua, la cual consiste que a partir del conjunto de datos se generan distintos pares de subconjuntos para entrenamiento y prueba. Los subconjuntos de entrenamiento parten desde el inicio y anteceden a los de prueba ya que son series de tiempo y dependen del orden de sus observaciones (Hyndman & Athanasopoulos, 2021). Para la evaluación se generarán 5 pares de subconjuntos como se ejemplifica en la figura 3. Luego se entrenará un modelo por cada conjunto y se evaluarán sus métricas. El promedio de estas métricas nos servirá para comparar entre los modelos propuestos.

Figura 3

Diagrama de subconjuntos de datos de entrenamiento y prueba usando validación cruzada



Las métricas que se emplearán son el error absoluto medio (MAE), la raíz cuadrada de error cuadrático medio (RMSE) y el error absoluto porcentual medio (MAPE), las cuales están descritas en las ecuaciones 9, 10 y 11. Se consideraron estas métricas basándose de los trabajos de Cook & Smalter Hall (2017), Nguyen y Nguyen (2020) y Kelany et al. (2020).

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - y_i^*| \tag{9}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - y_i^*)^2} \tag{10}$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - y_i^*}{y_i} \right| \times 100 \tag{11}$$

5. RESULTADOS Y DISCUSIONES

Se presentan los resultados por cada país en cada subsección, la cual contiene tablas con los promedios de las métricas de desempeño de los modelos en las evaluaciones y se resaltan los valores más bajos. Las gráficas de cajas y bigotes muestran la distribución de las métricas en las evaluaciones individuales. También se grafica la predicción a lo largo de las series, usando los modelos con mejores resultados.

5.1 Argentina

Según los resultados de la tabla 8, en promedio CNN tiene mejor desempeño en datos anuales y MLP para datos trimestrales. En las figuras 4, 5 y 6 observamos que las

métricas individuales no están muy dispersas para los modelos de mejor desempeño a comparación de los modelos restantes. En la figura 7 se pronostica el PBI a lo largo de todo el conjunto de datos y se compara con el estado original del PBI.

Tabla 8

Promedio de métricas en la evaluación de modelos para Argentina

		CNN	GRU	LSTM	MLP	Seq2Seq
Anual	MAE	0,039549	0,055836	0,076992	0,048956	0,060864
	RMSE	0,049235	0,068924	0,086598	0,057614	0,072513
	MAPE	8,110362	11,101573	14,114396	9,857729	10,306283
Trimestral	MAE	0,045966	0,051465	0,049192	0,038778	0,058814
	RMSE	0,054223	0,062224	0,059240	0,046067	0,066914
	MAPE	6,813480	7,614213	7,445939	5,817453	8,917134

Figura 4

Distribución de MAE en la evaluación de modelos de Argentina

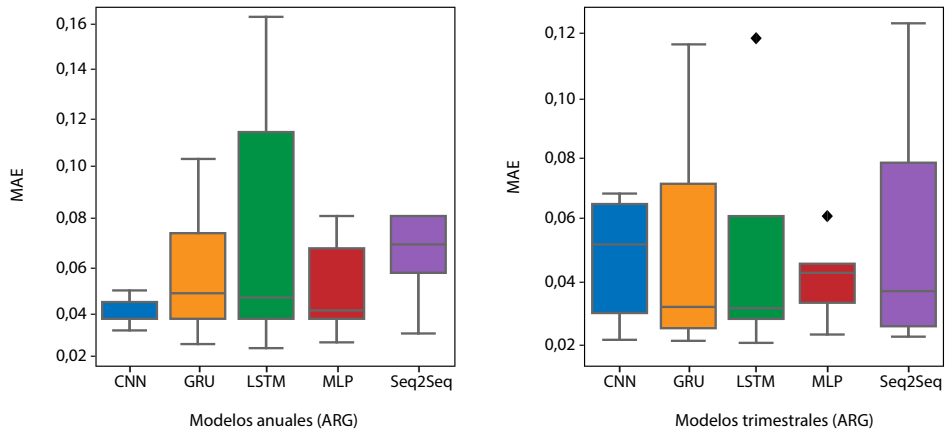


Figura 5

Distribución de RMSE en la evaluación de modelos de Argentina

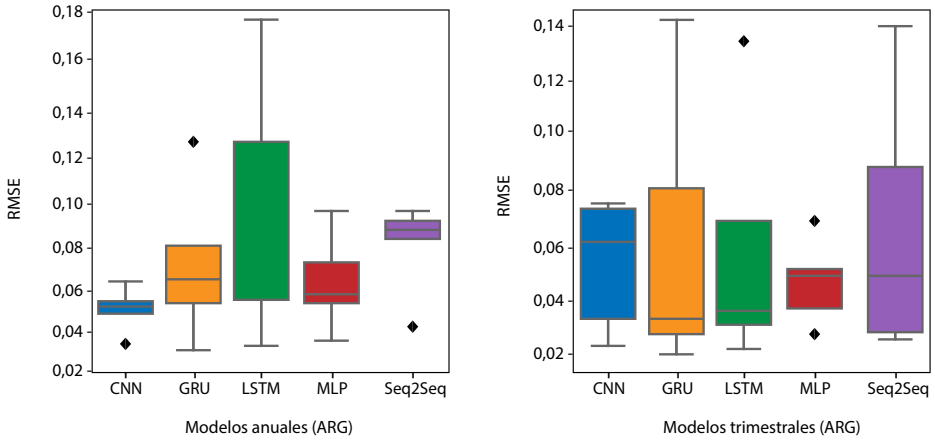


Figura 6

Distribución de MAPE en la evaluación de modelos de Argentina

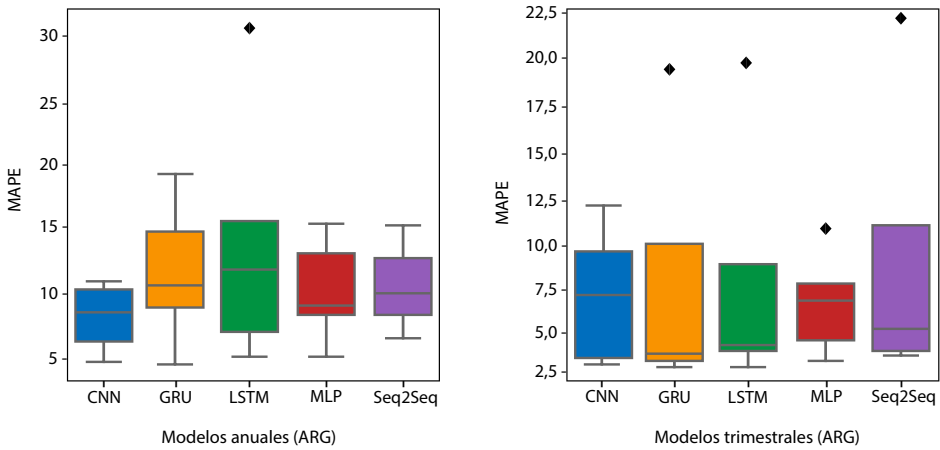
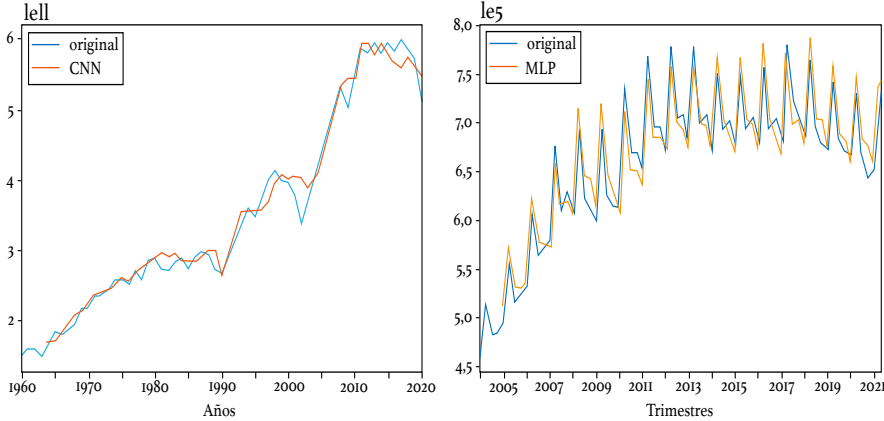


Figura 7

Predicción del PBI usando los modelos con mejores resultados para Argentina



4.2 Brasil

En los resultados de la tabla 9 observamos que, para datos anuales, si bien MLP tiene mejor promedio en MAE, es el modelo Seq2Seq que supera en las otras métricas (RMSE y MAPE). Para datos trimestrales, MLP presenta los mejores promedios. En las figuras 8, 9 y 10 observamos que Seq2Seq tiene menores valores individuales para el caso anual, y lo mismo con MLP en trimestrales. En la figura 11 se pronostica con todo el conjunto de datos y observamos que el modelo Seq2Seq anual predice por debajo de lo esperado en los últimos años.

Tabla 9

Promedio de métricas en la evaluación de modelos para Brasil

		CNN	GRU	LSTM	MLP	Seq2Seq
Anual	MAE	0,031627	0,052510	0,037535	0,026192	0,026595
	RMSE	0,035953	0,059284	0,043709	0,032024	0,029517
	MAPE	6,310547	9,748143	7,639060	4,934667	3,642307
Trimestral	MAE	0,030357	0,029789	0,024329	0,023548	0,025053
	RMSE	0,035878	0,036870	0,029941	0,028094	0,030952
	MAPE	4,921294	5,376683	4,354568	4,155495	4,331294

Figura 8

Distribución de MAE en la evaluación de modelos de Brasil

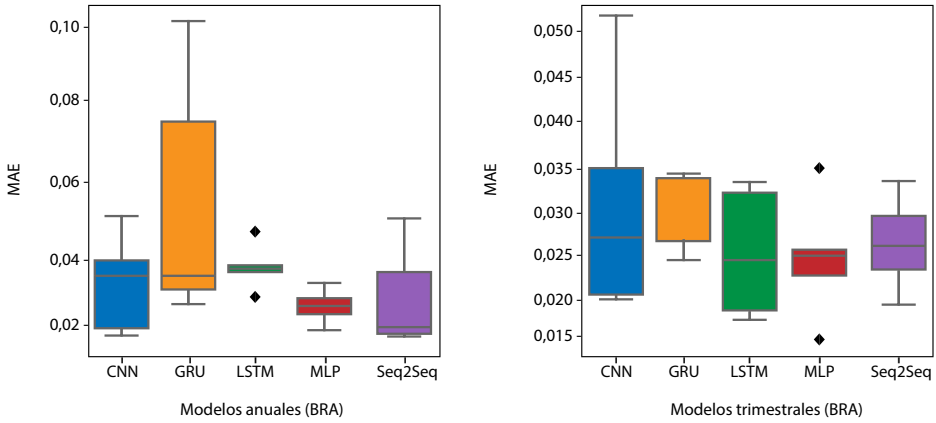


Figura 9

Distribución de RMSE en la evaluación de modelos de Brasil

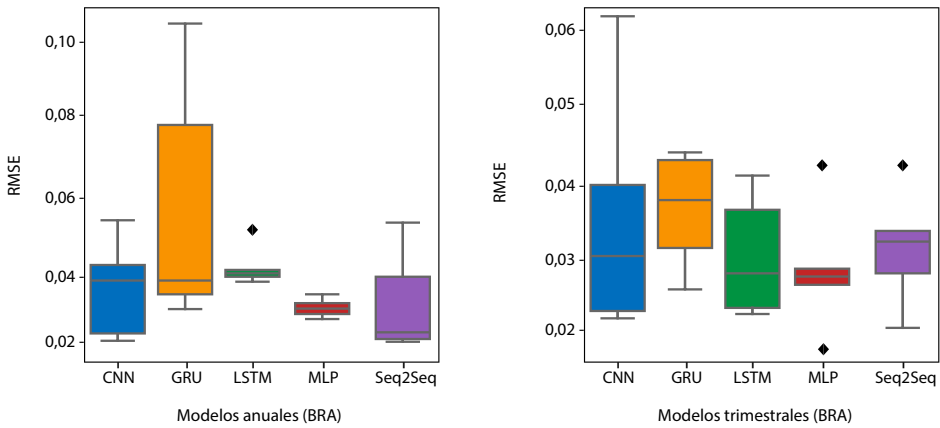


Figura 10

Distribución de MAPE en la evaluación de modelos de Brasil

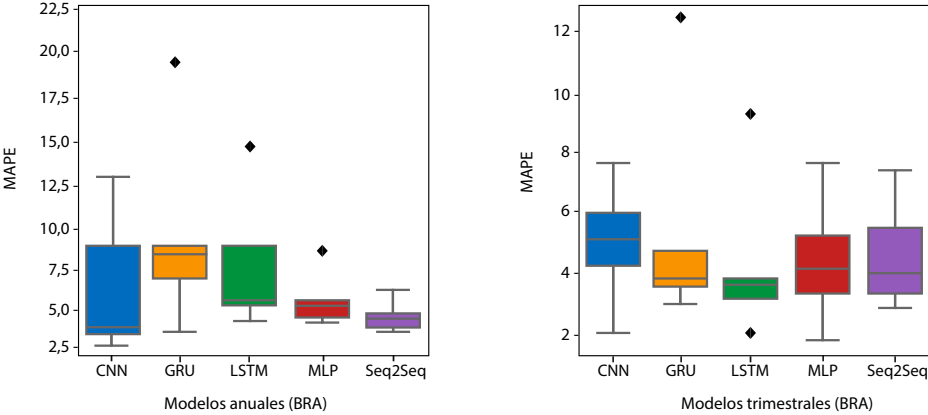
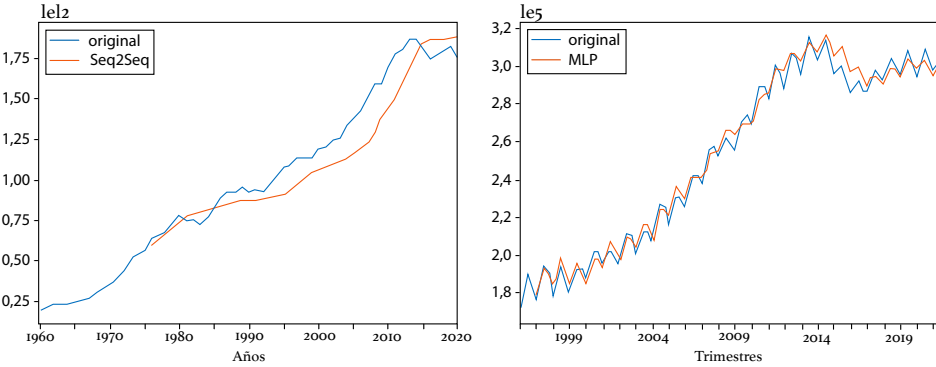


Figura 11

Predicción del PBI usando los modelos con mejores resultados para Brasil



4. Chile

Según los resultados de la tabla 10, en datos anuales, en promedio, MLP tiene mejor desempeño en MAE y RMSE, y Seq2Seq en MAPE. Para datos trimestrales, MLP supera a los restantes. En las figuras 12, 13 y 14 observamos que MLP tiene mejores métricas individuales en los dos casos. La figura 15 pronostica el PBI con los modelos MLP respectivamente.

Tabla 10

Promedio de métricas en la evaluación de modelos para Chile

		CNN	GRU	LSTM	MLP	Seq2Seq
Anual	MAE	0,034325	0,033571	0,033604	0,021722	0,024758
	RMSE	0,041179	0,040875	0,039730	0,027051	0,031637
	MAPE	14,892479	12,863911	14,133017	9,772485	7,144571
Trimestral	MAE	0,024589	0,027807	0,031122	0,021260	0,032499
	RMSE	0,031262	0,034194	0,038984	0,026748	0,041087
	MAPE	5,865346	6,656868	7,215769	5,031725	7,081380

Figura 12

Distribución de MAE en la evaluación de modelos de Chile

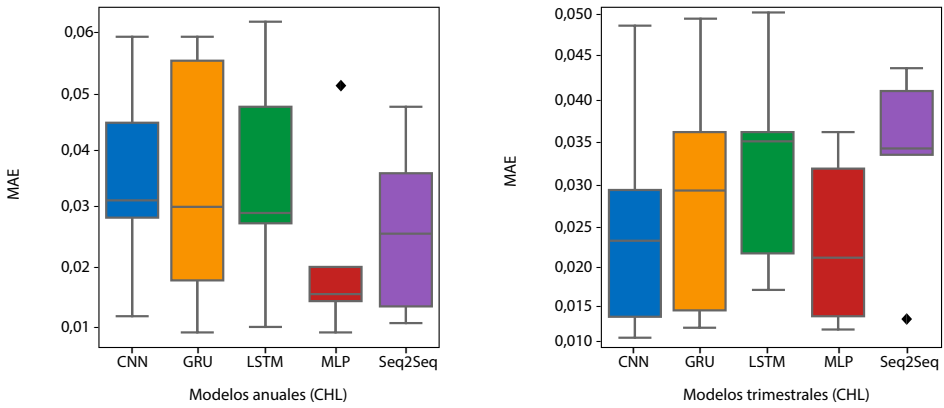


Figura 13

Distribución de RMSE en la evaluación de modelos de Chile

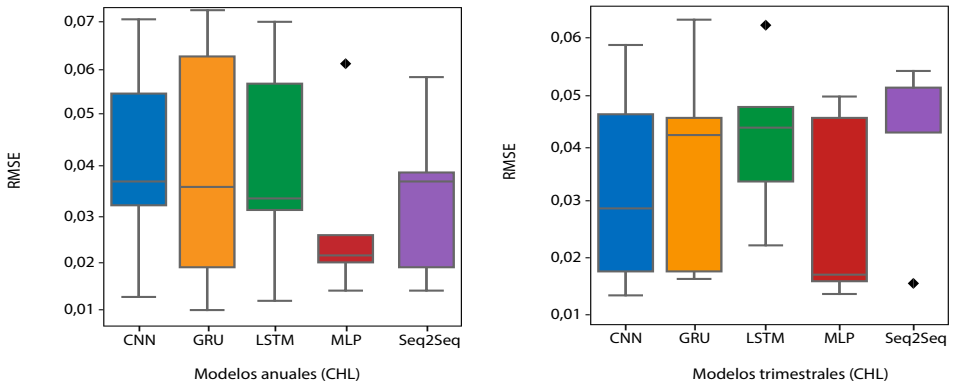


Figura 14

Distribución de MAPE en la evaluación de modelos de Chile

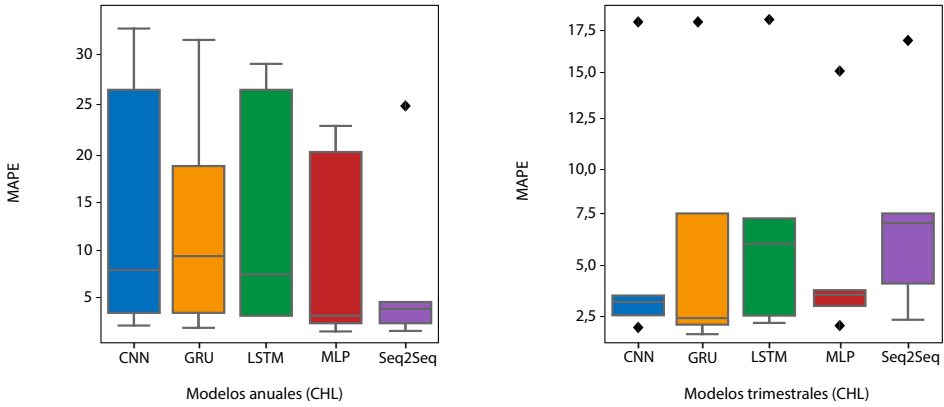
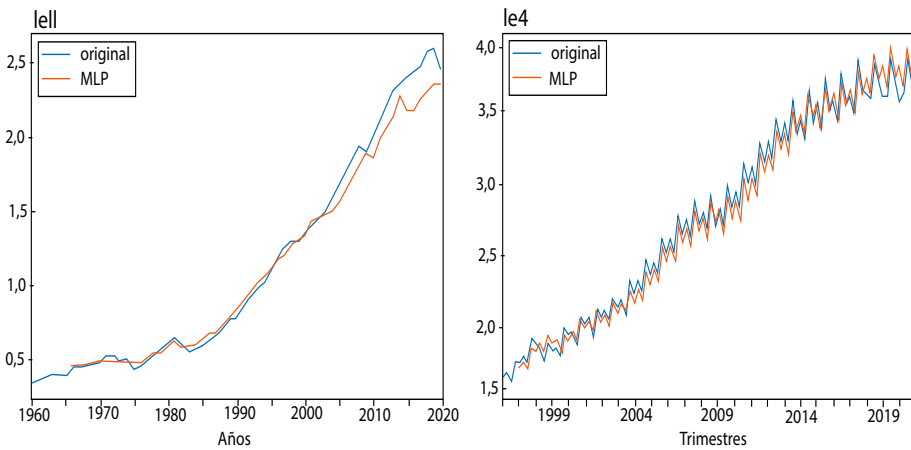


Figura 15

Predicción del PBI usando los modelos con mejores resultados para Chile



4.4 Colombia

En la tabla 11, observamos que, en promedio, para datos anuales, CNN tiene mejor desempeño en MAE y RMSE, y Seq2Seq en MAPE; y para trimestrales, MLP supera a los demás modelos. En las figuras 16, 17 y 18, CNN tiene mejores valores individuales para datos anuales y MLP en trimestrales. En la figura 19 se pronostica el PBI y observamos que CNN se aleja del valor original en los últimos años.

Tabla 11

Promedio de métricas en la evaluación de modelos para Colombia

		CNN	GRU	LSTM	MLP	Seq2Seq
Anual	MAE	0,023863	0,036272	0,036205	0,038001	0,026422
	RMSE	0,026892	0,046048	0,041337	0,044638	0,029779
	MAPE	6,390575	9,403895	9,142741	9,585601	5,378279
Trimestral	MAE	0,026070	0,031893	0,040862	0,020612	0,039732
	RMSE	0,031688	0,038621	0,048573	0,024807	0,049222
	MAPE	4,851385	5,814529	7,737369	3,890099	6,512726

Figura 16

Distribución de MAE en la evaluación de modelos de Colombia

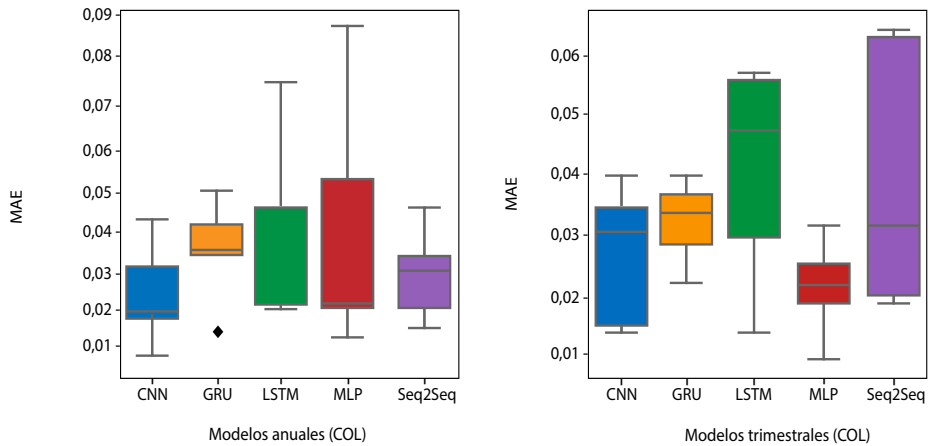


Figura 17

Distribución de RMSE en la evaluación de modelos de Colombia

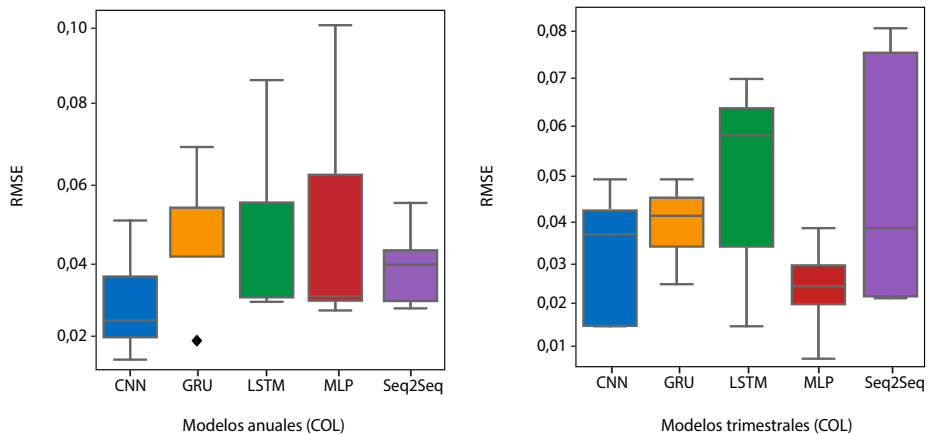


Figura 18

Distribución de MAPE en la evaluación de modelos de Colombia

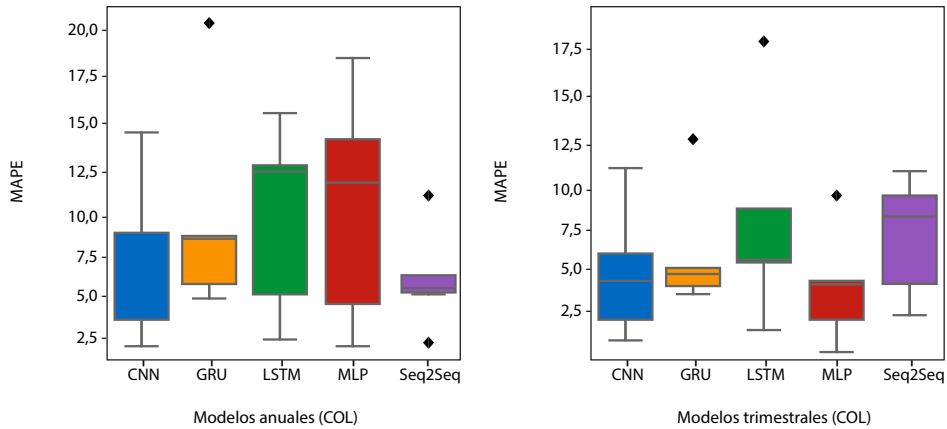
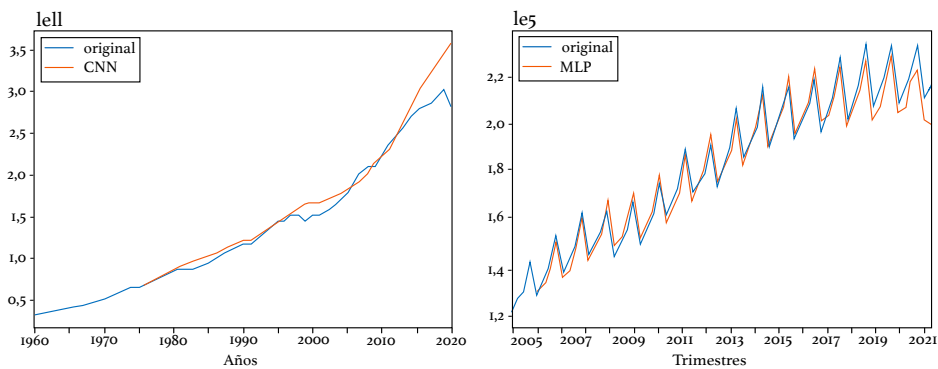


Figura 19

Predicción del PBI usando los modelos con mejores resultados para Colombia



4.5. México

En los resultados de la tabla 12, observamos que el modelo Seq2Seq tiene mejor desempeño en promedio para datos anuales; para trimestrales, el mejor es MLP. En las figuras 20, 21 y 22 se observa que Seq2Seq y MLP tienen los mejores valores individuales respectivamente. En la figura 23 se pronostica el PBI y se realiza la comparación de los resultados a lo largo de todo el conjunto de datosn anuales y trimestrales.

Tabla 12

Promedio de métricas en la evaluación de modelos para México

		CNN	GRU	LSTM	MLP	Seq2Seq
Anual	MAE	0,031604	0,044983	0,038439	0,037753	0,029141
	RMSE	0,039263	0,054898	0,045722	0,043099	0,034944
	MAPE	6,539337	8,356517	7,242875	8,492722	4,779532
Trimestral	MAE	0,023512	0,023032	0,024003	0,014631	0,024424
	RMSE	0,028603	0,028353	0,028500	0,017840	0,029421
	MAPE	4,350519	4,232047	4,283476	2,713248	4,448161

Figura 20

Distribución de MAE en la evaluación de modelos de México

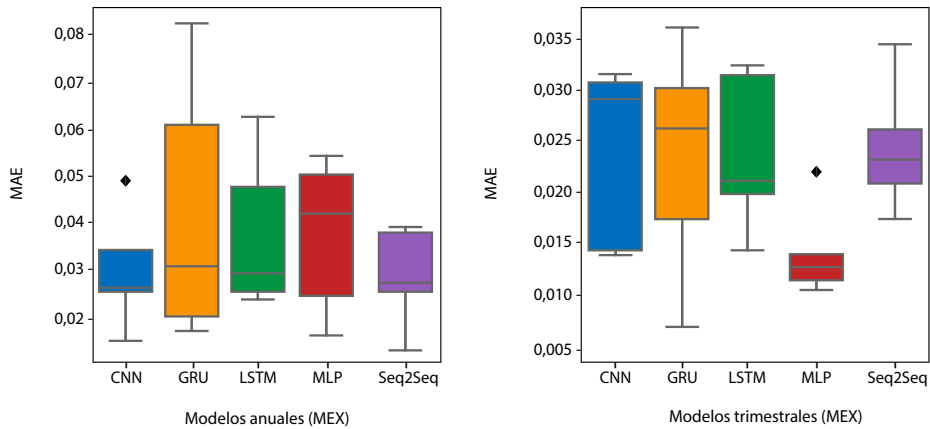


Figura 21

Distribución de RMSE en la evaluación de modelos de México

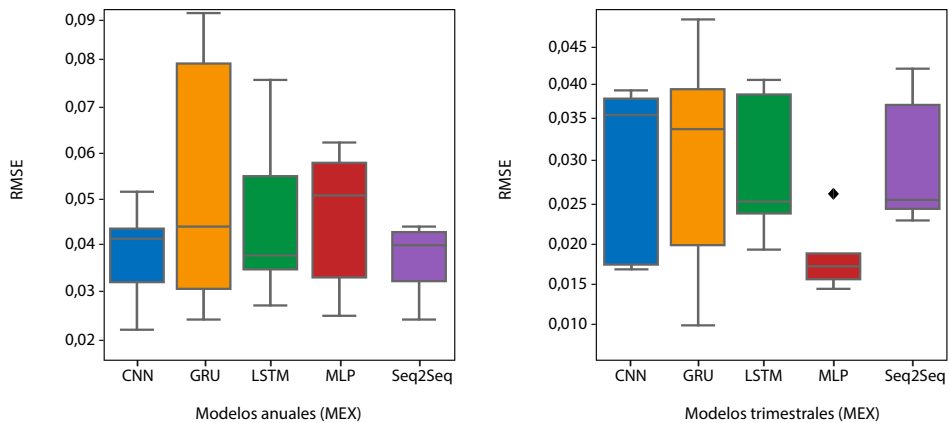


Figura 22

Distribución de MAPE en la evaluación de modelos de México

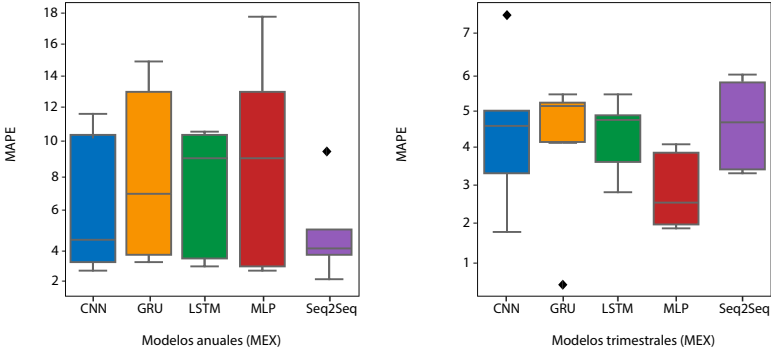
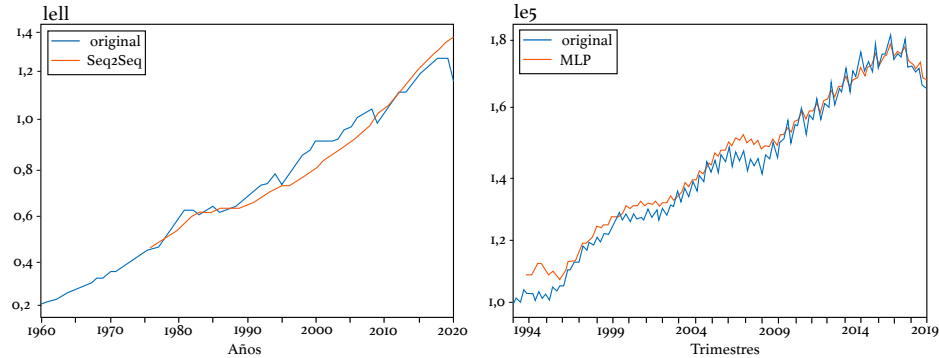


Figura 23

Predicción del PBI usando los modelos con mejores resultados para México



4.6 Perú

Según la tabla 13, en promedio, Seq2Seq tiene mejor desempeño en MAE y MAPE, y MLP en RMSE para datos anuales; mientras que, en trimestrales, MLP supera al resto. En las figuras 24, 25 y 26 se observan que los valores individuales de Seq2Seq no tienen mucha dispersión que los demás en datos anuales, y MLP tiene los valores más bajos en los datos trimestrales del PBI. La figura 27 pronostica el PBI, y en el modelo trimestral se sobreestima el valor predicho en los últimos años.

Tabla 13

Promedio de métricas en la evaluación de modelos para Perú

		CNN	GRU	LSTM	MLP	Seq2Seq
Anual	MAE	0,031455	0,043288	0,053703	0,032230	0,030423
	RMSE	0,039819	0,049165	0,061968	0,038643	0,039474
	MAPE	8,236680	11,280637	13,827532	9,438972	7,486841
Trimestral	MAE	0,025705	0,030373	0,032632	0,021909	0,030306
	RMSE	0,031833	0,036847	0,037814	0,026027	0,035709
	MAPE	4,137657	4,513389	4,620843	3,199139	4,560172

Figura 24

Distribución de MAE en la evaluación de modelos de Perú

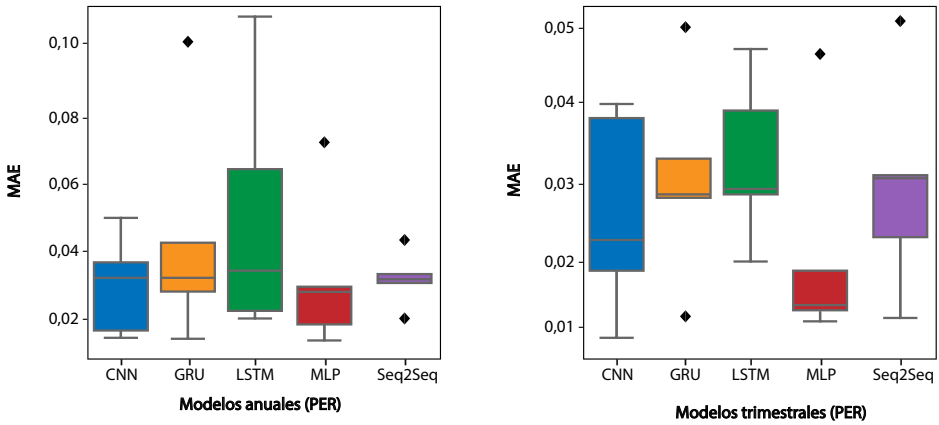


Figura 25

Distribución de RMSE en la evaluación de modelos de Perú

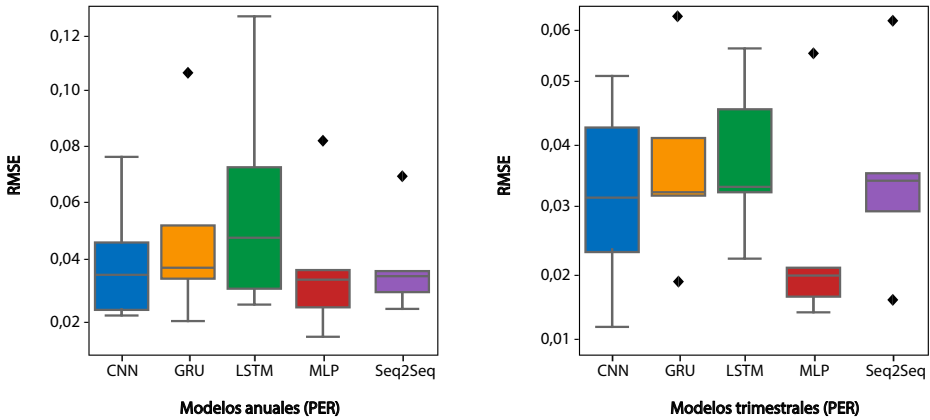


Figura 26

Distribución de MAPE en la evaluación de modelos de Perú

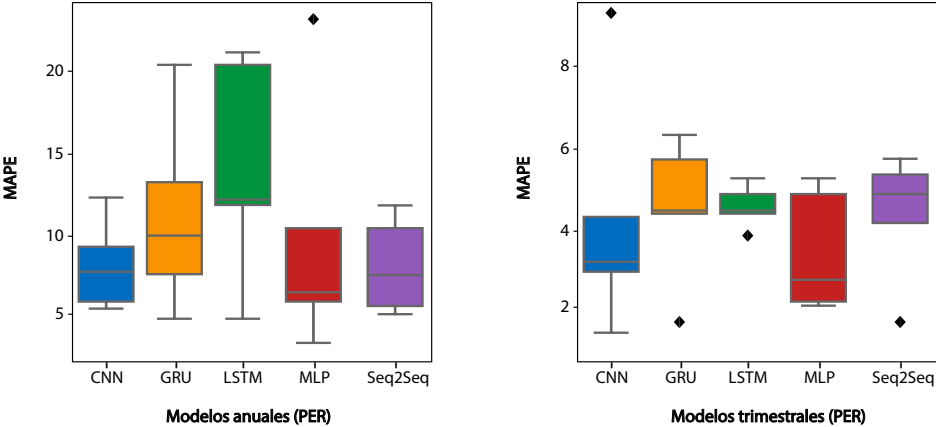
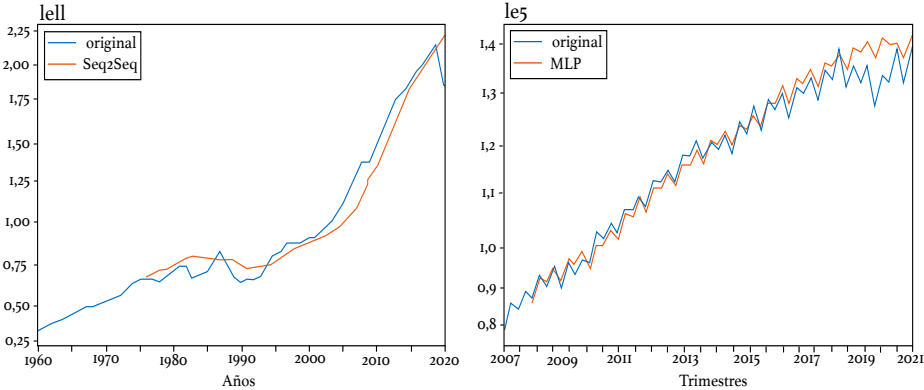


Figura 27

Predicción del PBI usando los modelos con mejores resultados para Perú



4.7 Discusión final

Se halló que los modelos CNN, MLP y Seq2Seq se desempeñan mejor en datos anuales, y que MLP supera a los demás en datos trimestrales. Mientras que, en la predicción del PBI, los modelos trimestrales se aproximan más al original, a excepción del caso de Perú (figura 27) que en el periodo de pandemia de COVID-19 no predice la crisis. Lo mismo sucede con los modelos anuales en las figuras 19, 23 y 27, al sobreestimar el último año que corresponde a la crisis.

5. CONCLUSIONES

Los indicadores macroeconómicos de América Latina tomados de las series anuales del Banco Mundial y trimestrales de la Comisión Económica para América Latina y el Caribe (CEPAL) fueron estudiados en el presente trabajo. El preprocesamiento consiste en la descomposición de las series trimestrales y detección de datos atípicos. Los distintos modelos de *deep learning* usados para pronóstico de las series temporales de los indicadores macroeconómicos fueron el perceptrón multicapa (MLP), red convolucional (CNN), LSTM, GRU y Seq2Seq. Estos modelos se aplicaron a cada país y tipo de conjunto de datos (anual o trimestral). Al evaluar su desempeño con técnica de validación cruzada continua y las métricas MAE, RMSE y MAPE, los modelos tienen diferente desempeño en distintos conjuntos de datos. Para los datos anuales y por país los modelos de mejor desempeño fueron: CNN (Argentina y Colombia), MLP (Chile) y Seq2Seq (Brasil, México y Perú). Y en datos trimestrales fue MLP (los 6 países). Se pronosticó con mayor cercanía el siguiente trimestre en los datos trimestrales. Sin embargo, el pronóstico del siguiente año para los datos anuales se alejó porque el último año observado (2020) corresponde a la pandemia.

Como trabajo a futuro se debería pronosticar también a largo plazo empleando ya no un paso a futuro sino más pasos como salida del modelo. También se debería construir un modelo general el cual se entrene con datos de un grupo de países, y probar con otros países no incluidos en el entrenamiento. Respecto a los datos de entrada, se puede experimentar con series de mayor frecuencia, por ejemplo, mensuales. Además, se pueden incluir más indicadores macroeconómicos no solamente relacionados directamente con el PBI. Esto permitirá experimentar si con más variables mejoraría el pronóstico.

Para el modelamiento, se puede mejorar explorando un mayor espacio de búsqueda de hiperparámetros, incluyendo aquellos que se habían fijado para este trabajo, como, por ejemplo, variar la cantidad de capas de convolución para CNN, la cantidad de capas ocultas en LSTM y GRU; probar otras funciones de activación, de pérdida y optimizadores; buscar el número de épocas óptimas para evitar el sobreajuste.

REFERENCIAS

- Banco Mundial. (2021). Banco de datos. Recuperado el 24 de octubre de 2021, de <https://databank.bancomundial.org/>
- Brownlee, J. (2017). *Deep learning for time series forecasting: Predict the future with MLPs, CNNs and LSTMs in Python*. Machine Learning Mastery.
- Chollet, F. (2018). *Deep learning with Python*. Manning Publications Co.

- Cleveland, R. B., Cleveland, W. S., McRae, J. E., & Terpenning, I. J. (1990). STL: A seasonal trend decomposition procedure based on loess. *Journal of Official Statistics*, 6(1), 3-33.
- Comisión Económica para América Latina y el Caribe. (2021). CEPALSTAT: Bases de Datos y Publicaciones Estadísticas. Recuperado el 14 de diciembre de 2021 de <https://statistics.cepal.org/portal/cepalstat/dashboard.html>
- Cook, T., & Smalter Hall, A. (2017). Macroeconomic indicator forecasting with deep neural networks. *The Federal Reserve Bank of Kansas City Research Working Papers*. <https://doi.org/10.18651/rwp2017-11>
- Guimarães, R. R. S. (2022). *Deep learning macroeconomics* [Tesis de maestría, Universidade Federal do Rio Grande do Sul]. <http://hdl.handle.net/10183/239533>
- Han, J., Kamber, M., & Pei, J. (2012). *Data mining: Concepts and techniques* (3.ª ed.). Elsevier/Morgan Kaufmann.
- Hyndman, R. J., & Athanasopoulos, G. (2021). *Forecasting: principles and practice* (3.ª ed.), OTexts. Recuperado el 10 de mayo de 2022 de <https://OTexts.com/fpp3>
- Jamieson, K., & Talwalkar, A. (2016). Non-stochastic best arm identification and hyperparameter optimization. *Proceedings of the 19th International Conference on Artificial Intelligence and Statistics. Proceedings of Machine Learning Research 51*, 240-248. <https://proceedings.mlr.press/v51/jamieson16.html>
- Jung, J.-K., Patnam, M., & Ter-Martirosyan, A. (2018). An algorithmic crystal ball: Forecasts-based on machine learning. *IMF Working Papers*, 18(230), 1-33. <https://doi.org/10.5089/9781484380635.001>
- Kelany, O., Aly, S., & Ismail, M. A. (2020, November). Deep learning model for financial time series prediction. *2020 14th International Conference on Innovations in Information Technology (IIT)*, 120-125.
- Lazzeri, F. (2020). *Machine learning for time series forecasting with Python*. Wiley.
- Li, L., Jamieson, K., DeSalvo, G., Rostamizadeh, A., & Talwalkar, A. (2018). Hyperband: A novel bandit-based approach to hyperparameter optimization. *Journal of Machine Learning Research*, 18(185), 1-52. <http://jmlr.org/papers/v18/16-558.html>
- Mankiw, N. G. (2014). *Macroeconomía*, (8.ª ed.). Antoni Bosch.
- Nguyen, H. T., & Nguyen, D. T. (2020). Transfer learning for macroeconomic forecasting. *2020 7th NAFOSTED Conference on Information and Computer Science (NICS)*, 332-337. <https://doi.org/10.1109/NICS51282.2020.9335848>
- Skansi, S. (2018). *Introduction to deep learning: From logical calculus to artificial intelligence*. Springer.

- Sutskever, I., Vinyals, O., & Le, Q. V. (2014). Sequence to sequence learning with neural networks. *Proceedings of the 27th International Conference on Neural Information Processing Systems (NIPS'14)*, 2, 3104-3112. <https://dl.acm.org/doi/10.5555/2969033.2969173>
- Tan, P.-N., Steinbach, M., & Kumar, V. (2006). *Introduction to data mining*. Pearson Education.
- Viswanath, S., Saha, M., Mitra, P., & Najundiah R. S. (2019), Deep learning based LSTM and SeqToSeq Models to detect monsoon spells of India. En João M. F. Rodrigues, Pedro J. S. Cardoso, Jânio Monteiro, Roberto Lam, Valeria V. Krzhizhanovskaya, Michael H. Lees, Jack J. Dongarra, Peter M. A. Sloot (Eds.), *Computational Science – ICCS 2019 (Part II, vol. 11537, pp. 204-218)*. https://doi.org/10.1007/978-3-030-22741-8_15
- Zyatkov, N., & Krivorotko, O. (2021). Forecasting recessions in the US Economy using machine learning methods. *2021 17th International Asian School-Seminar "Optimization Problems of Complex Systems (OPCS)"*, 139-146, <https://doi.org/10.1109/OPCS53376.2021.9588678>

CLASIFICACIÓN DE ORGANISMOS EN REINOS UTILIZANDO FRECUENCIA DE CODONES DE ADN

LUIS BELTRÁN PALMA TTITO
luis.palma@unsaac.edu.pe / ORCID: 0000-0002-0950-5369
Departamento Académico de Ingeniería Informática
Universidad Nacional de San Antonio Abad del Cusco, Perú

Resumen

Este estudio tiene por objetivo utilizar clasificadores de machine learning para predecir el reino al que pertenece un organismo por la frecuencia de uso de codones de ADN. Para ello se ha tomado 13 028 datos de organismos del GenBank distribuidos en once reinos y se los redujo a seis reinos (arqueas, bacterias, invertebrados, plantas, virus y vertebrados) con 9027 datos reagrupados. El proceso requirió la depuración de atributos irrelevantes, el empleo de métricas de medición de clasificadores de exactitud, precisión, sensibilidad y puntuación, así como el ajuste de hiperparámetros de los modelos. Los algoritmos de clasificación fueron *voting*, *bagging*, *boosting* y *stacking*, usando KNN, AD, MLP, SVC y RF. La selección de atributos se hizo con random forest. El ensamble *stacking*, con sus modelos, predice mejor la clasificación de organismos en el presente estudio.

PALABRAS CLAVE: *machine learning* / ensambles / frecuencia de codones ADN / reino

CLASSIFICATION OF ORGANISMS INTO KINGDOMS USING DNA CODON FREQUENCY

Abstract

This study aims to use machine learning classifiers to predict the kingdom to which an organism belongs by the frequency of use of DNA codons. The study used 13,028 data from GenBank organisms distributed in eleven kingdoms and reduced them to six kingdoms (archaea, bacteria, invertebrates, plants, viruses, and vertebrates) with 9,027 regrouped data. The process required cleaning irrelevant attributes, using measurement metrics of accuracy, precision, sensitivity, and score classifiers, and the adjustment of hyperparameters of the models. The classification algorithms were *voting*, *bagging*, *boosting*, and *stacking*, using KNN, AD, MLP, SVC, and RF. Random forest was used in selecting the attributes. The *stacking* ensemble, with its models, better predicts the classification of organisms in the present study.

KEYWORDS: *machine learning* / ensembles / DNA codon frequency / kingdom /

1. INTRODUCCIÓN

Un enfoque novedoso para la clasificación de organismos y también de genes es la frecuencia de uso de codones en el ADN codificante. Se conoce que los diferentes organismos tienen la tendencia a utilizar determinados codones para la producción de aminoácidos, por lo tanto esta particularidad es innata a cada grupo de organismos (Im & Choi, 2017; Sharp, 2010), entonces ¿el análisis de la frecuencia de uso de un determinado codón será capaz de clasificar a los organismos en reinos? Para responder a la pregunta definimos como objetivo el predecir el reino al que pertenece un organismo por la frecuencia de uso de codones, utilizando clasificadores de *machine learning*.

Para esta investigación se utilizan las frecuencias de cada una de las 9027 secuencias de codificación proteica completas (CDS) que se han compilado a partir de las divisiones taxonómicas de la base de datos de secuencias de ADN del GenBank. Los archivos de datos se pueden obtener de los sitios ftp anónimos de DDBJ, Kazusa y EBI.

2. ESTADO DEL ARTE

Diversos genes y secuencias de ADN han sido utilizadas para realizar la clasificación taxonómica y filogenética de los organismos, así se utilizan secuencias de los genes ribosomales, rRNA 16S para los organismos procariotas y rRNA 18S para muchos organismos eucariotas, lo cual ha definido la clasificación actual en tres dominios de la naturaleza: bacteria, archaea y eucaria. Sin embargo, el uso de dichas secuencias no responde a la diferenciación para que pueda ser utilizado de forma universal.

Hay antecedentes de uso del *codon usage* como herramienta para predecir y clasificar características genómicas y evolutivas entre los tres dominios de la naturaleza. Por otro lado, también ha resultado útil para predecir el tipo de ADN, como, por ejemplo, ADN nuclear, mitocondrial o de cloroplastos, incluso se propuso su uso al interior de las secuencias de determinadas proteínas, de tal forma que ayuden a predecir anomalías. Utilizando el sesgo de uso de codones, se puede identificar orígenes evolutivos y composición genética, pero se tiene que tomar en cuenta que la alta data disponible dificulta su manipulación, por lo que métodos de clasificación de *machine learning* empiezan a cobrar importancia, como una herramienta que nos permita sistematizar y clasificar los datos genómicos.

Se pueden analizar variaciones intraespecies en cuanto al uso del codón, por ejemplo, se ha reportado que los niveles de producción de proteínas pueden ser predecidos de secuencias del genoma completo utilizando las bases de datos de proteínas ribosomales. No se necesita analizar todas las proteínas, el análisis de un solo grupo puede dar respuestas, sin embargo, hay otros datos que indican que el uso de determinados codones está aún relacionado a la expresión de determinadas proteínas,

lo cual lo complica más, sin embargo, es un modo interesante para ver la expresión de proteínas no solo a nivel de especies sino también a nivel de grupos de proteínas dentro de un mismo organismo (Nakamura et al., 2000)

En el mundo globalizado donde se comparten datos genómicos, se cuenta con bases de datos que compila los codones de proteínas completas, estos datos son utilizados para la anotación de genomas, y a su vez la secuencia de los genomas contiene información importante que puede ser utilizada en la probabilidad de uso como un modo de conseguir no solo clasificar a los organismos, sino identificarlos a través de los sesgos de un código genético degenerado.

3. DISEÑO DEL EXPERIMENTO

3.1 Descripción del conjunto de datos

El conjunto de datos del presente estudio consta de 13 028 datos de organismos, los que registran la frecuencia de usabilidad de cada uno de los codones de ADN, siendo la cantidad de atributos de 69, que detallamos a continuación:

- *Kingdom*: Reino al que pertenece el organismo, con una codificación de 3 caracteres de acuerdo al código CUTG.
- *DANtype*: Tipo o fuente de origen del ADN, expresado con valores numéricos.
- *SpeciesID*: Identificador único en números enteros.
- *Ncodons*: Cantidad de codones de la muestra.
- *SpeciesName*: Nombre descriptivo del organismo.
- *64 Codones*: Frecuencia relativa de los 64 codones de ADN, con 5 cifras significativas.

Podemos apreciar un porcentaje de los datos del estudio en la tabla 1.

Tabla 1

Conjunto de datos del estudio

Kingdom	DNAtype	SpeciesID	Ncodons	SpeciesName	UUU	UGA
vrl	0	100220	1474	Bohle iridovirus	0,02714		0
vrl	0	100755	4862	Sweet potato feaf curl virus	0,01974		0,00144
vrl	0	100880	1915	Northern cereal mosaic virus	0,01775		0

(continúa)

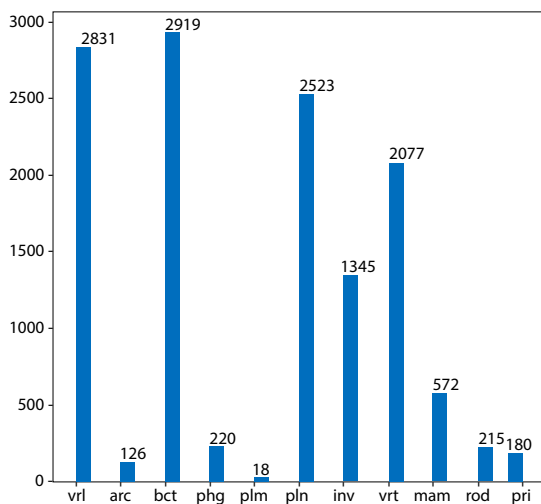
(continuación)

vrl	0	100887	22831	Soil-borne cereal mosaic virus	0,02816	0,00131
vrl	0	101029	5274	Human adenovirus type 7d	0,02579	0,00209
vrl	0	101688	3042	Apple latent spherical virus	0,04635	0
vrl	0	101764	2801	Aconitum latent virus	0,02285	0,00071
vrl	0	101947	2897	Pseudoarabies virus Ea	0,01105	0,00138
vrl	0	10249	61247	Vaccina virus Copenhagen	0,03411	0,00103
vrl	0	10253	55330	Vaccina virus Tian Tan	0,03441	0,00112

Siendo el interés del presente estudio, la clasificación de organismos en reinos, la figura 1, muestra la distribución de los organismos en 11 reinos.

Figura 1

Distribución de organismos por reino



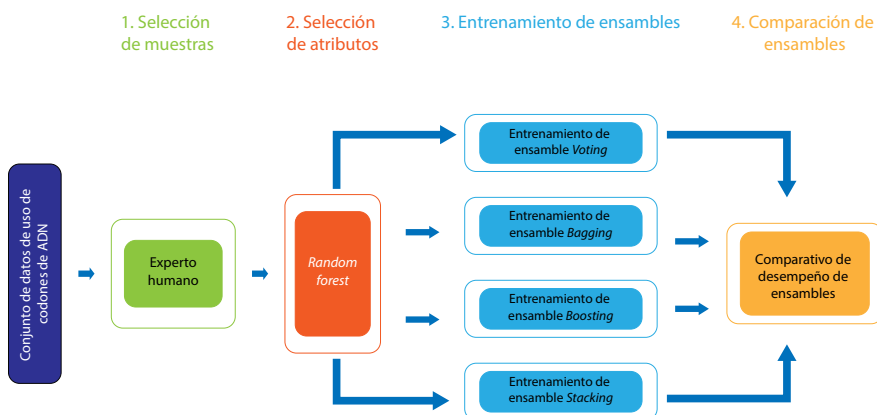
3.2 Metodología

La figura 2 muestra la metodología empleada para el desarrollo del presente estudio, en la primera fase se selecciona un conjunto de muestras del conjunto de datos;

posteriormente, se aplica la selección de atributos utilizando *random forest*; en la tercera, se realiza entrenamiento de cuatro ensambles de clasificación, para finalmente realizar el comparativo de los modelos entrenados.

Figura 2

Metodología del estudio



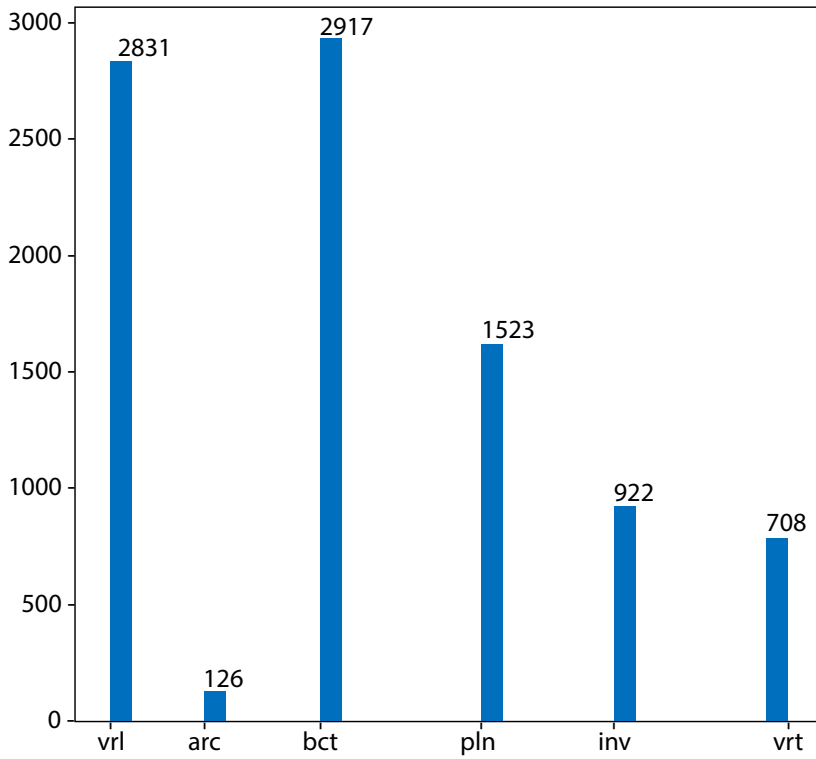
3.2.1 Selección de muestras

El conjunto de datos originales posee 11 clases en el atributo reino, de las cuales dos fueron retiradas: plásmidos (plm) y bacteriófagos (phg), ya que no fueron considerados en la clasificación del presente estudio. Luego se re-etiquetaron mamíferos (mam), primates (pri) y roedores (rod) por vertebrados (vrt). Quedándonos con 6 clases: arqueas (arc), bacterias (bct), invertebrados (inv), vertebrados (vrt), plantas (pln), virus (Vrl). También fueron eliminados los datos que no corresponden al ADN genómico.

Finalmente, la distribución de clases se aprecia en la figura 3, con un total de 9027 datos.

Figura 3

Distribución de reagrupamiento de clases



3.2.2 Cantidad de datos de entrenamiento y test

La cantidad de datos utilizados para el entrenamiento y test se aprecia en la tabla 2.

Tabla 2

Distribución de datos de entrenamiento y test

	Total	Train	Test
Arqueas	126	95	31
Bacterias	2917	2178	739
Invertebrados	922	709	213
Plantas	1523	1133	390
Virus	2831	2110	721
Vertebrados	708	545	162
Σ	9027	6770	2256

3.2.3 Selección manual de atributos

Ya que el propósito del presente estudio es la identificación del reino al que pertenece un organismo, en función al uso de codones, fueron eliminados los atributos: DANtype, SpeciesID, NCodons y SpeciesName, por sugerencia del experto humano.

3.2.4 Selección de métricas de medición de clasificadores

Por tener una distribución de clases desequilibrada, se propone utilizar diferentes métricas: exactitud, precisión, sensibilidad y puntuación F1, las que se calculan por macro promedio:

$$\text{Exactitud} = \frac{VP + VN}{VP + VN + FP + FN} \quad (1)$$

$$\text{Sensibilidad} = \frac{VP}{VP + FN} \quad (2)$$

$$\text{Precisión} = \frac{VP}{VP + FP} \quad (3)$$

$$\text{Puntuación F1} = 2 * \frac{\text{Precisión} * \text{Sensibilidad}}{\text{Precisión} + \text{Sensibilidad}} \quad (4)$$

Donde: *VP* es verdadero positivo; *VN* es verdadero negativo; *FP* es falso positivo; *FN* es falso negativo.

3.2.5 Estrategia de validación y ajuste de hiperparámetros

Para la búsqueda de ajuste de hiperparámetros de los diferentes modelos, se utiliza la búsqueda aleatoria en cuadrícula con validación cruzada, con un máximo de 100 iteraciones y una validación cruzada de 3.

3.2.6 Algoritmos de clasificación

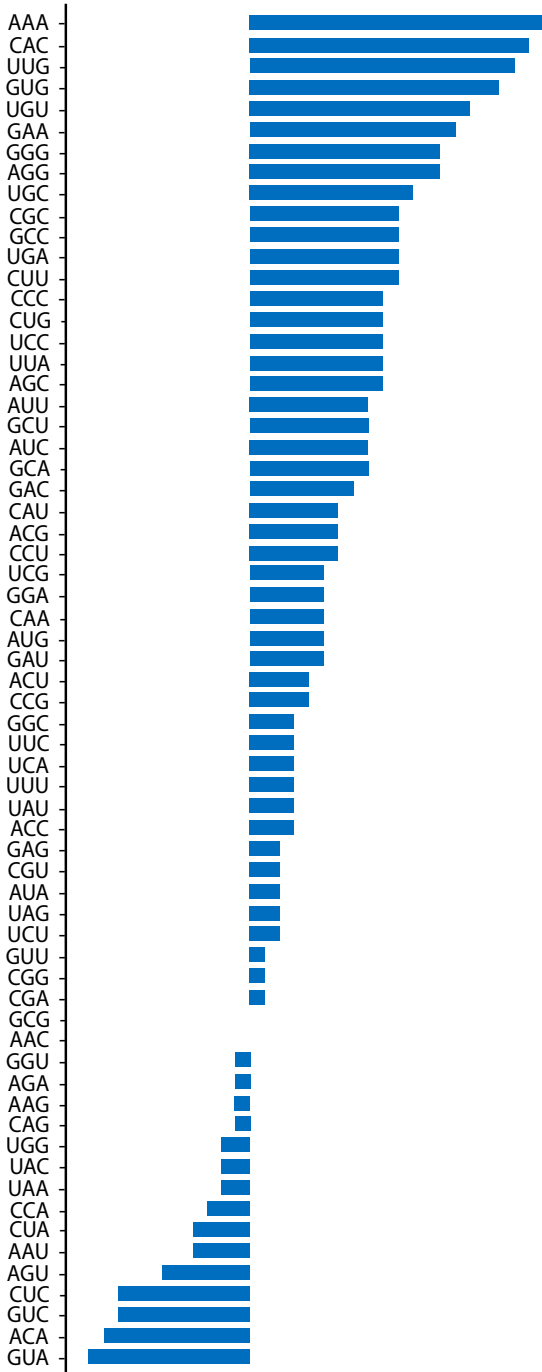
Los algoritmos de clasificación a utilizar corresponden a cuatro arquitecturas de ensamble: *voting*, *bagging*, *boosting* y *stacking*, en las que se hace uso de k-vecinos más cercanos (KNN), árbol de decisión (AD), red neuronal perceptrón multicapa (MLP), máquina de vector de soporte para clasificación (SVC) y ensamble, bosque aleatorio (RF).

3.2.7 Selección de atributos mediante random forest

Se hace uso de bosque aleatorio (*random forest*) para determinar variables predictoras, las que podemos apreciar en la figura 4, esta información es utilizada en los experimentos de diferentes modelos de clasificación.

Figura 4

Importancia de variables en la predicción del reino



4. EXPERIMENTACIÓN Y RESULTADOS

La tabla 3 muestra los mejores modelos encontrados para cada ensamble.

Tabla 3

Ensamblados encontrados en la clasificación de seis reinos a partir de frecuencias de codones de ADN

	Hiperparámetro de ensamble	Modelo	Hiperparámetro de modelo	Exactitud
	pesos de modelos = 1	AD	criterio separación = entropía, prof. árbol = 20, vc =5	
	votación mayoría simple	KNN	k vecinos = 3, métrica distancia = euclidiana, vc =5	
<i>Voting</i>	cv folds = 5	SVC	gamma = 5, kernel = función de base radial, costo = 300, cv = 5	89,46%
	repeticiones = 10	MLP	capas ocultas = (1000,500,200,80,10), alfa = 0.01, fun. activación = tanh, cv =5	
<i>Bagging</i>	cantidad svc = 30 cv folds = 5 repeticiones = 10	SVC	Costo = 1, kernel = función de base radial, degree = 3, gamma = scale,	90,17%
<i>Boosting</i>	cv folds = 5 repeticiones = 10		cant árboles = 90, prof. Árbol = 30, coef. entrenamiento = 0.2	88,29%
	cv folds = 5		k vecinos = 3, métrica distancia = euclidiana	
	repeticiones = 10		prof. árbol = 50, criterio separación = entropía	
<i>Stacking</i>			gamma = 5, costo = 300, kernel = función de base radial	90,97%
			capas ocultas = (500,200,100,50), alfa = 0.01, activación = thanh	
			n.º árboles = 300	

Tabla 4*Presencia de algunos de los datos mal clasificados*

Kingdom	Pred	UUU	UUC	UUA	...	UAA	UAG	UGA
inv	vrt	0,00717	0,01402	0,00163		0,00049	0,00016	0,00033
pln	vrl	0,03234	0,00667	0,029		0,00128	0,00205	0,00051
vrt	inv	0,01131	0,01654	0,00522		0	0,00087	0,00087
arc	bct	0,02566	0,01537	0,01202		0,00266	0,0015	0,00162
inv	vrt	0,01782	0,02858	0,00928		0,00074	0	0,00148
arc	bct	0,00089	0,022	0,00268		0,00089	0,00059	0,00119
vrl	pln	0,02682	0,02778	0,00479		0	0,00096	0,00287
inv	vrt	0,01828	0,02611	0		0,00261	0	0
vrl	bct	0,00488	0,03245	0,00096		0,00188	0,00092	0,00218
pln	bct	0,01427	0,02379	0,00666		0	0,00095	0
bct	pln	0,02274	0,01005	0,01546		0,0022	0,0004	4,00E-05
pln	vrl	0,01279	0,01918	0,0024		0	0,0024	0
bct	vrl	0,01004	0,03161	0		0	0,00355	0,00059
bct	inv	0,01495	0,02415	0,0046		0,00173	0,00115	0,00115

Nota. La tabla muestra algunos de los datos mal clasificados. La columna *Kingdom* registra la etiqueta correcta y la columna *Pred* registra la predicción del mejor modelo.

En la tabla 4 se muestran las predicciones incorrectas, cuya frecuencia relativa se aprecia en la tabla 5.

Tabla 5*Frecuencia relativa de organismos mal clasificados por el mejor modelo encontrado*

	Total de predicciones	Predicciones incorrectas	Frecuencia %
Arqueas	31	5	16,13
Bacterias	739	26	3,52
Invertebrados	213	40	18,78
Plantas	390	8	2,05
Virus	721	17	2,36
Vertebrados	163	16	9,82
Σ	2257	112	

5. DISCUSIÓN DE LA EXPERIMENTACIÓN Y RESULTADOS

Para realizar una aproximación más estricta de la clasificación de los organismos en estudio, se descartaron algunos datos de la matriz inicial, porque no cumplían con la posición taxonómica de reino, otros datos se reclasificaron para conseguir una mejor agrupación y únicamente se trabajó con datos de ADN genómico, para evitar el sesgo de excepciones que tiene el código genético en estructuras intracelulares como los plásmidos, mitocondrias y/o cloroplastos. Para conocer la capacidad predictora de los 64 codones, estos se evaluaron utilizando RF y se determinaron 17 codones que no tenían capacidad predictora por lo que se realizaron ensayos sin utilizarlos (datos no mostrados). Los resultados indican que el uso de estos 17 codones no incrementa mayor predicción en la clasificación de reinos, por lo que se puede prescindir de su uso. Por otro lado, será necesario un análisis más profundo para conocer las razones biológicas de su poca o ninguna capacidad de clasificación en grupos taxonómicos.

Khomtchouk (2020), con la misma base de datos, en el trabajo “Los niveles de sesgo en el uso de codones predicen la identidad taxonómica y la composición genética”, muestra que los niveles de sesgo de uso de codones de un organismo pueden servir como predictor y clasificador de varias características genómicas y evolutivas en los tres dominios de la vida (arqueas, bacterias, eucaria). Teniendo los siguientes resultados:

Tabla 6

Resultados de la clasificación del reino según Khomtchouk

<i>Model</i>	<i>Precision</i>	<i>Recall</i>	<i>Micro F1-Score</i>	<i>Macro F1-Score</i>	<i>Accuracy</i>	<i>AUC</i>
<i>k-Nearest Neighbors</i>	0,9660	1	0,9827	0,9293	0,9669	0,9792
<i>Random Forests</i>	0,9298	1	0,9636	0,8611	0,9298	0,9954
<i>Extreme Gradient Boosting</i>	0,9502	1	0,9745	0,8846	0,9502	0,9970
<i>Artificial Neural Networks</i>	0,9132	1	0,9546	0,8425	0,9132	0,9901
<i>Naive Bayes</i>	0,7200	0,3529	0,4737	0,5487	0,6561	0,8410

Nota. Resultado obtenidos de la clasificación del reino realizada por Khomtchouk (2020)

Donde su mejor clasificador es KNN el cual tiene una precisión de 0,9660 para una clasificación de 3 dominios, lo cual significa su capacidad de discernimiento en una clasificación taxonómica superior a reino. En comparación con nuestra investigación, con el ensamble de *stacking*, tiene una precisión de 0,9097 para la clasificación de 6 reinos, lo cual es muy buena.

A pesar de alcanzar unos índices de precisión y exactitud altos, se observa que con los modelos empleados no logramos la clasificación de la totalidad de los organismos ensayados, han sido 135 organismos que no son correctamente clasificados, una posible explicación es la distribución de datos, que se aprecia en la figura 3, y es la disponibilidad de datos desbalanceada, que podría inducir a los modelos de aprendizaje y aplicación, una predicción incorrecta, probablemente con un mejor balance de distribución de la frecuencia de datos podrían mejorar los resultados obtenidos. Por otro lado, la explicación podría encontrarse en la característica del código genético degenerado, particularidades que deben de ser analizadas con cuidado.

6. CONCLUSIONES

En primer lugar, las técnicas de *machine learning* no responden de forma adecuada a la heterogeneidad de los datos o desbalance en el número de datos. En segundo lugar, se ha comprobado la existencia de 17 codones que tienen menos importancia en el agrupamiento en reinos. Finalmente, se necesita una corrección de excepcionalidades del código genético universal previo a la aplicación de la técnicas de *machine learning*.

7. TRABAJOS FUTUROS

Para trabajos futuros proponemos el uso de *deep learning* y de redes convolucionales, utilizando optimizaciones de *simple gradient descent update* y *mini-batch gradient descent*. Además, el de incrementar datos del reino *arqueas*, utilizando las herramientas de generación de datos sintéticos como DataProf y Iri Rougen, y comprobar estos datos con el algoritmo de generación de datos sintéticos basados en reglas.

REFERENCIAS

- Khomtchouk, B. B. (2020). Codon usage bias levels predict taxonomic identity and genetic composition. *BioRxiv. The Preprint Server for Biology*. <https://doi.org/10.1101/2020.10.26.356295>.
- Im, E.-H., & Choi, S. S. (2017). Synonymous codon usage controls various molecular aspects. *Genomic & Informatics*, 15(4), 123-127. <https://doi.org/10.5808/GI.2017.15.4.123>.

- Nakamura, Y, Gojobori, T, & Ikemura, T. (2000). Codon usage tabulated from international DNA sequence databases: status for the year 2000. *Nucleic Acids Research*, 28(1), 292. <https://doi.org/10.1093/nar/28.1.292>
- Parvathy, S. T., Udayasuriyan, V., & Bhadana, V. (2021). Codon usage bias. *Molecular Biology Reports*, 49, 539-565. <https://doi.org/10.1007/s11033-021-06749-4>
- Sharp, P. M., Emery, L. R., & Zeng, K. (2010). Forces that influence the evolution of codon bias. *Philosophical Transactions of the Royal Society B. Biological Sciences*, 365(1544), 1203-1212. <https://doi.org/10.1098/rstb.2009.0305>
- Wang, F.-P., & Li, H. (2009). Codon-pair usage and genome evolution. *Gene*, 433(1-2), 8-15. <https://doi.org/10.1016/j.gene.2008.12.016>

Recibido: 17/2/2022 / Aceptado: 8/6/2022

doi: <https://doi.org/10.26439/interfases2022.n015.5778>

CIENCIA DE DATOS EN LA EVALUACIÓN DEL IMPACTO DE LAS POLÍTICAS PÚBLICAS: UNA REVISIÓN DE LA LITERATURA

GANIMEDES T. ROSALES REYES

grosalesr@unmsm.edu.pe / ORCID: 0000-0001-6415-6109

Universidad Nacional Mayor de San Marcos (UNMSM), Lima, Perú

Resumen

Lograr que las políticas públicas realmente logren resultados que mejoren la calidad de vida de las personas es un tema que siempre preocupa a los decisores públicos; más aún en un contexto de restricciones presupuestales como el que la mayoría de países atraviesa o atravesará en la situación de pospandemia. Por otro lado, la ciencia de datos, la inteligencia artificial, los datos abiertos y, en general, las tecnologías impulsadas por las grandes cantidades de datos cada día van ganando terreno en el ámbito de la gestión pública. El presente artículo tiene por objetivo conocer el estado actual de las investigaciones acerca de la interacción entre estos dos ámbitos, tomando como objeto de estudio la aplicación de la ciencia de datos a la evaluación del impacto de las políticas públicas, a fin de identificar los vacíos en la investigación existente. Luego de una revisión sistemática de la literatura, se encontró que el objeto de estudio propuesto no ha estado en el centro de la investigación académica. Lo que se ha investigado es el ciclo completo de desarrollo de políticas o la gestión pública en general. Se verificó también que esta interacción entre la cosa pública y la ciencia de datos aún es un campo emergente y, en opinión de muchos académicos, hacen falta investigaciones con una visión holística y con una mirada que vaya más allá de lo eminentemente técnico.

PALABRAS CLAVE: políticas públicas / *big data* / ciencia de datos / evaluación de impacto / modelo

DATA SCIENCE IN THE EVALUATION OF THE IMPACT OF PUBLIC POLICIES: A LITERATURE REVIEW

Abstract

Ensuring that public policies deliver results that improve people's quality of life has always been a central concern of public decision-makers, especially with budgetary restrictions such as the ones most countries are going through or will go through in the post-pandemic context. On the other hand, data science, artificial intelligence, open data, and, generally, technologies driven by large amounts of data are gaining ground in public management. This article aims to determine the current state of research on the interaction between these two disciplines, studying the application of data science to the impact assessment of public policies and identifying research gaps. A systematic literature review revealed that the proposed object of study has not been at the center of academic research. What has been investigated is the complete cycle of policy development or public management in general. The study also verified that the interaction between public affairs and data science is still an emerging field, and, in the opinion of many academics, what is lacking is research with a holistic vision that sees beyond the eminently technical.

KEYWORDS: public policies / big data / data science / impact assessment / model

1. INTRODUCCIÓN

En principio, es preciso aclarar que si bien el objeto de estudio de la presente investigación se enmarca en la evaluación del impacto de las políticas públicas y las tecnologías impulsadas por las grandes cantidades de datos como la ciencia de datos, *big data*, inteligencia artificial (IA) o el aprendizaje máquina, términos que en adelante se usarán, indistintamente, como elementos que la soportan, en la literatura revisada se verifica que en muchos casos se ha abordado el asunto de una manera más amplia, enfocándose en la interacción de estas tecnologías, con la administración pública en general o con el ciclo completo de desarrollo de políticas públicas, e incluso con los esfuerzos que despliegan los países para el logro de los objetivos de desarrollo sostenible (ODS), compromisos que deben ser logrados en el año 2030. En esta sección se describirán cuales son las brechas, sendas o espacios por explorar al respecto, identificados por los autores de la literatura revisada y que, en adelante, denominaremos simplemente como “problemas”, es decir, aquellos que potencialmente se podrían convertir en objetos de estudio para investigaciones futuras.

Varios autores de los artículos revisados han coincidido, en mayor o menor grado, en la identificación de ciertos problemas de investigación y dentro de dichas coincidencias se observa un espectro muy variado en cuanto al alcance de los mismos. Para Ballester (2021), el cual podría ser ubicado en el espectro más amplio de la problemática, existe una ausencia de investigación con una visión panorámica sobre las características del sector público para adoptar las tecnologías habilitadas por la IA; de otro lado, en el extremo más acotado podríamos ubicar los problemas asociados a la incertidumbre generada por los modelos que se utilizan para la toma de decisiones automatizada en el sector público; así para Knüsel et al. (2020) es necesario investigar más acerca de la cuantificación de la incertidumbre de los modelos basados en datos, a las cual denomina “incertidumbre de segundo orden”.

Dada la gran cantidad y variedad de problemática identificada en la literatura revisada, para facilitar su comprensión estos se han agrupado en cuatro temas. El primero de ellos, al que se ha denominado “interacción entre políticas públicas y ciencia de datos”, está relacionado propiamente al objeto de estudio planteado para la presente investigación, es decir, a la relación existente entre las políticas públicas, ya sea como ciclo completo o como alguna de sus etapas, y las tecnologías habilitadas por los grandes datos. El segundo tema denominado “ausencia de una visión holística” está relacionado a la ausencia de una visión amplia e integral de las investigaciones realizadas sobre la administración pública y la ciencia de datos y tecnologías asociadas como habilitador. Para definir el tercer y cuarto tema, se tomaron en cuenta los roles que desempeña el sector público frente a estas tecnologías, de acuerdo con Pi (2021), los sectores públicos tienen un doble papel, como usuarios y como reguladores; en tal sentido, en el tercer tema, al que se le denomina “condiciones a favor y en

contra”, se agruparán los problemas identificados que están relacionados a aquellas barreras o habilitadores que dificultan o facilitan su adopción en el sector público, es decir, cuestiones relacionadas a su rol como usuario. Finalmente, el cuarto tema denominado “impacto en el entorno” agrupará los problemas relacionados a posibles efectos negativos de la adopción de estas tecnologías en el entorno en general, es decir, efectos no deseados del uso de estas tecnologías en la sociedad, medio ambiente, economía, etc., y que el sector público, en su rol de regulador, debe tener en cuenta.

De acuerdo con Zhang et al. (2021), a pesar de que existe una interacción relativamente activa entre las disciplinas de análisis de políticas y ciencia de datos, dicha interacción aún se encuentra en una fase inicial, en tal sentido, es de esperar que haya poca bibliografía que ahonde de manera íntegra dicha interacción; por tal motivo, en el presente estudio se han considerado artículos que aborden tanto el tema desde un punto de vista amplio, sector público en general, así como artículos que han tratado dicha interacción desde un sector en particular del ámbito público.

2. METODOLOGÍA

Para la presente investigación, se plantean las siguientes 4 preguntas, con las cuales se pretende lograr conocer el estado actual de las investigaciones llevadas a cabo sobre la ciencia de datos y su interacción con las políticas públicas:

Pregunta 1: *¿Cuál es la interacción entre políticas públicas y ciencia de datos?*

Pregunta 2: *¿Existe una ausencia de visión holística en la investigación sobre la interacción de ambas disciplinas?*

Pregunta 3: *¿Cuáles son las condiciones a favor y en contra para que el sector público, en general, y la evaluación de impacto de políticas públicas, en particular, obtengan beneficios de la ciencia de datos?*

Pregunta 4: *¿Cuál es el impacto en el entorno que ocasionaría la adopción de la ciencia de datos en el sector público?*

Además, con el propósito de lograr una adecuada recopilación de toda la información relevante que permita tener una idea de la situación actual sobre el conocimiento relacionado a la ciencia de datos y su aplicación a la evaluación de impacto de las políticas públicas, se consultaron las bases de datos Scopus, Science Direct, Web of Science, IEEE, ACM y Ebsco, utilizando como palabras clave de búsqueda las siguientes; “public policies”, “evaluation”, “assessment”, “model”, “data science” y “big data”. Con ese mismo fin se establecieron como criterios de exclusión aquellos documentos que no correspondían a artículos de investigación y aquellos que fueron publicados con anterioridad al año 2017.

Resultado de la búsqueda, se encontraron en total 261 documentos que cumplían con las características descritas y que por lo tanto potencialmente podrían contribuir al estudio; luego de una revisión minuciosa del título, resumen y conclusiones de cada uno de ellos, se eligieron 49 artículos como relevantes y, finalmente, luego de una revisión de la introducción y en algunos casos de las metodologías, se seleccionaron 20 artículos que conforman el universo para la presente investigación.

La base de datos que más contribuye con cantidad de artículos al presente estudio es Scopus, con 9 seleccionados, y en la que no se pudo encontrar ninguno fue Web Of Science; además de los 20 artículos, 16 corresponden a la condición de arbitradas y 4 no arbitradas; respecto al año de publicación, 8 fueron publicados en el año 2021, 5 en el año 2020, 4 en el 2019, 2 en el 2018 y uno en el año 2017. En cuanto a la categoría de cuartiles, 12 de los artículos han sido publicados en revistas categorizadas en el primer cuartil (Q1), 3 tanto en el segundo (Q2) y tercer cuartil (Q3), y uno en el cuarto cuartil (Q4). Cabe mencionar que uno de los artículos seleccionados fue publicado por la *ACM International Conference Proceeding Series*, indexada a Scopus, a la cual no se le ha asignado ningún cuartil a la fecha.

En cuanto a las revistas consultadas, las siguientes 19 fueron las que aportaron los artículos para la presente investigación: *ACM International Conference Proceeding Series, Economic Annals-XXI, EJournal of EDemocracy and Open Government, Environmental Modelling and Software, Government Information Quarterly, Habitat International, Heliyon, IEEE Instrumentation and Measurement Magazine, IEEE Transactions on Engineering Management, International Journal of Information Management, Journal of Cleaner Production, Journal of Data and Information Quality, Knowledge Organization, Policy Sciences, Sustainability (Switzerland), Sustainable Cities and Society, Telecommunications Policy, The Computer Journal y Transportation Research*. La que más aportó fue la revista *Government Information Quarterly* con 2 artículos.

3. RESULTADOS

A continuación, con base a la revisión de literatura realizada se dan respuesta a las cuatro preguntas de investigación planteadas.

3.1 Interacción entre políticas públicas y ciencia de datos

En cuanto a este tema, los investigadores de la literatura revisada han abordado sus investigaciones tomando como objeto de estudio a las políticas públicas en general, así como aquellas políticas públicas en sectores específicos, para nuestro caso en concreto, sector salud, transporte y medio ambiente.

Las investigaciones relacionadas con las políticas en general se han orientado a dilucidar principalmente tres aspectos: a) evaluar el impacto de la inteligencia artificial (IA) en el marco del ciclo de las políticas públicas, b) cómo la ciencia de datos puede contribuir al proceso de definición de políticas públicas, desde un punto de vista sociotécnico, y b) la forma en que la ciencia de datos se ha involucrado en el análisis de políticas.

Para Valle-Cruz et al. (2020), el ciclo de las políticas públicas tradicional se dinamiza a través del conocimiento que se genera impulsado por las tecnologías de las grandes cantidades de datos, sin embargo, indica como necesario llevar a cabo análisis más exhaustivos de los efectos de la IA en las políticas públicas, utilizando diferentes métodos de investigación y considerando aspectos legales, morales y éticos, así como la legitimidad, transparencia, opacidad, discriminación algorítmica y discrecionalidad en la toma de decisiones.

De otro lado, Arnaboldi y Azzone (2020) investigan, desde una perspectiva sociotécnica, la forma en que la ciencia de datos puede contribuir al proceso de definición de políticas públicas y, haciendo uso de la metodología investigación-acción, proponen un procedimiento y un conjunto de roles necesarios para adecuar la demanda y la oferta de datos dentro del proceso de definición de políticas, considerando los aspectos sociales que rodean a la interacción, entre los actores dentro de este proceso, los responsables políticos y los especialistas en análisis de datos; sin embargo sugieren que hace falta más investigación al respecto, reconociendo que el proceso y el resultado de las políticas varían en función de los datos empleados y el nivel de vaguedad o ambigüedad que rodea a la necesidad política, por ello recomiendan estudiar más casos para confirmar sus resultados exploratorios.

Finalmente, Zhang et al. (2021), en su investigación de naturaleza bibliométrica inteligente, busca proporcionar información empírica sobre la participación de la ciencia de datos en el análisis de políticas, concluyendo que se trata de un campo de estudio de interés emergente para la comunidad de ciencia, tecnología e innovación política (STIP, por sus siglas en inglés), pero la literatura que aborda esta interfaz es limitada, lo cual supone un campo fértil por explotar para los investigadores.

Respecto a las investigaciones de políticas públicas y la ciencia de datos en sectores específicos, se observa que han tenido por objetivo proponer nuevos o mejorar modelos existentes para desarrollar políticas en su sector específico. Bassolas et al. (2019), haciendo uso de los registros de teléfonos móviles que alimentan un modelo de transporte basados en actividad (MATSim: Multiagent Transport Simulation), demuestran que es posible tomar mejores decisiones acerca del transporte privado y recomienda explorar, con técnicas similares a la utilizada, el impacto de estas políticas en el transporte público y cómo la demanda puede migrar entre uno u otro tipo de transporte. De otro lado, Anisetti et al., (2018), en su investigación, la cual es una extensión de un trabajo propio anterior, cuyo propósito es sugerir una forma mejorada de desarrollar políticas

de salud pública, considerando un equilibrio entre calidad de vida y la privacidad, en el marco de una ciudad inteligente en Europa, las cuales están afectas al cumplimiento del Reglamento General de Protección de Datos (GDPR, por sus siglas en inglés), proponen el modelo Model-Based Big Data Analytics as a Service (MBDAaaS), el cual según los autores, aumenta la transparencia y reduce la complejidad del diseño y la gestión de las campañas de *big data*. Producto de este trabajo, dichos autores sugieren mayor exploración para cerrar la brecha que existe entre las políticas de salud pública, que toman forma de leyes, reglamentos y directrices que los investigadores proponen y las que finalmente se promulgan y aplican, asimismo indican que para la definición de políticas públicas de salud no solo se pueden tomar datos de las fuentes de soluciones tradicionales, sino incluso de las redes sociales, a fin de integrar la dimensión social en el análisis del escenario de la salud personal, previéndola de una visión holística a la misma; en ese sentido, cobra relevancia realizar mayores investigaciones para asegurar los aspectos éticos de los datos, confidencialidad, anonimato en su análisis para evitar interpretaciones erróneas y conclusiones inadecuadas.

3.2 Ausencia de una visión holística

Los autores han identificado oportunidades de mayor investigación, debido que a su juicio las que se han realizado no tuvieron una visión holística; así, por ejemplo, para Ballester (2021) existe una ausencia de investigación panorámica sobre las particularidades del sector público para utilizar o adoptar las tecnologías habilitadas por la IA. Según este autor, se han realizado investigaciones en su rol de regulador, pero no de usuario, y esto es de suma importancia ya que en el sector público, a diferencia del privado, las innovaciones deben tener muy poco margen para el error debido a las restricciones presupuestales. De otro lado, Y. K. Dwivedi et al. (2021) afirman que muchos de los debates de investigación en torno al sector público y la IA se centran en los aspectos técnicos y el desempeño de la tecnología, por lo que es necesario un debate más amplio que considere las repercusiones culturales y sociales de la tecnología en el contexto de la vida de las personas. Finalmente, del Río Castro et al. (2021), en su búsqueda de las brechas que existen en la investigación sobre si la digitalización y sus paradigmas pueden contribuir a superar las deficiencias en la consecución de los ODS al 2030 por parte de los países, identificaron 7 categorías de brechas por estudiar, entre las cuales se encuentra la existencia de una falta de entendimiento de la naturaleza sistémica e interrelación entre sostenibilidad y digitalización.

3.3 Condiciones a favor y en contra

De acuerdo con Manny et al. (2021), la digitalización en sí misma no suele aportar beneficios inmediatos, por el contrario, produce despilfarro si las grandes cantidades de datos no se comparten, tratan y analizan adecuadamente. En este grupo de problemas,

recopilamos aquellas condiciones o factores que facilitan o dificultan la obtención de beneficios de la digitalización por parte del sector público y que han sido identificados por los investigadores como aspectos que merecen mayor investigación. Cabe señalar que, de la literatura revisada, este es el grupo donde los investigadores ha encontrado un mayor potencial para futuras investigaciones.

Se identificaron 19 problemas que encuadran en este grupo, dada esa cantidad, variedad y, en muchos casos, nivel de especificidad de los mismos, estos se han subcategorizado siguiendo el esquema de análisis de una organización basado en tres elementos: personas, procesos y tecnología.

En la subcategoría relacionada a personas, se identificaron 5 problemas, con mayor y menor grado de amplitud en su planteamiento; así, por ejemplo, desde un punto de vista amplio, Pi (2021) manifiesta una preocupación por el estudio de la transformación cultural de las instituciones públicas hacia una cultura impulsada por la IA y Dwivedi et al. (2021) indican que los dirigentes de la administración pública parecen ser lentos a la hora de reaccionar ante los cambios tecnológicos, lo que demostraría la existencia de un déficit de conocimientos y la necesidad de un cambio cultural en el sector público. Por otro lado, en esta misma subcategoría, pero desde un punto de vista más acotado, del Río Castro et al. (2021) manifiestan que uno de los problemas que merecen mayor investigación son los roles y responsabilidades confusos y la pobre coordinación entre ellos. A su vez, Arnaboldi y Azzone (2020) sostienen que es necesario profundizar en el estudio de las competencias y conocimientos de los 3 roles que identificaron en su investigación, que no son puramente técnicos (traductor, negociador de datos, *scrum master*), a fin de evaluar si el sector público está plenamente preparado para asumir las nuevas funciones requeridas. Finalmente van der Voort et al. (2019) sugieren una investigación más amplia de la dinámica de interacción entre los analistas de datos y los responsables de la toma de decisiones, indicando además que esto se puede entender mejor con las ideas de las ciencias del comportamiento y estudiando el fenómeno en múltiples niveles, tales como el organizativo, el de procesos y el individual.

Respecto a la subcategoría de procesos, los investigadores identificaron 5 problemas relacionados a deficiencias en el diseño de la administración pública para el logro de sus objetivos, monitoreo y rendición de cuentas, así como asimetrías de información y transparencia. Para del Río Castro et al. (2021), en su investigación relacionada a la interrelación que existe entre la sostenibilidad y la digitalización, identificaron una brecha de investigación existente a los problemas de diseño y desequilibrio, argumentando que la agenda de los países es el resultado de un complejo proceso político en lugar de un diálogo basado en la ciencia, lo que lleva a una ambición diluida a la hora de establecer los objetivos finales; además, remarcan que a pesar de la pretendida ambición holística, los objetivos y las metas no están bien conectados, permaneciendo inspirados en una filosofía basada en silos, dominada por la dimensión económica y debilitando la realización

de los objetivos. Asimismo, Dwivedi et al. (2021) plantean que la política tradicional de estrategias a largo plazo, que normalmente optan los gobiernos centrales y los departamentos del sector público, no funciona con tecnologías que cambian rápidamente como la IA, sería mejor si adoptaran una estrategia de planes a corto y medio plazo que pueda ser lo suficientemente flexible como para hacer frente al cambio tecnológico y a los probables avances. En cuanto a monitoreo y rendición de cuentas, mientras que Manny et al. (2021) sugieren que es necesario tener consideraciones en las investigaciones acerca de la rendición de cuentas, ya que la evaluación de datos y la publicación de los resultados de la evaluación pueden aumentar la transparencia y, por tanto, la responsabilidad de las autoridades públicas responsables; del Río Castro et al. (2021) identifican desafíos en el monitoreo y evaluación, debido al vacío de datos, métricas inadecuadas, metodologías de evaluación y rendición de cuentas defectuosas. Por otro lado, van der Voort et al. (2019) indican que, entre los analistas de datos y los responsables de la toma de decisiones, existen asimetrías de información y problemas de transparencia en los procesos de *big data*, que es necesario estudiar desde una perspectiva más allá de la técnica.

Finalmente, respecto a las barreras de naturaleza técnica o tecnológica con los que la administración pública se enfrenta para adoptar la ciencia de datos, se identificaron 9 problemas de investigación, relacionados a la opacidad de los algoritmos de IA, la incertidumbre de estos modelos y dificultades en el tratamiento de las grandes cantidades de datos. Para Pi (2021), muchos de los beneficios en torno a la inteligencia artificial aún no se han materializado en los sectores públicos, debido a su naturaleza opaca y los requisitos de rendición de cuentas de los gobiernos, por lo que se puso en marcha la inteligencia artificial explicable (XAI, por sus siglas en inglés), que es un campo multidisciplinario en la intersección de las ciencias sociales y la inteligencia artificial, pero que la investigación actual tiende a adoptar una visión centrada en los algoritmos, ignorando las necesidades específicas de los usuarios del mundo real. En esa misma línea de problemas relacionados con la poca transparencia de la inteligencia artificial, Dwivedi et al. (2021) sugieren una mayor investigación para responder a las preguntas: ¿qué niveles de garantía algorítmica son necesarios?, ¿cómo pueden los humanos confiar en un enfoque de caja negra de la IA?, ¿cómo podemos saber qué niveles de prueba y escenarios aplicados han sido usados para validar el algoritmo de IA?, ¿qué niveles de soporte tendrían los humanos si las decisiones tomadas por la IA son cuestionables? Por otro lado, respecto a la incertidumbre, Knüsel et al. (2020) recomiendan realizar una mayor investigación sobre los principios que deben tenerse en cuenta para la toma de decisiones en situación de incertidumbre, así como en métodos para la cuantificación de la incertidumbre de segundo orden; en esta misma línea, Astanakulov (2020) también señala que las incertidumbres se consideran factores que afectan negativamente al proceso de modelización, conducen a diversos errores y reducen la calidad de los resultados finales, por lo que estas deben ser investigadas sobre todo con enfoques

cuantitativos. Finalmente, respecto a las dificultades para el tratamiento de las grandes cantidades de datos, Casaccia et al. (2021) señala que existen varios retos de ingeniería acerca de las fuentes de las grandes cantidades de datos que hay que tener en cuenta, especialmente en términos de integrabilidad de los sensores, conectividad y rendimiento metrológico, con el fin de obtener sistemas de medición fiables y precisos. También en ese sentido, Bassolas et al. (2019) indican que el potencial de las nuevas fuentes de datos (de grandes cantidades) conlleva una serie de retos que debe ser investigada; por un lado, se disponen de muestras mucho más grandes que las obtenidas con las encuestas tradicionales y, por otro lado, los datos no se han producido originalmente con el fin de recoger información que se requiere, por lo que a menudo son ruidosos y/o están sesgados; así mismo, Anisetti et al. (2018) plantean la cuestión de cómo gestionar los datos procedentes de los dispositivos IoT. Con una visión un poco más amplia, Shah et al. (2021) manifiestan que los resultados de su investigación, acerca del ecosistema de grandes datos del Gobierno (GBDE, por sus siglas en inglés), requieren de una verificación práctica para asegurar su relevancia para la industria y proponen para ello aplicar sus resultados a una organización del sector público utilizando un estudio de caso, así mismo sugiere que algunas áreas críticas de la GBDE, como, por ejemplo, el ecosistema de *big data* (BDE, por sus siglas en inglés), necesitan una atención especial por parte de la comunidad investigadora pertinente. Sobre este mismo concepto BDE, Victorino et al. (2018), en su investigación, donde proponen un ecosistema de *big data* que gestiona el almacenamiento de datos de diferentes fuentes y formatos para la gestión pública brasileña, y que aseguran proporcionará un valioso apoyo para la evaluación de programas sociales y la gestión de políticas públicas, sugieren que es necesario continuar con las investigaciones para abordar el desafío de dotar a las aplicaciones de la capacidad suficiente para manejar este “entorno de *big data*”, de modo que la información vital y decisiva sea fácilmente accesible.

3.4 Impacto en el entorno

Los problemas que se han identificado en este grupo están relacionados a dos elementos donde las externalidades que ocasionaría la adopción de la IA son negativas: el medio ambiente y la sociedad. Para del Río Castro et al. (2021), en la interacción, digitalización y sostenibilidad, las tecnologías no solamente desempeñan un rol positivo, sino que están acompañadas de problemas aún desconocidos, sobre todo éticos, sociales y medioambientales, que pueden poner en peligro el desarrollo de las tecnologías medioambientales que a su vez ponen en peligro el logro de los ODS. En ese mismo sentido, Kostoska y Kocarev (2019) sostienen que las tecnologías tienen el potencial de aumentar las amenazas a nuestro medio ambiente, la seguridad y la sensación de bienestar, y para mitigar estas amenazas se requiere una transformación de las sociedades más rápida y sustancial que en el pasado. Por otro lado, Dwivedi et al. (2021) sostienen que es necesario seguir investigando para determinar cómo pueden extenderse los

beneficios de la IA al conjunto de la sociedad, teniendo en cuenta que existen pocas pruebas que demuestren que los gobiernos poseen una estrategia tangible o un conocimiento profundo para regular la IA y facilitar que esto ocurra. Del mismo modo, Engin y Treleaven (2019), en su investigación sobre cómo desarrollar sistemas para transformar los servicios públicos, sugieren que deberían llevarse a cabo más debates intersectoriales para considerar otras dimensiones sociales importantes en relación con el modo en que los ciudadanos se relacionan con el Gobierno, para ellos un mayor trabajo cualitativo y tecnológico, en torno a las cuestiones de la equidad, la transparencia, la privacidad y la responsabilidad, sería complementario para construir sistemas integrales en el espacio de las políticas públicas. Finalmente, Kuziemski y Misuraca (2020) indican que a medida que los gobiernos, municipios y los organismos públicos de todo el mundo recurren a la toma de decisiones automatizada en sectores tan diversos como la sanidad, la aplicación de la ley y los servicios sociales —a veces con resultados subóptimos o injustos—, es fundamental considerar en las investigaciones futuras, las direcciones deseadas del desarrollo del campo y examinar las prácticas algorítmicas existentes, planteándose preguntas como ¿qué objetivos deben perseguir las organizaciones del sector público al encargar sistemas de decisión automatizados? y ¿a qué beneficios se debe dar prioridad? Estos autores señalan que a pesar de los trabajos existentes, sobre los sistemas de apoyo a la toma de decisiones y la interacción persona-ordenador (HCI), los dilemas a los que se enfrentan actualmente los responsables políticos distan mucho de ser sencillos y binarios.

4. DISCUSIÓN DE LOS RESULTADOS

De lo encontrado en la literatura revisada, se puede afirmar que el objeto de estudio con el que se inició la presente investigación, es decir, la evaluación de políticas públicas a través de la ciencia de datos, no ha sido tratado de manera específica, sino que las investigaciones se han realizado tomando como objeto de investigación, al ciclo completo de las políticas públicas o a la gestión pública en general. Otros de los aspectos a resaltar, es que todos los autores revisados dan cuenta, en mayor o menor medida, de que la interacción entre la ciencia de datos y la gestión pública, en general, o las políticas públicas, en particular, es un campo emergente que aún no ha sido estudiado lo suficiente y que dicha interfaz necesita ser abordada con una visión holística y considerando aspectos más allá de los netamente técnicos. Una tercera idea, que puede ser muy útil para enmarcar toda la discusión acerca de este tema, es la de considerar al sector público con respecto a la IA en dos roles distintos, la de usuario y la de regulador; si tomamos esto en cuenta, es posible afirmar que, en ambos roles, el sector público enfrenta desafíos de distinta naturaleza y los autores coinciden, en su mayoría, que ha habido más investigaciones en su rol de regulador que como usuario. Finalmente, respecto a los problemas identificados, podemos hacer una síntesis afirmando que existen amplias oportunidades de investigación, ya sea abordando el tema desde un punto de vista holístico, es decir,

tomando al sector público o al ciclo completo de las políticas públicas en su conjunto, o desde un punto de vista más acotado, como el estudio de métodos para reducir la opacidad o desvelar el concepto de caja negra de la IA en la toma de decisiones automatizada en el sector público; en esa misma línea, otro aspecto importante de estudio para investigaciones futuras y, en especial, para la evaluación de políticas públicas son los desafíos que representa el tratamiento de grandes cantidades de datos, ya que cada vez más los gobiernos adoptan políticas de datos abiertos, sin embargo aún no existen marcos o métodos que permitan manejarlos de una manera adecuada, y aunque existen investigaciones en las que se demuestra su aplicación, como la llevada a cabo por Lin et al. (2018) donde verifican que utilizando métodos y marcos de *big data* a los métodos tradicionales, se puede mejorar la capacidad de gestionar los riesgos de inundaciones y evaluar la efectividad de los planes de respuesta en una ciudad China, sin embargo estos estudios son realizados en sectores específicos, quedando pendiente investigar sobre cómo generalizar estas prácticas.

Respecto a las metodologías utilizadas por los investigadores, se observa que una gran mayoría ha utilizado como diseño predominante de su investigación los estudios de caso (encontrados en 11 artículos), luego el análisis de contenido (encontrados en 6 artículos), y en menor frecuencia, los experimentos (encontrados en 2 artículos) e investigación acción (solo en 1 artículo), aunque para ser más exactos, muchas de esas investigaciones utilizaron una combinación de métodos, como el estudio llevado a cabo por Knüsel et al. (2020), quienes con el fin de proponer un marco para la evaluación de la incertidumbre, en primer lugar llevaron a cabo una revisión de contenido de los *frameworks* existentes, para en base a ello proponer un modelo (al que denominaron “ejemplo de juguete”) que luego evaluaron a través de un estudio de caso.

Por otro lado, respecto a los marcos teóricos utilizados, se ha encontrado una gran variedad en cuanto a enfoques y alcance, que es lo que se esperaba dada la variedad en la amplitud y enfoque de los problemas que han abordado los investigadores revisados. Entre los marcos teóricos con un amplio alcance que pueden ser destacados se encuentran: el marco TIC para abordar los ODS, cuyo objetivo es gestionar la complejidad de los ODS y proporcionar conocimientos para construir plataformas de TIC propuestas por Kostoska y Kocarev (2019); el marco propuesto por Ballester (2021), denominado “marco de adopción-clasificación de IA para instancias gubernamentales”, cuyo propósito es proveer de criterios tangibles para evaluar el despliegue de proyectos de IA en el sector público; el modelo de tres niveles (individual, organizativo e institucional) y 4 condiciones (ausencia de una visión, falta de recursos, ausencia de cultura de digitalización y fragmentación administrativa) propuesto por Manny et al. (2021), resulta relevante pues puede ser un punto de partida para la identificación de las barreras que enfrenta el sector público para la adopción de la IA, y finalmente el ciclo dinámico de las políticas públicas (DPPC, por sus siglas en inglés) tratado por Valle-Cruz et al. (2020), quienes

integran la IA al ciclo tradicional y cuyo objetivo es explicar el impacto de estas tecnologías al desarrollo de políticas públicas que se complementarían bastante bien con el ciclo de *big data*, referido por Arnaboldi y Azzone (2020) en su investigación.

Del mismo modo, entre los marcos teóricos asociados a la problemática más específica o de menor alcance podemos citar como relevantes a la inteligencia artificial explicable (XIA, por sus siglas en inglés), tratada en el artículo de Pi (2021), cuyo objetivo es contribuir a reducir la opacidad de los algoritmos IA, el cual según su proponente es en última instancia un problema de interacción persona-ordenador (HCI, por sus siglas en inglés); el marco basado en argumentos, para evaluar la incertidumbre en los modelos basados en datos estudiado por Knüsel et al. (2020), resulta interesante para las investigaciones orientadas a resolver los problemas de incertidumbre de segundo orden; los modelos basados en actividad, que con respecto a los basados en procesos representan ventajas para estimar los efectos de las políticas, que en el caso estudiado por Bassolas et al. (2019) se aplicó al transporte privado, pero que es posible trasladarlo a otros sectores; otro marco conceptual, aunque orientado a aplicaciones específicas del sector salud, resulta relevante para las investigaciones relacionadas a evaluar el impacto de políticas públicas, es el entorno vital inteligente (SLE, por sus siglas en inglés), que representa un marco para diseñar entornos que faciliten las mediciones a través de sensores del bienestar de las personas estudiado por Casaccia et al. (2021); y finalmente en este grupo podemos mencionar, como un marco relevante para investigaciones futuras relacionadas al tema, a los ecosistemas de grandes datos del Gobierno (GBDE) cuyo fin es lidiar con el desafío de tratar con grandes cantidades de datos de diferentes fuentes y formatos, estudiado con amplitud por Shah et al. (2021).

5. CONCLUSIONES

La principal conclusión a la que se llega en la presente investigación es que la evaluación de políticas públicas mediante IA no ha sido tratada de manera específica por los investigadores de la literatura revisada, sino más bien se observa que los estudios se han enfocado ya sea a todo el ciclo de desarrollo de políticas o al sector público en general. También se encontraron estudios sobre políticas públicas en sectores específicos, que es donde se ha hecho mayor referencia a la evaluación de impacto de las mismas, pero no como un tema central sino como uno complementario, a la mejora del desarrollo de las mismas impulsada por IA. Esta evidencia podría estar sugiriendo que existen dificultades para un estudio general, en lugar de aplicaciones a sectores específicos, sobre la evaluación de políticas públicas y ciencia de datos, las cuales pueden estar relacionadas, a su vez, a una ausencia de marcos teóricos sólidos que permitan estudiar la evaluación de impacto de políticas con una visión amplia e integrada del sector público. De otro lado, reconociendo que el presente estudio tuvo limitaciones que pueden haber ocasionado

la omisión de artículos de relevancia, se sugiere una mayor investigación de literatura, utilizando palabras clave como inteligencia artificial, aprendizaje maquina, entre otras, y métodos bibliométricos inteligentes, para confirmar que los estudios sobre evaluación de políticas públicas solo se han dado en sectores específicos del ámbito público y tratar de dilucidar por qué los mismos no se han podido generalizar.

REFERENCIAS

- Anisetti, M., Ardagna, C., Bellandi, V., Cremonini, M., Frati, F., & Damiani, E. (2018). Privacy-aware Big Data Analytics as a service for public health policies in smart cities. *Sustainable Cities and Society*, 39, 68-77. <https://doi.org/10.1016/J.SCS.2017.12.019>
- Arnaboldi, M., & Azzone, G. (2020). Data science in the design of public policies: dispelling the obscurity in matching policy demand and data offer. *Heliyon*, 6(6), e04300. <https://doi.org/10.1016/J.HELIYON.2020.E04300>
- Astanakulov, O. (2020). National projects and government programmes: Functional algorithm for evaluating and modelling using the Data Science methodology. *Economic Annals-XXI*, 183(5-6), 51-59. <https://doi.org/10.21003/EA.V183-05>
- Ballester, O. (2021). An artificial intelligence definition and classification framework for public sector applications. *ACM International Conference Proceeding Series*, 67-75. <https://doi.org/10.1145/3463677.3463709>
- Bassolas, A., Ramasco, J. J., Herranz, R., & Cantú-Ros, O. G. (2019). Mobile phone records to feed activity-based travel demand models: MATSim for studying a cordon toll policy in Barcelona. *Transportation Research Part A: Policy and Practice*, 121, 56-74. <https://doi.org/10.1016/J.TRA.2018.12.024>
- Casaccia, S., Revel, G. M., Cosoli, G., & Scalise, L. (2021). Assessment of domestic well-being: from perception to measurement. *IEEE Instrumentation and Measurement Magazine*, 24(6), 58-67. <https://doi.org/10.1109/MIM.2021.9513641>
- del Río Castro, G., González Fernández, M. C., & Uruburu Colsa, Á. (2021). Unleashing the convergence amid digitalization and sustainability towards pursuing the Sustainable Development Goals (SDGs): A holistic review. *Journal of Cleaner Production*, 280(Part 1). <https://doi.org/10.1016/J.JCLEPRO.2020.122204>
- Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., Duan, Y., Dwivedi, R., Edwards, J., Eirug, A., Galanos, V., Ilavarasan, P. V., Janssen, M., Jones, P., Kar, A. K., Kizgin, H., Kronemann, B., Lal, B., Lucini, B., ... Williams, M. D. (2021). Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and

- policy. *International Journal of Information Management*, 57, 101994. <https://doi.org/10.1016/J.IJINFOMGT.2019.08.002>
- Engin, Z., & Treleaven, P. (2019). Algorithmic government: Automating public services and supporting civil servants in using data science technologies. *The Computer Journal*, 62(3), 448-460. <https://doi.org/10.1093/COMJNL/BXY082>
- Knüsel, B., Baumberger, C., Zumwald, M., Bresch, D. N., & Knutti, R. (2020). Argument-based assessment of predictive uncertainty of data-driven environmental models. *Environmental Modelling and Software*, 134. <https://doi.org/10.1016/J.ENVSOFT.2020.104754>
- Kostoska, O., & Kocarev, L. (2019). A novel ICT framework for sustainable development goals. *Sustainability*, 11(7). <https://doi.org/10.3390/SU11071961>
- Kuziemski, M., & Misuraca, G. (2020). AI governance in the public sector: Three tales from the frontiers of automated decision-making in democratic settings. *Telecommunications Policy*, 44(6). <https://doi.org/10.1016/j.telpol.2020.101976>
- Lin, T., Liu, X., Song, J., Zhang, G., Jia, Y., Tu, Z., Zheng, Z., & Liu, C. (2018). Urban waterlogging risk assessment based on internet open data: A case study in China. *Habitat International*, 71, 88-96. <https://doi.org/10.1016/j.habitatint.2017.11.013>
- Manny, L., Duygan, M., Fischer, M., & Rieckermann, J. (2021). Barriers to the digital transformation of infrastructure sectors. *Policy Sciences*, 54, 943-983. <https://doi.org/10.1007/S11077-021-09438-Y>
- Pi, Y. (2021). Machine learning in Governments: Benefits, challenges and future directions. *EJournal of EDemocracy and Open Government*, 13(1), 203-219. <https://doi.org/10.29379/JEDEM.V13I1.625>
- Shah, S. I. H., Peristeras, V., & Magnisalis, I. (2021). Government Big Data Ecosystem: Definitions, types of data, actors, and roles and the impact in public administrations. *Journal of Data and Information Quality*, 13(2), 1-25. <https://doi.org/10.1145/3425709>
- Valle-Cruz, D., Criado, J. I., Sandoval-Almazán, R., & Ruvalcaba-Gomez, E. A. (2020). Assessing the public policy-cycle framework in the age of artificial intelligence: From agenda-setting to policy evaluation. *Government Information Quarterly*, 37(4). <https://doi.org/10.1016/J.GIQ.2020.101509>
- van der Voort, H. G., Klievink, A. J., Arnaboldi, M., & Meijer, A. J. (2019). Rationality and politics of algorithms. Will the promise of big data survive the dynamics of public decision making? *Government Information Quarterly*, 36(1), 27-38. <https://doi.org/10.1016/J.GIQ.2018.10.011>

- Victorino, M., de Holanda, M. T., Ishikawa, E., Oliveira, E. C., & Chhetri, S. (2018). Transforming open data to linked open data using ontologies for information organization in big data environments of the Brazilian Government: The Brazilian Database Government Open Linked Data DBgoldbr. *Knowledge Organization*, 45(6), 443-466. <https://doi.org/10.5771/0943-7444-2018-6-443>
- Zhang, Y., Porter, A. L., Cunningham, S., Chiavetta, D., & Newman, N. (2021). Parallel or intersecting lines? Intelligent bibliometrics for investigating the involvement of data science in policy analysis. *IEEE Transactions on Engineering Management*, 68(5), 1259-1271. <https://doi.org/10.1109/TEM.2020.2974761>

Recibido: 16/5/2022 / Aceptado: 2/7/2022

doi: <https://doi.org/10.26439/interfases2022.n015.5898>

APLICACIÓN DE *MACHINE LEARNING* EN LA GESTIÓN DE RIESGO DE CRÉDITO FINANCIERO: UNA REVISIÓN SISTEMÁTICA

JULER ANDERSON HERMITAÑO CASTRO

juleranderson@gmail.com / orcid: 0000-0001-6580-880X

Universidad de Lima, Perú

Resumen

La gestión de riesgos bancarios puede ser dividida en las siguientes tipologías: riesgo crediticio, riesgo de mercado, riesgo operativo y riesgo de liquidez, siendo el primero el tipo de riesgo más importante para el sector financiero. El presente artículo tiene como objetivo mostrar las ventajas y desventajas de la implementación de los algoritmos de *machine learning* en la gestión de riesgos de crédito y, a partir de esto, mostrar cuál tiene mejor rendimiento, señalando también las desventajas que puedan presentar. Para ello se realizó una revisión sistemática de la literatura con la estrategia de búsqueda PICO y se seleccionaron doce artículos. Los resultados reflejan que el riesgo de crédito es el de mayor relevancia. Además, algunos de los algoritmos de *machine learning* ya han comenzado a implementarse, sin embargo, algunos presentan desventajas resaltantes como el no poder explicar el funcionamiento del modelo y ser considerados como caja negra. En ese sentido, desfavorece la implementación debido a que los organismos regulatorios exigen que un modelo deba ser explicable, interpretable y transparente. Frente a ello, se ha optado por realizar modelos híbridos con algoritmos que no son sencillos de explicar, como aquellos modelos tradicionales de regresión logística. También, se presenta como alternativa utilizar métodos como SHAPley Additive exPlanations (SHAP) que ayudan a la interpretación de dichos modelos.

PALABRAS CLAVE: *machine learning* / ML / gestión, riesgo / crédito / algoritmo

Application of Machine Learning in Financial Credit Risk Management: A systematic review

Abstract

Banking risk management can be divided into the following typology: credit risk, market risk, operational risk, and liquidity risk, the first being the most important type of risk for the financial sector. This article aims to show the advantages and disadvantages of implementing Machine Learning algorithms in credit risk management. A systematic literature review was carried out with the PICO search strategy, and 12 articles were selected. The results show that credit risk is the most relevant. In addition, some of the Machine Learning algorithms have already begun to be implemented; however, some have significant disadvantages, such as not being able to explain the model's operation and are considered a black box. In this sense, it discourages implementation because regulatory bodies require that a model be explainable, interpretable and transparent. Faced with this, it has been decided to make hybrid models between algorithms that are not easy to explain with traditional ones, such as logistic regression. Also, it is presented as an alternative to using methods such as SHAPley Additive exPlanations (SHAP) that help the interpretation of these models

KEYWORDS: machine learning / ML / management / risk / credit / algorithm

1. INTRODUCCIÓN

El *machine learning* (ML) es una rama de la inteligencia artificial que puede ser usada en múltiples áreas de investigación académica, desarrollo tecnológico y predicción de datos para empresas (Enterprise.nxt Staff, 2021), ya que su proceso consiste en crear un modelo algorítmico que permite ingresar datos históricos, hacer un entrenamiento según el enfoque que se requiera y, a partir de este, encontrar patrones para predecir alguna información futura. Esto genera que las empresas pongan énfasis en el desarrollo del *machine learning*, pero es importante que se identifique cuáles son los objetivos que se quieren lograr con su implementación (SAS: Analytics, Artificial Intelligence and Data Management, 2017). Según una publicación en *Forbes* (2021), el 43 % de empresas ha detectado que la implementación de estos algoritmos ha sido más útil de lo que creían inicialmente y el 50 % ha planeado invertir más en estos modelos el presente año, lo cual empieza a demostrar que ponerlos en funcionamiento es de gran importancia en cualquier sector. Un sector que importa y que genera el movimiento en la economía, es el financiero, el cual lleva implementándolo para mejorar la experiencia al usuario a través de soluciones personalizadas o *chatbots* como destacó *Forbes* a la BBVA Corporation (2019). Por otro lado, una actividad que es de gran importancia para cualquier organización o entidad financiera es la gestión de riesgos que puede ser dividida según su taxonomía en: riesgo de seguros y demográfico, riesgo de mercado, riesgo de crédito y riesgo operacional (Mashrur et al., 2020). Al que más enfoque se ha prestado es a los riesgos crediticios para poder mitigar las pérdidas ya sea un banco o una fintech dedicada a préstamos, ya que en el último Risk Dashboard Q1 2021 de la European Banking Authority (EBA) se enfatizó que los bancos seguían siendo vulnerables a los movimientos adversos del riesgo crediticio (KPMG Company, 2021). Por este motivo, se percibe un cambio en el paradigma de no solo optar por los modelos estadísticos tradicionales, sino que cada vez procuren ser más sofisticados como el *machine learning* para cuantificar y mitigar el riesgo de manera correcta (Mashrur et al., 2020).

Por esta razón, es importante realizar una revisión sistemática de la literatura de investigaciones previas, que permita identificar y centralizar la información de qué algoritmos de *machine learning* están implementándose dentro de la gestión de riesgos de crédito. Por lo tanto, el objetivo de la presente investigación es mostrar las soluciones que se han propuesto y determinar las ventajas y desventajas que presentan cada uno de estos, ya sea de los modelos tradicionales o algoritmos de *machine learning*. Asimismo, seguir como modelo una revisión previa, como la encontrada en el artículo de Bhatore, Mohan y Reddy, (2020).

Este documento está estructurado de la siguiente manera: la sección 2 explica los conceptos previos del estudio, la sección 3 describe la metodología de investigación que se empleó para obtener la información, la sección 4 presenta los resultados del

estudio, respondiendo a las preguntas planteadas en la sección previa, y, finalmente, la sección 5 muestra la discusión, conclusiones y trabajos futuros.

2. CONCEPTOS Y TÉCNICAS DE MACHINE LEARNING

2.1 Inteligencia artificial y *machine learning*

El primer concepto se refiere a que un sistema o máquina tiene como objetivo imitar la inteligencia humana al realizar alguna función específica y mejorar conforme se recopilen datos. Asimismo, el *machine learning* es una parte de la IA que permite lograr ese objetivo deseado, a través de un aprendizaje de los datos y no una programación explícita (IBM, 2018). El *machine learning* es un subconjunto de la IA, que se puede asociar con una metáfora que consiste en que, “una vez resuelto el problema, es capaz de reconocer la situación problemática y reaccionar usando la estrategia aprendida” (Moreno et al., 1994, p. 6). Los paradigmas que se han planteado para dividir han variado a lo largo del tiempo. Por ejemplo, Moreno et al. señalaban que se podrían dividir entre aprendizaje deductivo, aprendizaje analítico, aprendizaje analógico, aprendizaje inductivo, aprendizaje mediante descubrimiento, entre otros. Sin embargo, como se mencionó, actualmente se ha categorizado de la misma manera por la forma en que aprende el algoritmo: aprendizaje supervisado, aprendizaje no supervisado, aprendizaje semisupervisado, aprendizaje por reforzamiento (Zhang, 2010). La explosión en la investigación que ha tenido el ML, desde los años noventa, se debe a causas como, en primer lugar, “las comunidades de investigación separadas en el aprendizaje automático simbólico, la teoría del aprendizaje computacional, las redes neuronales, las estadísticas y el reconocimiento de patrones se han descubierto entre sí y han comenzado a trabajar juntas” (Dietterich, 1997). En segundo lugar, “las técnicas de aprendizaje automático se están aplicando a nuevos tipos de problemas, incluido el descubrimiento de conocimientos en bases de datos, procesamiento de lenguaje, control de robots y optimización combinatoria, así como a problemas más tradicionales como el reconocimiento de voz, reconocimiento facial, reconocimiento de escritura a mano, médicos, análisis de datos y juego” (Dietterich, 1997).

2.2 Aprendizaje supervisado

Este conjunto de algoritmos utiliza variables etiquetadas para proceder con un entrenamiento. A través de esto, el modelo puede ser capaz de realizar predicciones de ejemplos sin etiquetar. Principalmente, se asocia con problemas de clasificación, regresión y “*ranking problem*” (Mashrur et al., 2020). Algunos tipos de algoritmos que utilizan el aprendizaje supervisado son regresión lineal, regresión logística, *k-nearest neighbor* (KNN), *support vector machine* (SVM), *random forest*, *árboles de decisión* y Naive Bayes (Cunningham et al., 2008). Cada uno de estos presentan un grado de dificultad distinto y su utilización depende del contexto en que se necesite.

2.3 Aprendizaje no supervisado

Este conjunto de algoritmos tiene como principal tarea detectar patrones a partir de datos no etiquetados. En el entrenamiento no existen datos etiquetados disponibles y el principal enfoque que se les otorga es resolver problemas de “*clustering*”, detección de valores atípicos, reducción de dimensionalidad y detección de anomalías (Mashrur et al., 2020). Algunos algoritmos que usan este tipo de aprendizaje son el análisis factorial, *principal component analysis* (PCA), mezclas de gaussianos, ICA, modelos de Markov ocultos, modelos de espacio de estado (Ghahramani, 2004). Asimismo, *K-means*, *mean shift*, *K-medoids*.

2.4 Aprendizaje profundo (*deep learning*)

Este tipo de aprendizaje se encuentra dentro del *machine learning* que tiene un campo de estudio muy amplio, debido a la complejidad de sus modelos y que muchas veces es representado como “caja negra” que extrae representaciones de diferentes capas de características. Se aproxima a una función compuesta no lineal que forma una transformación jerárquica de características en etiquetas (Mashrur et al., 2020).

3. METODOLOGÍA

Para la realización del presente trabajo, se adoptó una revisión sistemática de la literatura (SRL, por sus siglas en inglés). Se optó por realizar una metodología similar a la propuesta por Nina et al. (2021), la cual consta de una planificación para identificar el interés, la población y el contexto actual, el planteamiento de preguntas relevantes para la revisión, una búsqueda de artículos por palabras clave y una construcción de cadenas de búsqueda, y, finalmente, una filtración y selección de artículos. Los pasos detallados son los siguientes:

3.1 Planificación de estudio

En primer lugar, se requirió realizar una planificación del estudio utilizando la técnica PICO:

- Población: Gestión de riesgos de crédito
- Interés: Aplicación de *machine learning* en el riesgo de crédito
- Contexto: Sector bancario, fintech o financieras, plataformas de préstamos
- Como interés de la presente revisión, se lograron identificar 4 preguntas:
- ¿Qué relevancia presenta el riesgo de crédito dentro de la gestión de riesgos en el sector financiero?

- ¿Cómo el *machine learning* mejora la predicción del riesgo de crédito financiero frente a métodos tradicionales?
- ¿Qué retos presenta el *machine learning* frente a los métodos tradicionales?
- ¿Cómo identificar el algoritmo de *machine learning* más eficaz para la gestión de riesgo de crédito financiero?

3.2 Búsqueda de artículos

En segundo lugar, se utilizó la base de datos de bibliografía indexada Scopus, la cual es frecuente para estudios de revisión sistemática de la literatura. Dado que cuenta con una herramienta que permite realizar búsquedas avanzadas, se optó por identificar palabras claves que permitan realizar una cadena de búsqueda y así realizar un filtrado de artículos. Todas las palabras fueron traducidas al inglés, debido al mayor alcance que este idioma permite.

- Palabras clave: *risk management, credit, financiera, bank, fintech, lending, machine learning, artificial intelligence*
- Cadena de búsqueda: se generaron a partir de conectores lógicos como *OR* para los sinónimos y *AND* para la combinación de palabras clave. La cadena que se ingresó fue la siguiente: TITLE-ABS (“RISK MANAGEMENT” AND “CREDIT” AND “FINANCIAL” AND (“BANK” OR “FINTECH” OR “LENDING PLATFORM”) AND (“MACHINE LEARNING” OR “ML” OR “ARTIFICIAL INTELLIGENCE” OR “AI”)) AND (LIMIT-TO (PUBYEAR, 2022) OR LIMIT-TO (PUBYEAR, 2021) OR LIMIT-TO (PUBYEAR, 2020) OR LIMIT-TO (PUBYEAR, 2019)). Con esta cadena, se obtuvieron 27 resultados.
- Filtrado de artículos: el primer filtrado se realizó en la cadena de búsqueda de acuerdo con los años, en la que se permitieron artículos publicados desde el 2019. Luego, se prosiguió con la lectura de los títulos y *abstracts* de cada artículo.

3.3 Selección de artículos

En tercer lugar, se realizaron criterios de inclusión y exclusión:

- Criterios de inclusión
 - Los artículos científicos informan sobre el uso de *machine learning* en la gestión de riesgos en el rubro financiero.
 - Los artículos científicos fueron revisados por pares.
 - Los artículos científicos fueron publicados en revistas de cuartiles Q1, Q2 o Q3.

- Los artículos científicos fueron publicados en entre los años 2019 y 2022.
- Criterios de exclusión
 - Los artículos científicos no informan sobre el uso del *machine learning* en la gestión de riesgos financieros.
 - Los artículos científicos no fueron revisados por pares.
 - Los artículos científicos son de publicaciones anteriores al 2019.

Luego de aplicar estos criterios, se analizaron 12 artículos y 1 capítulo de libro.

Tabla 1

Artículos revisados para el análisis de las preguntas

Item	Año	Nombre	Autores	Journal
1	2022	Machine learning for credit scoring: Improving logistic regression with non-linear decision-tree effects	Dumitrescu, E., Hué, S., Hurlin, C., & Tokpavi, S.	<i>European Journal of Operational Research</i>
2	2022	Financial risk management and explainable, trustworthy, responsible AI	Fritz-Morgenthal, S., Hein, B. & Papenbrock, J.	<i>Frontiers in Artificial Intelligence</i>
3	2021	Prospects of artificial intelligence and machine learning application in banking risk management	Milojević, N., & Redzepagic, S	<i>Journal of Central Banking Theory and Practice</i>
4	2021	Explainable machine learning in credit risk management	Bussmann, N., Giudici, P., Marinelli, D., & Papenbrock, J.	<i>Computational Economics</i>
5	2021	Classical machine learning methods in economics research: Macro and micro level examples	Babenko, V., Panchyshyn, A., Zomchak, L., Nehrey, M., Artym-Drohomyretska, Z., & Lahotskyi, T.	<i>WSEAS Transactions on Business and Economics</i>
6	2021	Research on Credit Risk Control of Commercial Banks Based on Data Mining Technology	Cui, H	<i>IEEE CBFDF Conference</i>
7	2021	Comparative Analysis of Machine Learning Models on Loan Risk Analysis	Srinivasa Rao, Sekhar, & Bhattacharyya	<i>Advances in Intelligent Systems and Computing</i>

(continúa)

(continuación)

8	2020	Machine learning for financial risk management: A survey	Mashrur, A., Luo, W., Zaidi, N., & Robles-Kelly, A.	<i>IEEE Access</i>
9	2020	Explainable AI in fintech risk management	Bussmann, N., Giudici, P., Marinelli, D., & Papenbrock, J.	<i>Frontiers in Artificial Intelligence</i>
10	2020	Risk attribution and interconnectedness in the EU via CDS data	Giacometti, Torri, Farina, & De Giuli	<i>Computational Management Science</i>
11	2020	DGHNL: A new deep genetic hierarchical network of learners for prediction of credit scoring	Plawiak, Abdar, Plawiak, Makarenkov, & Acharya	<i>Information Sciences</i>
12	2020	Estimation procedures of using five alternative machine learning methods for predicting credit card default	Teng, H-W. & Lee, M.	[Libro] <i>Handbook of financial econometrics, mathematics, statistics, and machine learning</i>
13	2019	Machine learning in banking risk management: A literature review	Leo, M., Sharma, S., & Maddulety, K.	<i>Risks</i>

4. RESULTADOS

La presente sección responde a cada una de las preguntas planteadas previamente en la primera sección de la metodología.

4.1 ¿Qué relevancia presenta el riesgo de crédito dentro de la gestión de riesgos en el sector financiero?

El riesgo crediticio es el mayor riesgo al que se puede enfrentar un banco, ya que evita pérdidas y maximiza ganancias, lo cual es vital para cualquier empresa (Mashrur et al., 2020). Asimismo, es importante determinar el riesgo de crédito desde una perspectiva macroprudencial considerando al sector bancario como un sistema complejo y no como instituciones individuales (Giacometti et al., 2020).

La gestión de riesgos financieros (FRM) puede ser dividida en gestión de riesgos demográficos y de seguros, gestión de riesgo crediticio, gestión de riesgo de mercado, y gestión de riesgo operacional (Mashrur et al., 2020). Sin embargo, el desarrollo e implementación de *machine learning* se ha priorizado en el segundo tipo de gestión mencionado, debido a que es el más importante. También, se percibe a través de la

mención de su importancia en 7 de los 12 papers consultados en la revisión, tales como los ítems 3, 6, 7, 8, 10, 11 y 13 de la tabla 1. Tal como se señala Cui (2021) y Srinivasa et al. (2021) la gestión del riesgo de crédito en los bancos es crucial para evitar el riesgo de perder dinero debido a los malos préstamos.

Como se sabe, el riesgo de crédito es la incertidumbre que involucra la capacidad de cumplir cierta obligación monetaria de algún prestatario. El prestatario puede ser una persona u organización (Mashrur et al., 2020). Por lo tanto, es necesario poder realizar un análisis que permita conocer al prestador, el nivel de pérdidas que puede llegar a tener en caso ocurriera algún incumplimiento de los acreditados. Este puede ser determinado con un factor importante como es la calificación crediticia, ya que es crucial para el análisis del préstamo, pues permite identificar a quiénes se debería otorgar el crédito (Plawiak et al., 2020). Además, otro objetivo de esta medición es la estimación de determinar las pérdidas esperadas y no esperadas (Elizondo, 2003). Como se detalla en el libro *Medición integral del riesgo de crédito*, “representa el monto de capital que podría perder una institución, como resultado de la exposición al riesgo de crédito” (Elizondo, 2003, p. 50). De acuerdo a los Acuerdos de Basilea III, las técnicas para la gestión del riesgo de crédito pueden ser tipo estandarizadas o basadas en calificaciones internas. Estas últimas se calculan mediante la multiplicación de probabilidad de incumplimiento (PD), pérdida en caso de incumplimiento (LGD) y la exposición al incumplimiento (EAD) (Superintendencia de Banca, Seguros y AFP, s. f.). Entonces, se permite que cada entidad pueda desarrollar sus propios modelos de riesgo con los parámetros de riesgo clave lo cual puede resultar en una pérdida esperada (EL) con la multiplicación de estos tres (Bank for International Settlements, 2005). Asimismo, si no se cuenta con una buena práctica de la gestión de riesgos de créditos, cualquier entidad corre riesgo, ya que es una de las principales oportunidades para mejorar el desempeño general de la organización y asegurar una ventaja competitiva frente a los demás (SAS: Analytics, Artificial Intelligence and Data Management, 2015).

4.2 ¿Cómo el *machine learning* mejora la predicción del riesgo de crédito frente a métodos tradicionales?

Los algoritmos clasificadores de *machine learning* funcionan de forma significativamente más precisa que los métodos estadísticos estándar en la calificación crediticia, lo cual ofrece más valor a la gestión de riesgos.

Dentro de la gestión de riesgo crediticio, inicialmente las decisiones de admisión de créditos se tomaban con evaluaciones subjetivas, dependiendo de la experiencia. Luego fue se reemplazó por métodos estadísticos tradicionales, de los cuales el más usado ha sido el análisis discriminante lineal (LDA) o también regresión logística para determinar la probabilidad de incumplimiento (Leo et al., 2020). Sin embargo, en los

últimos años se ha implementado dos tipos de métodos de aprendizaje automático: supervisado y no supervisado. De los primeros, se utilizan los clasificadores únicos, de los cuales se señala que muchos estudios han demostrado que hay dos que se pueden utilizar eficazmente para predecir la quiebra o calificar el crédito, los cuales son las *support vector machine* (SVM) y las redes neuronales. De los segundos, el método que se usa es de agrupación y se puede utilizar para identificar el riesgo de quiebra o incumplimiento crediticio, además ayuda a identificar grupos de solicitantes de préstamos y ofrece un modelo de puntuación dinámico basado en conglomerados, el cual mejora la precisión de puntuación (Mashrur et al., 2020). Además, Dumitrescu et al. (2021) agregan que los métodos de árboles de decisión y *random forest* proporcionan un mejor rendimiento de clasificación que los modelos de regresión logística estándar. Como mencionan Leo et al., (2020), en el enfoque de *machine learning*, el algoritmo *support vector machine*, al igual que otros, tiene éxito al clasificar a los clientes que incumplen e incluso son competitivos en el descubrimiento de características significantes para determinar el riesgo de incumplimiento. En adición a esto, se pueden implementar los algoritmos de ML a los métodos cotidianos, tal como lo realizaron Dumitrescu et al. (2021), en el que combinan la regresión logística tradicional con el algoritmo de árboles de decisión. Esto último logra transmitir que no se opte solo por algún método, sino que los modelos tradicionales pueden trabajar correctamente con el ML.

4.3 ¿Qué retos presenta el *machine learning* frente a los métodos tradicionales?

Actualmente, existen inconvenientes con la adopción de estos nuevos modelos, ya que, al inicio, a algunos de estos se les solía etiquetar como “caja negra”, término que se refería a que era difícil poder explicar el funcionamiento del modelo y cómo lograr el resultado que se podía obtener. En ese sentido, las entidades regulatorias no permiten la implementación de este tipo de modelos, ya que se requiere explicabilidad, interpretabilidad y transparencia de los modelos (Fritz-Morgenthal et al., 2022).

El término “caja negra” se le atribuye a la tecnología, ya que su auge ha sido exponencial y cada vez más compleja. Especialmente, se les etiqueta este término a los algoritmos de aprendizaje profundo (*deep learning*), ya que resulta el más complejo de entender (Castelvecchi, 2016), pero ciertos algoritmos de *machine learning* también pueden contener esta característica. Además, Mashrur et al. (2020) contribuyen a esta afirmación sobre que estos algoritmos poseen una naturaleza estocástica y no se puede lograr una completa explicación de modelo en algunos casos, lo cual representa un problema, ya que en el dominio financiero se necesita que las decisiones sean tomadas con fines éticos y regulatorios y muchas veces no es compatible con los requisitos regulatorios financieros (Dumitrescu et al., 2021). Por ejemplo, en el caso de Perú, se debe informar a la SBS la explicabilidad de algún modelo implementado. Bussmann et al. (2021) refuerzan indicando que la inteligencia artificial no es adecuada en servicios

financieros regulados, porque no se cumple que sea “explicable” e “interpretable”. El primer término corresponde a que se puedan responder preguntas sobre el funcionamiento de un modelo según el grupo de interés al que se dirigen; en cuanto al segundo término, corresponde a que la parte interesada pueda comprender los impulsores de la decisión tomada por el modelo. Además, el mismo *paper* señala que, si bien la IA mejora la conveniencia y accesibilidad de los servicios financieros, también se desencadenan nuevos riesgos. Los modelos de aprendizaje estadísticos “simples” pueden proporcionar alta interpretabilidad, pero limitada precisión predictiva. A diferencia de estos, los modelos de aprendizaje automático “complejos”, como las redes neuronales y los modelos de árbol, son totalmente opuestos: alta precisión predictiva, pero interpretación limitada. La misma conclusión pudieron obtener los mismos autores en la investigación realizada para fintechs (Bussmann et al., 2020). Además, otro enfoque para lograr la explicabilidad necesaria de los modelos se está optando por metodologías como SHAP (SHAPley Additive exPlanations), la cual es una herramienta de visualización usada para este objetivo. Si bien es entendida por científicos de datos o desarrolladores o validadores del modelo, con capacitación se puede lograr que otras personas relevantes y participantes del negocio, como auditores, organismos de supervisión o clientes, puedan comprender y mencionar qué variables son las más importantes para el resultado que se obtenga, y mostrar que el resultado no es de causalidad (Fritz-Morgenthal et al., 2022).

Por otro lado, se conoce que para el entrenamiento del modelo se requieren datos previos y etapas de selección de variables, análisis descriptivos, estimación de parámetros y selección del modelo adecuado (Babenko et al., 2021), lo cual es realizado por algún analista y esto podría caer en sesgos al seleccionar los datos (Leo et al., 2020), asimismo puede haber una falta de equidad, ya que existen estos sesgos externos que pueden instigar a los modelos de calificación crediticia (Mashrur et al., 2020). Por ejemplo, en esta misma investigación de Babenko et al., se consideraron a las variables independientes edad, sexo, estado civil, número de dependientes, ingreso mensual, residencia, tipo de apartamento, garantes, número de préstamos devueltos, y otros datos más del prestatario; la variable dependiente o “*target*” fue “*creditability*”, pero esto no indica que sean las variables que siempre se deban considerar en cualquier modelo, ya que es muy dependiente de los datos con los que se pueda contar. Adicionalmente, se requiere de disponibilidad de personal capacitado, ética en la recopilación, protección y transparencia de los datos.

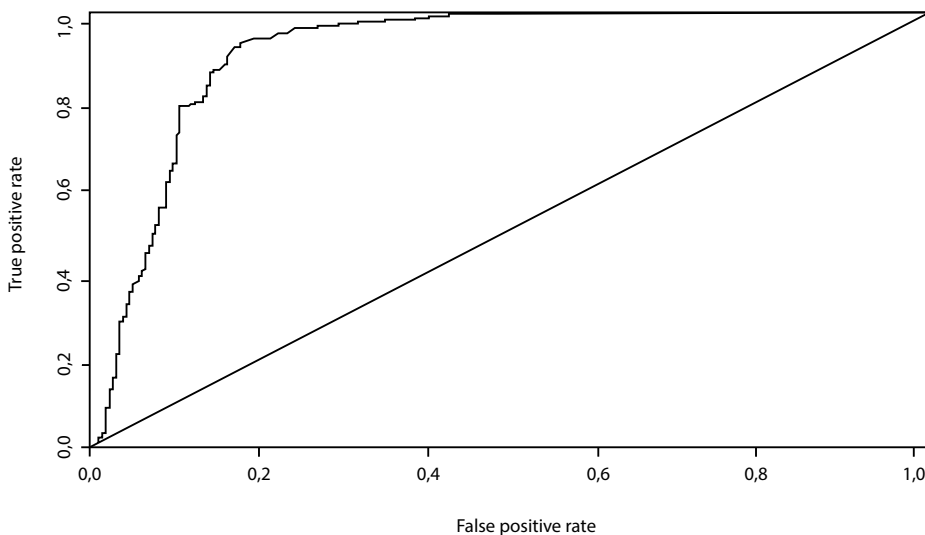
4.4 ¿Cómo identificar el algoritmo de *machine learning* más eficaz para la gestión de riesgo de crédito financiero?

Son diversos algoritmos de *machine learning* e identificar uno de estos como el más adecuado es una tarea complicada; sin embargo, se realizan diversas acciones y métricas para poder determinar si el modelo es adecuado. Para esta función, se toma en cuenta la matriz de confusión y, a partir de esta, poder determinar la sensibilidad y especificidad del modelo. Asimismo, se analiza la curva ROC, o AUC en inglés, para calcular el área bajo su curva.

El término “sensibilidad” se refiere a la proporción de predicción correcta de que cumplen con el caso (verdaderos positivos) sobre el número total de personas que cumplen el caso. Con respecto a la “especificidad”, se refiere a la proporción de predicción correcta de que no cumplen con el caso (verdaderos negativos) sobre el número total de personas que no cumplen el caso (Visa et al., 2011). Por otro lado, también se puede utilizar una curva ROC para obtener una determinación más rigurosa. El proceso para hacerlo es el siguiente: se deben tabular las sensibilidades y especificidades para diferentes valores de una medida de prueba continua. “Luego, la curva ROC gráfica se produce trazando la sensibilidad (tasa de verdaderos positivos) en el eje Y, contra la especificidad 1 (tasa de falsos positivos) en el eje X para los diversos valores tabulados” (Hui Hoo et al., 2017). El área bajo la curva ROC (AUC) es la capacidad de una prueba para discriminar si una condición específica está presente o no. Si el AUC es de 0,5, representa una prueba sin capacidad de discriminación. Si el AUC resulta de 1,0, representa una prueba con discriminación perfecta (Hui Hoo et al., 2017). Por ejemplo, en el artículo científico de Babenko et al. (2021) emplearon un modelo clásico de regresión lineal cuyas variables independientes fueron edad, sexo, estado civil, número de dependientes, ingreso mensual, residencia, tipo de apartamento, garantes, número de préstamos devueltos, y otros datos más del prestatario; la variable dependiente o “*target*” fue “*creditability*”. La sensibilidad entregada del modelo fue de 91 % y la especificidad del 80 %, lo cual significa que el 91 % de prestatarios confiables recibirán el préstamo (verdaderos positivos) y el 9 % restante son prestatarios no confiables que recibirán el préstamo (falsos positivos). Para evaluar la calidad del modelo se usó una matriz de clasificación y la curva ROC, la cual muestra la dependencia del número de resultados positivos correctamente clasificados del número de consecuencias negativas incorrectamente clasificadas. El área bajo la curva resultó 0,9, lo cual muestra alto poder predictivo y fiabilidad del modelo (Babenko et al., 2021).

Figura 1

Curva ROC para el modelo de puntuación de regresión logística



Nota. Área bajo la curva de la curva ROC para el modelo generado de regresión logística, cuyo resultado fue 0,9. De "Classical machine learning methods in economics research: Macro and micro level examples", por V. Babenko, A. Panchyshyn, L. Zomchak, M. Nehrey, Z. Artym-Drohomyretska, y T. Lahotskyi, 2021, *WSEAS Transactions on Business and Economics*, 18, p. 213 (<https://doi.org/10.37394/23207.2021.18.22>).

Sin embargo, determinar el algoritmo óptimo es muy variable, más incluso considerando los desafíos que se mostraron en la pregunta anterior, ya que no solo depende del nivel de sensibilidad, especificidad o el área bajo la curva ROC. Algunos de estos modelos de ML ya están interactuando con los tradicionales. Por ejemplo, en el caso de Dumitrescu et al. (2021), implementaron un algoritmo basado en un método de calificación crediticia que sea interpretable y de alto rendimiento al cual denominan regresión de árbol logístico penalizado (PLRT), donde concluyen que, mediante una simulación de Montecarlo, el modelo PLRT predice de manera significativamente más precisa que la regresión logística, regresión logística no lineal, regresión logística no lineal con lazo adaptativo y *support vector machine* se comparan a un nivel similar con *random forest*.

Otro ejemplo, es el artículo "Explainable machine learning in credit risk management", en el que se propone un modelo de inteligencia artificial basado en redes de correlación a los valores de Shapley que puede ser utilizado en la gestión de riesgo crediticio y, en particular, midiendo riesgos cuando ocurre que el crédito es tomado empleando plataformas de préstamos entre pares. Para este objetivo, se tomó una muestra de 15 mil pequeñas y medianas empresas (riesgosos y no riesgosos) que solicitaron crédito, los

cuales van a poder agruparse de acuerdo con sus características financieras similares y así otorgar puntajes crediticios y predecir comportamientos futuros. La principal contribución es mejorar los valores de Shapley partiendo de la interpretación del resultado predictivo de un modelo de aprendizaje automático mediante modelos de red de correlación.

De acuerdo con los artículos científicos mencionados anteriormente, se puede determinar que los algoritmos de *machine learning* basados en modelos estadísticos “simples” como la regresión logística son mejores que los métodos tradicionales, como el análisis de discriminante lineal, que se aplicaban en el pasado; asimismo, los algoritmos más “complejos” actuales representan superioridad en cuanto a mejor predicción de las variables a predecir: riesgoso o no riesgoso. Sin embargo, estos tipos de modelos son catalogados como una “caja negra”; es decir, no se sabe con exactitud cómo ocurre el proceso de selección a través de clasificación o “clusterización”. Debido a estos problemas de explicabilidad e interpretabilidad, se ha optado por el desarrollo de algoritmos que representan un híbrido para que se puedan cumplir (los estándares de regulación a nivel financiero).

A continuación, se muestran los posibles eventos que conllevan la implementación de los diversos algoritmos tradicionales y de *machine learning*, mencionados a través de la revisión sistemática de literatura, y los riesgos que pueden representar:

Tabla 2

Ventajas y desventajas de cada algoritmo

Técnica	Ventaja	Desventaja
Estadística tradicional (cartera atrasa y refinanciada, morosidad según tipo de crédito, cobertura de atrasos, cobertura de morosidad)	La capacitación del personal es menos especializada en algoritmos y se basa en conocimientos estadísticos.	Puede haber subjetividad o sesgos en la recopilación de datos.
Análisis de discriminante lineal (LDA)	Se considera como método tradicional a pesar de ser un algoritmo de clasificación, ya que lleva tiempo utilizándose y no presenta mucha incertidumbre (Mashrur et al., 2020).	La eficiencia y eficacia del modelo está limitada y se puede presentar subjetividad al momento de seleccionar los datos históricos (Mashrur et al., 2020).

(continúa)

(continuación)

Regresión logística	Es el algoritmo que suele ser más empleado actualmente, ya que se considera como transparente y no identificado como “caja negra”, debido a que no tiene la característica de alta complejidad (Leo et al., 2020).	Queda superada por algoritmos más complejos como SMV, Naive Bayes o random forest (Bussmann et al., 2020, Explainable AI in fintech risk management).
Support vector machine	Tiene éxito en la clasificación e incluso son competitivos en el descubrimiento de características significantes para determinar el riesgo de incumplimiento (Leo et al., 2020).	Es considerado de naturaleza estocástica, ya que existe incertidumbre en el modelo (Mashrur et al., 2020). Su funcionamiento empeora a medida que hay un aumento de datos de ingreso (Teng & Lee, 2020).
Árboles de decisión (decision tree)	Posee una precisión (accuracy en inglés) regular, mejor que los métodos tradicionales o los mencionados anteriormente, y logra tener una explicabilidad ante organismos reguladores (Dumitrescu et al., 2021). Obtuvo mayor precisión y funcionó más rápido frente a SVM y KNN (Teng & Lee, 2020).	Solo se ha demostrado que es superior a los anteriores considerando la métrica de accuracy (Teng & Lee, 2020).
Random forest	Posee mejor rendimiento que la regresión logística estándar, ya que se adapta mejor a efectos de umbral univariados y multivariados no observados (Dumitrescu et al., 2021). Se beneficia de la partición recursiva subyacente de los árboles de decisión (Dumitrescu et al., 2021).	Posee problemas de explicabilidad del modelo y requiere de un conjunto de datos que no sea pequeño e incluso puede caer en un sobreajuste del modelo (Mashrur et al., 2020).
Regresión de árbol logístico penalizado (PLRT)	Es más preciso que regresión logística, support vector machine y a un nivel similar que random forest (Dumitrescu et al., 2021). Es interpretable y de alto rendimiento, lo cual podría cumplir los requisitos reguladores financieros (Dumitrescu et al., 2021).	Requiere un alto grado de especialización para la utilización del algoritmo. Además, no es estándar, sino que es una propuesta de los autores, por lo cual se tendría que analizar en cada caso particular (Dumitrescu et al., 2021).
Redes neuronales	Tiene la capacidad propia de encontrar patrones y no se necesita de etiquetas previas (Mashrur et al., 2020).	El modelo puede sufrir dificultades de generalización y con ello producir modelos donde exista sobreajuste. (Zhang, 2020, p. 617). Representa el nivel más alto para considerarse como “caja negra”, además de un nivel de especialización mucho mayor que en los algoritmos anteriores, ya que la dificultad es alta (Mashrur et al., 2020).

5. CONCLUSIONES

Los algoritmos de *machine learning* se llevan utilizando un periodo de tiempo considerable, por lo cual es notoria la evolución que han tenido, asimismo ha habido un aumento importante en la gestión de riesgos de cualquier entidad que se dedique al rubro financiero, pero ha existido un mayor énfasis en la gestión de riesgo de crédito, debido a que representa el mayor riesgo, ya que evita pérdidas y ganancias en las organizaciones. Asimismo, se ha podido determinar que el *machine learning* representa ventajas frente a los métodos tradicionales, ya que obtienen mejor éxito para clasificar a los potenciales clientes. Sin embargo, es importante señalar que algunos de estos modelos no son fáciles de explicar en cuanto a su funcionamiento, por lo que se les etiqueta como “caja negra” y eso representa una desventaja para su implementación, ya que primero debe lograrse un visado por parte del sistema regulatorio financiero, además de necesitar personal especializado, ética en la recopilación, protección, transparencia y tratamiento de datos; sin embargo, estos problemas también podrían ocurrir en el cálculo tradicional. A pesar de esto, se han comenzado a desarrollar algoritmos “híbridos” entre modelos complejos y regresión logística para que se puedan cumplir los estándares regulatorios que se requieren en el rubro financiero, por lo que se puede observar una tendencia positiva en la adopción de modelos. El que obtuvo mejor rendimiento en casos aislados fue la propuesta híbrida de regresión de árbol logístico penalizado, pero, como se detalla en el cuadro, se necesita capacitación de personal que maneje algoritmos de *machine learning*, ya que es una propuesta nueva; cabe mencionar que esta superioridad fue mostrada con comparaciones que se realizaron para casos aislados, por lo que se recomienda tomarlo como una referencia teórica. Asimismo, es importante que estos modelos sean entrenados no solo con un conjunto de datos, sino que se deben usar múltiples veces para ilustrar la solidez frente a cualquier caso que se presente. Finalmente, se mostraron las métricas y métodos para poder determinar si un algoritmo es óptimo y se recomienda utilizar estas herramientas como corroboración. De igual manera, es importante para la organización o entidad poder analizar el mejor algoritmo a implementar de acuerdo con sus características propias, los datos con los que se cuente y en nivel de explicabilidad que se pueda brindar.

REFERENCIAS

- Babenko, V., Panchyshyn, A., Zomchak, L., Nehrey, M., Artym-Drohomyretska, Z., & Lahotskyi, T. (2021). Classical machine learning methods in economics research: Macro and micro level examples. *WSEAS Transactions on Business and Economics*, 18, 209-2017. <https://doi.org/10.37394/23207.2021.18.22>

- Bank for International Settlements. (2005). *An explanatory note on the Basel II IRB Risk Weight Functions*. Basel Committee on Banking Supervision. <https://www.bis.org/bcbs/irbriskweight.htm>
- BBVA Corporation. (2019, 6 de febrero). Forbes destaca los usos que BBVA hace de la inteligencia artificial. BBVA. <https://www.bbva.com/es/forbes-destaca-los-usos-que-bbva-hace-de-la-inteligencia-artificial/>
- Bhatore, S., Mohan, L., & Reddy, R. (2020). *Machine learning techniques for credit risk evaluation: a systematic literature review*. *Journal of Banking and Financial Technology*, 4, 111-138. <https://doi.org/10.1007/s42786-020-00020-3>
- Bussmann, N., Giudici, P., Marinelli, D., & Papenbrock, J. (2020). Explainable AI in fintech risk management. *Frontiers in Artificial Intelligence*, 1-5. <https://doi.org/10.3389/frai.2020.00026>
- Bussmann, N., Giudici, P., Marinelli, D., & Papenbrock, J. (2021). Explainable machine learning in credit risk management. *Computational Economics*, 57(1), 203-216. <https://doi.org/10.1007/s10614-020-10042-0>
- Castelvecchi, D. (2016). Can we open the black box of AI? *Nature*, 538, 20-23. <https://doi.org/10.1038/538020a>
- Columbus, L. (2021, 17 de enero). 76% of enterprises prioritize AI & machine learning in 2021 IT budgets. *Forbes*. <https://www.forbes.com/sites/louis-columbus/2021/01/17/76-of-enterprises-prioritize-ai--machine-learning-in-2021-it-budgets/?sh=6f9c20c618a3>
- Cui, H. (2021). Research on Credit Risk Control of Commercial Banks Based on Data Mining Technology. *2021 International Conference on Computer, Blockchain and Financial Development*, (pp. 383-386). <https://doi.org/10.1109/CBFD52659.2021.00083>
- Cunningham, P., Cord, M., & Delany, S. J. (2008). Supervised learning. En M. Cord y P. Cunningham (Eds.), *Machine learning techniques for multimedia. Cognitive technologies* (pp. 21-49). https://doi.org/10.1007/978-3-540-75171-7_2
- Dietterich, T. (1997). Machine-learning research. *AI Magazine*, 18(4), 97-136. <https://doi.org/10.1609/aimag.v18i4.1324>
- Dumitrescu, E., Hué, S., Hurlin, C., & Tokpavi, S. (2021). Machine learning for credit scoring: Improving logistic regression with non-linear decision-tree effects. *European Journal of Operational Research*, 297(3), 1178-1192. <https://doi.org/10.1016/j.ejor.2021.06.053>
- Elizondo, A. (2003). *Medición integral del riesgo de crédito*. Limusa.

- Enterprise.nxt Staff. (2021, 19 de agosto). The rise of artificial intelligence and machine learning. *Hewlett Packard Enterprise*. <https://www.hpe.com/us/en/insights/articles/the-rise-of-artificial-intelligence-and-machine-learning-2108.html>
- Fritz-Morgenthal, S., Hein, B., & Papenbrock, J. (2022). Financial risk management and explainable, trustworthy, responsible AI. *Frontiers in Artificial Intelligence*, 5, 1-14. <https://doi.org/10.3389/frai.2022.779799>
- Ghahramani, Z. (2004). Unsupervised learning. En O. Bousquet, U. von Luxburg, y G. Rätsch (Eds.), *Advanced lectures on machine learning. Lecture Notes in Computer Science* (vol. 3176, pp. 72-112). Springer. https://doi.org/10.1007/978-3-540-28650-9_5
- Giacometti, R., Torri, G., Farina, G., & De Giuli, M. (2020). Risk attribution and interconnectedness in the EU via CDS data. *Computational Management Science*, 17, 549-567. <https://doi.org/10.1007/s10287-020-00385-2>
- Hui Hoo, Z., Candlish, J., & Teare, D. (2017). What is an ROC curve? *Emergency Medicine Journal*, 34, 357-359. <http://dx.doi.org/10.1136/emmermed-2017-206735>
- IBM. (2018, 26 de diciembre). ¿Qué es *machine learning*? <https://www.ibm.com/pe-es/analytics/machine-learning>
- KPMG Company. (2021, mayo). With asset quality on the brink, supervisors' focus on credit risk is growing. <https://home.kpmg/xx/en/home/insights/2021/05/with-asset-quality-on-the-brink-supervisors-focus-on-credit-risk-is-growing.html>
- Leo, M., Sharma, S., & Maddulety, K. (2020). Machine learning in banking risk management: A literature review. *Risks*, 7(1), 1-22. <https://doi.org/10.3390/risks7010029>
- Mashrur, A., Luo, W., Zaidi, N., & Robles-Kelly, A. (2020). Machine learning for financial risk management: A survey. *IEEE Access*, 8, 203203-203223. <https://doi.org/10.1109/ACCESS.2020.3036322>
- Milojević, N., & Redzepagic, S. (2021). Prospects of artificial intelligence and machine learning application in banking risk management. *Journal of Central Banking Theory and Practice*, 10(3), 41-57. <https://doi.org/10.2478/jcbtp-2021-0023>
- Moreno, A., Armengol, E., Béjar, J., Belanche, L., Cortés, U., Gavaldà, R., Gimeno, J. M., López, B., Martín, M. & Sánchez, M. (1994). *Aprendizaje automático*. Edicions UPC.
- Nina, H., Pow-Sang, J. A., & Villavicencio, M. (2021). Systematic mapping of the literature on secure software development. *IEEE Access*, 9, 36852-36867. <https://doi.org/10.1109/ACCESS.2021.3062388>
- Plawiak, P., Abdar, M., Plawiak, J., Makarenkov, V., & Acharya, U. (2020). DGHNL: A new deep genetic hierarchical network of learners for prediction of credit scoring. *Information Sciences*, 516, 401-418. <https://doi.org/10.1016/j.ins.2019.12.045>

- SAS: Analytics, Artificial Intelligence and Data Management. (2017). Machine learning and artificial intelligence in a brave new world. https://www.sas.com/en_au/insights/articles/analytics/machine-learning-and-artificial-intelligence-in-a-brave-new-world.html
- SAS: Analytics, Artificial Intelligence and Data Management. (2015). Credit risk management - What it is and why it matters. Risk Insights. https://www.sas.com/el_gr/insights/risk-management/credit-risk-management.html
- Srinivasa Rao, M., Sekhar, C., & Bhattacharyya, D. (2021). Comparative Analysis of Machine Learning Models on Loan Risk Analysis. *Advances in Intelligent Systems and Computing*, 1280, 81–90. https://doi.org/10.1007/978-981-15-9516-5_7
- SuperintendenciadeBanca,SegurosyAFP.(s.f.).Métodosbasadosencalificacionesinternas. https://www.sbs.gob.pe/Portals/0/jer/pres_doc_basilea/III%20METODO%20IRB.pdf
- Teng, H.-W., & Lee, M. (2020). Estimation procedures of using five alternative machine learning methods for predicting credit card default. En C. Lee, & J. Lee, *Handbook of financial econometrics, mathematics, statistics, and machine learning* (vol. 4, pp. 3545-3572). https://doi.org/10.1142/9789811202391_0101
- Visa, S., Ramsay, B., Ralescu, A., & van der Knaap, E. (2011). Confusion matrix-based feature selection. *Proceedings of the 22nd Midwest Artificial Intelligence and Cognitive Science Conference 2011*, 120-127.
- Zhang, X. D. (2020). *A matrix algebra approach to artificial intelligence*. Springer.
- Zhang, Y., (Ed.). (2010). *New Advances in Machine Learning*. IntechOpen. <https://doi.org/10.5772/225>

DATOS DE LOS AUTORES

ÁLVARO ANTONIO AURES GARCÍA

Ingeniero de Computación y Sistemas. Magíster en Ciencia de la Información y doctor en Ingeniería de Sistemas. Exmiembro del equipo de los proyectos de Incubadoras de Empresas de Base Tecnológica en el Technologie Park Herzogenrath (Alemania), Gestión del Conocimiento en la Universidad Málaga (España), MétróMéta de la Universidad de Montreal (Canadá) y Procesamiento de Imágenes Médicas (Brasil). Estudios de especialización en Gestión de la Innovación Tecnológica (Fraunhofer Institute), Gestión del Conocimiento y Tecnologías de la Información en países como España, Alemania, Canadá, Brasil y Uruguay. Responsable de la gestión del conocimiento en el sector público. Docente en universidades peruanas.

ADELAIDA ESTELA BALVÍN LANDEO

Ingeniera empresarial y de sistemas por la Universidad San Ignacio de Loyola. Máster en Dirección de Sistemas por la Universidad de Barcelona de España. Especialista en tecnologías de la información. Interés por la gestión del conocimiento en las organizaciones.

FRANCI SUNI-LÓPEZ

Docente e investigador de la Universidad de Lima. Candidato a doctor en Ciencias de la Computación por la Universidad Nacional de San Agustín de Arequipa. Realizó un *joint master degree* en Ciencia de la Computación en la Universidad Libre de Ámsterdam (Holanda) y en la Universidad Católica San Pablo. Su línea de investigación se centra

en el campo de la computación afectiva, la cual procesa las emociones humanas para potenciar la capacidad de autoadaptación de los servicios de *software*. Asimismo, tiene interés en las áreas de ingeniería de *software* para la sostenibilidad y la inteligencia artificial aplicada al procesamiento de imágenes y señales.

HÉCTOR EDUARDO UGARTE ROJAS

Máster en Ciencias de la Computación por la Rheinische Friedrich-Wilhelms-Universität Bonn (Alemania). Ingeniero informático y de sistemas graduado de la Universidad Nacional de San Antonio Abad del Cusco (Perú). Su interés de investigación incluye cadenas de bloque, web semántica, *linked data* y *distributed ledgers*.

BORIS CHULLO LLAVE

Magíster en Ciencia de la Computación por el Instituto de Matemáticas y Estadística de la Universidad de São Paulo (USP) (Brasil). Interesado en trabajar en áreas como optimización, algoritmos, investigación operativa y teoría de la computación.

JOSE MARTIN LOZANO APARICIO

Doctor en Informática por la Universidad de Lille, profesor auxiliar en la Facultad de Ciencias de la Universidad Nacional de Ingeniería. Interesado en las ontologías, optimización de consultas, bases de datos de grafos, procesamiento de lenguaje natural y *data exchange*.

JOSE MALPARTIDA VALVERDE

Estudiante de la Facultad de Ciencias, en la especialidad de Ciencia de la Computación, en la Universidad Nacional de Ingeniería. Interesado en el ambiente virtual (RA y RV), las interfaces con interacciones no convencionales y la robótica.

VÍCTOR AUGUSTO ALEGRE IBÁÑEZ

Estudiante de la Facultad de Ciencias de la Universidad Nacional de Ingeniería. Sus intereses son el campo de la ciencia de datos y aprendizaje profundo.

LUIS BELTRÁN PALMA TTITO

Maestro en Ciencias Mención Informática por la Universidad Nacional de San Antonio Abad del Cusco. Docente ordinario del Departamento Académico de Ingeniería Informática de la Universidad Nacional de San Antonio Abad del Cusco. Sus intereses son algoritmos genéticos, *machine learning* y *deep learning*.

GANIMEDES T. ROSALES REYES

Magíster en Administración de Negocios, es ingeniero industrial, consultor independiente y docente universitario. Más de catorce años de experiencia en posiciones ejecutivas y de dirección, realizando y liderando actividades vinculadas a la formulación y ejecución de estrategias, mejora de rendimiento organizacional e implementación de tecnologías de la información y comunicaciones (TIC), tanto en el sector público como privado. Posee habilidades para coordinar y comunicar resultados al más alto nivel, así como para conformar y liderar equipos multidisciplinarios. Tiene experiencia en los campos de dirección y mejora de rendimiento organizacional, incluyendo el diseño de modelos de negocios para gestionar oficinas de proyectos (PMO) y oficinas de gestión de la excelencia (*business process office*), así como el diseño e implementación de arquitecturas empresariales y de procesos, sistemas de gestión de mejora continua bajo estándares internacionales. Cuenta con certificaciones PMP (*project management professional*), ITILf v3, Lean Six Sigma Black Belt - ASQ y especializaciones en las normas ISO9001.

JULER ANDERSON HERMITAÑO CASTRO

Estudiante de la carrera de Ingeniería de Sistemas en la Universidad de Lima. Intereses en la investigación y en el desarrollo de *machine learning* e inteligencia artificial.



UNIVERSIDAD
DE LIMA

Propuesta metodológica para
la identificación de expertos
en el ámbito organizacional

A Distributed Model for Computing
3D Mesh Local Descriptors
Based on k -Rings

Football Pitch Condition
Analysis Based on k -Means
Clustering

Estudio de la usabilidad de una
aplicación de realidad aumentada
en la construcción básica de un
brazo robótico

Aplicación de métodos de deep
learning en series de tiempo
para el pronóstico de la situación
macroeconómica en América Latina

Clasificación de organismos en
reinos utilizando frecuencia de
codones de ADN

Ciencia de datos en la evaluación
del impacto de las políticas públicas:
una revisión de la literatura

Aplicación de machine learning
en la gestión de riesgo de crédito
financiero: una revisión sistemática